

Characterising Exoplanets with *TESS* and HARPS

by

Ares Osborn

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

October 2023

THE UNIVERSITY OF
WARWICK

Contents

List of Tables	iv
List of Figures	v
Acknowledgments	ix
Declarations	xi
Abstract	xii
Abbreviations	xiii
Chapter 1 Introduction	1
1.1 A brief history of exoplanets	1
1.2 Exoplanet detection	3
1.2.1 The transit method	4
1.2.2 The radial velocity method	14
1.2.3 Further methods	19
1.3 Exoplanet characterisation	26
1.3.1 Determination of system parameters	26
1.3.2 Inference of further characteristics	26
1.3.3 Types of exoplanets	31
1.4 The exoplanet population	37
1.4.1 Population features	37
1.4.2 Occurrence rates	44
1.4.3 Formation and evolution	47
1.5 Thesis outline	51
Chapter 2 Methods	52
2.1 Transit detection with <i>TESS</i>	52

2.1.1	About <i>TESS</i>	52
2.1.2	Confirmation of transit signals	57
2.2	Radial velocity follow-up with HARPS	59
2.2.1	About HARPS	59
2.2.2	Scheduling a HARPS follow-up program	60
2.2.3	Confirmation of radial velocity signals	65
2.3	Joint fit modelling of exoplanet systems	72
2.3.1	Using exoplanet	72
2.3.2	Bayes and MCMC	73

Chapter 3 TOI-431/HIP 26013: a super-Earth and a sub-Neptune transiting a bright, early K dwarf, with a third RV planet

		75
3.1	Introduction	76
3.2	Observations	77
3.2.1	Photometry	79
3.2.2	Spectroscopy	81
3.2.3	High resolution imaging	83
3.2.4	Stellar analysis	87
3.2.5	Stellar activity monitoring	92
3.3	The joint fit	94
3.3.1	The third planet found in the HARPS data	94
3.3.2	Construction of the joint fit model	98
3.3.3	Fit results	107
3.4	Discussion	110
3.5	Conclusion	114

Chapter 4 Nomads: uncovering the origin of remnant planets in the hot Neptunian desert

		115
4.1	Abstract	116
4.2	Introduction	116
4.3	The sample	117
4.3.1	Stage 1: cuts based on desert boundaries and observability	121
4.3.2	Stage 2: manual vetting and ranking	123
4.3.3	Stage 3: the final sample	129
4.4	The observations	146
4.5	First-look analysis of target data	148
4.5.1	Systems already published - TOIs 332, 908, 2498	151
4.5.2	Systems that are in-prep - TOIs 3071, 5559, 1117, 2411	155

4.5.3	Another system with a less than 20 per cent error on mass - TOI-5005157	
4.5.4	Systems that may show a planet - TOIs 2358, 426, 1839	160
4.5.5	Systems that are double-lined spectroscopic binaries - TOIs 2539, 2224	173
4.5.6	Systems that show a null result - TOIs 422, 570, 855, 728, 355, 1975, 4537, 2365, 502, 835, 499, 2227	173
4.6	Discussion	208
4.6.1	Outcomes of the observation sample	208
4.6.2	Published planets in the sample	208
4.7	Future outlook	213

Chapter 5 TOI-332 b: a super dense Neptune found deep within the Neptunian

desert	214
5.1	Introduction 215
5.2	Observations 218
5.2.1	Photometry 218
5.2.2	Spectroscopy 224
5.2.3	High resolution imaging 229
5.3	Spectroscopic analysis and chemical abundances 231
5.4	The joint fit 234
5.5	Results and discussion 238
5.5.1	Interior structure 238
5.5.2	Co-orbital bodies 239
5.5.3	Evolution under XUV-driven escape 242
5.5.4	Other formation scenarios 244
5.5.5	Orbital decay rate 244
5.5.6	Future observation prospects 245
5.6	Conclusion 246

Chapter 6 Conclusion 247

6.1	Thesis summary 247
6.2	Future outlook 248

List of Tables

1.1	Introduction: planet categories by radius and mass	32
3.1	TOI-431: system details	78
3.2	TOI-431: HARPS spectroscopy data	82
3.3	TOI-431: HIRES spectroscopy data	82
3.4	TOI-431: stellar parameters	89
3.5	TOI-431: joint fit model prior distributions	100
3.6	TOI-431: planet parameters	109
3.7	TOI-431: interior structure models	112
4.1	Nomads: relevant spectroscopic disposition tags explained	118
4.2	Nomads: SG1 Disposition tags explained	120
4.3	Nomads: TOIs after basic cuts on desert boundaries and observability . . .	132
4.4	Nomads: final sub-sample of already published targets/known nulls	142
4.5	Nomads: final sub-sample of TOIs to observe	145
4.6	Nomads: target visibilities and observations made	147
5.1	TOI-332: system details	219
5.2	TOI-332: HARPS radial velocity data	229
5.3	TOI-332: stellar chemical abundances	232
5.4	TOI-332: joint fit model prior distributions	235
5.5	TOI-332: stellar and planetary parameters	237
5.6	TOI-332: co-orbital hypothesis results	241

List of Figures

1.1	Introduction: cumulative exoplanet detections per year	2
1.2	Introduction: known exoplanet periods versus radii and masses	3
1.3	Introduction: idealised exoplanet transit schematic	5
1.4	Introduction: edge-on exoplanet transit schematic showing trigonometry to determine impact parameter	7
1.5	Introduction: observer's view exoplanet transit schematic showing geometry between the first and last transit contact points	7
1.6	Introduction: 3D exoplanet transit schematic showing geometry between the first and last transit contact points	8
1.7	Introduction: transits of HD 209458 b at different wavelengths, demonstrat- ing the effect of limb darkening	10
1.8	Introduction: the first full transits observed of HD 209458 b	13
1.9	Introduction: the transit light curve of the first exoplanet detected via transits, OGLE-TR-56 b	13
1.10	Introduction: idealised radial velocity schematic	15
1.11	Introduction: the effect of eccentricity and argument of periastron on the shape of a radial velocity curve	16
1.12	Introduction: the radial velocity curve of the first exoplanet discovery, 51 Peg b	18
1.13	Introduction: transit timing variations of Kepler-9 b and c	22
1.14	Introduction: direct imaging of the four planets in the HR 8799 system . . .	23
1.15	Introduction: light curve of the microlensing event MOA-2022-BLG-249 .	24
1.16	Introduction: the astrometric motion of DE0823-49 caused by a brown dwarf	25
1.17	Introduction: mass-radius plot showing exoplanet compositions	28
1.18	Introduction: the Rossiter-McLaughlin effect for several orbital configurations	30
1.19	Introduction: the Neptunian desert in period-radius and period-mass space .	38
1.20	Introduction: the 2017 discovery of the radius gap	42
1.21	Introduction: the period-radius distribution showing the radius valley	43
1.22	Introduction: the negative slope of the radius valley in period-radius space .	43

1.23	Introduction: occurrence rates of different sizes of planets from <i>Kepler</i> (2013)	45
1.24	Introduction: occurrence rates of planets from <i>Kepler</i> in period-radius space (2019)	46
2.1	Methods: a diagram of the <i>TESS</i> camera and sector geometry.	54
2.2	Methods: a diagram of all sectors viewed by <i>TESS</i> up to the end of cycle 5.	56
2.3	Methods: a diagram of the equatorial coordinate system.	61
2.4	Methods: activity on the surface of the Sun	68
2.5	Methods: diagrams of the stellar activity indicators	70
3.1	TOI-431: contrast curves and reconstructed images from high-resolution imaging	84
3.2	TOI-431: an example of the blended source confidence analysis	88
3.3	TOI-431: spectral energy distribution	93
3.4	TOI-431: periodogram of WASP-South data	94
3.5	TOI-431: periodograms of HARPS RV data, sequentially removing contributions from the planets and stellar activity	95
3.6	TOI-431: periodograms of 2019 HARPS stellar activity indicators and window functions	96
3.7	TOI-431: periodograms of 2004-2015 HARPS stellar activity indicators and window functions	97
3.8	TOI-431: <i>TESS</i> data, with subsequent detrending and best fit model	102
3.9	TOI-431: LCO and NGTS data, with subsequent best fit models	103
3.10	TOI-431: Spitzer data, with subsequent pixel level decorrelation and best fit model	104
3.11	TOI-431: radial velocity data, with subsequent detrending and best fit models	106
3.12	TOI-431: radius valley plot	110
3.13	TOI-431: mass-radius plot	113
4.1	Nomads: all released TOIs within and surrounding the Neptunian desert	119
4.2	Nomads: TOIs after basic cuts on desert boundaries and observability	122
4.3	Nomads: the 272 TOIs showing merit function	126
4.4	Nomads: the 272 TOIs showing ranking dispositions	128
4.5	Nomads: the final full sample	130
4.6	Nomads: final sub-sample of TOIs to observe	131
4.7	Nomads: CCFs for all observed targets	149
4.8	Nomads: TOI-908, periodograms and fit to the RV data	153
4.9	Nomads: TOI-2498, periodograms and fit to the RV data	154

4.10	Nomads: TOI-5005.01, RV data, periodogram, and simple Keplerian fit . .	158
4.11	Nomads: TOI-5005.01, activity indicators	159
4.12	Nomads: TOI-2358.01, RV data, periodogram, and simple Keplerian fit . .	161
4.13	Nomads: TOI-2358.01, RV data and fit including activity	162
4.14	Nomads: TOI-2358.01, activity indicators	163
4.15	Nomads: TOI-426.01, RV data, periodogram, and drift	165
4.16	Nomads: TOI-426.01, RV data and fit	166
4.17	Nomads: TOI-426.01, activity indicator correlations	167
4.18	Nomads: TOI-426.01, activity indicator periodograms	168
4.19	Nomads: TOI-1839.01, RV data and periodogram	170
4.20	Nomads: TOI-1839.01, RV data and activity fit	171
4.21	Nomads: TOI-1839.01, activity indicators	172
4.22	Nomads: TOI-422.01, RV data, periodogram, and simple Keplerian fit . . .	174
4.23	Nomads: TOI-422.01, activity indicators	175
4.24	Nomads: TOI-570.01, RV data, periodogram, and simple Keplerian fit . . .	176
4.25	Nomads: TOI-570.01, activity indicators	177
4.26	Nomads: TOI-855.01, RV data, periodogram, drift, and simple Keplerian fit	179
4.27	Nomads: TOI-855.01, activity indicator correlations	180
4.28	Nomads: TOI-855.01, activity indicator periodograms	181
4.29	Nomads: TOI-728.01, RV data, periodogram, and simple Keplerian fit . . .	182
4.30	Nomads: TOI-728.01, activity indicators	183
4.31	Nomads: TOI-355.01, RV data, periodogram, drift, and simple Keplerian fit	185
4.32	Nomads: TOI-355.01, activity indicator correlations	186
4.33	Nomads: TOI-355.01, activity indicator periodograms	187
4.34	Nomads: TOI-1975.01, RV data, periodogram, and simple Keplerian fit . .	188
4.35	Nomads: TOI-1975.01, activity indicators	189
4.36	Nomads: TOI-4461.01, RV data, periodogram, and simple Keplerian fit . .	191
4.37	Nomads: TOI-4461.01, activity indicators	192
4.38	Nomads: TOI-4537.01, RV data, periodogram, and simple Keplerian fit . .	193
4.39	Nomads: TOI-4537.01, activity indicators	194
4.40	Nomads: TOI-2365.01, RV data, periodogram, and simple Keplerian fit . .	196
4.41	Nomads: TOI-2365.01, activity indicators	197
4.42	Nomads: TOI-502.01, RV data, periodogram, drift, and simple Keplerian fit	198
4.43	Nomads: TOI-502.01, activity indicator correlations	199
4.44	Nomads: TOI-502.01, activity indicator periodograms	200
4.45	Nomads: TOI-835.01, RV data, periodogram, and simple Keplerian fit . . .	202
4.46	Nomads: TOI-835.01, activity indicators	203

4.47	Nomads: TOI-499.01, RV data, periodogram, and simple Keplerian fit . . .	204
4.48	Nomads: TOI-499.01, activity indicators	205
4.49	Nomads: TOI-2227.01, RV data, periodogram, and simple Keplerian fit . .	206
4.50	Nomads: TOI-2227.01, activity indicators	207
4.51	Nomads: observation sample outcomes	209
4.52	Nomads: Neptunian desert plots of published planets in sample	211
4.53	Nomads: mass-radius plot of published planets in sample	212
5.1	TOI-332: Neptunian desert plot	217
5.2	TOI-332: target pixel file from TESS	220
5.3	TOI-332: <i>TESS</i> , LCO, and HARPS data, with subsequent best fit model . .	221
5.4	TOI-332: all six LCO transits, with subsequent best fit model	223
5.5	TOI-332: PEST transit light curve	224
5.6	TOI-332: periodograms of WASP-South data	225
5.7	TOI-332: WASP transit light curve	226
5.8	TOI-332: HARPS radial velocities versus stellar activity indicators	227
5.9	TOI-332: periodograms for the HARPS radial velocities and stellar activity indicators	228
5.10	TOI-332: reconstructed images and contrast curves from high-resolution imaging	230
5.11	TOI-332: spectral energy distribution	233
5.12	TOI-332: mass-radius plot	239
5.13	TOI-332: stellar rotation period and XUV luminosity over time	243
5.14	TOI-332: planet radius and envelope mass fraction over time	243

Acknowledgments

Wow. Despite everything (looking at you, global pandemic, and you, health issues, and you, growing hostility towards transgender people in the UK), I made it? But I did not do this alone, so here we go...

First, to my supervisor, Dave Armstrong: your guidance and support and the frankly mind-boggling amount of knowledge and expertise that you have shared has made me the astronomer I am today. Thank you for answering all my stupid questions, thank you for putting up with me constantly taking on too many responsibilities (a lot of them nothing to do with my PhD), and thank you for being kind about the various health issues I always seem to be plagued by. Oh, and thank you for always getting my pronouns right.

To my masters supervisor, Dan Bayliss: I may have already thanked you in my previous thesis, but I just have to do so again, because it was your decision to take me on as a student and your support and patience throughout that year that has led to me being here.

To the Astronomy & Astrophysics group at Warwick, both my fellow students and the members of staff: thank you for making this perhaps one of the best places on Earth to do an astronomy PhD. As well as the absolute wealth of science that everyone gets up to, there has been no shortage of kindness, support, and fun in my time working with you all. Team garlic forever.

To George King and Finnbar Keating: you are the best of friends anyone could ever ask for and then some. You've both supported me through thick and thin, and I love you so much. George, thank you for answering all my weird little astronomy questions, even when I asked them previously and promptly forgot the answer. Finnbar, sorry for scarring you forever by showing you some of my horrendous code.

To my Chillies: you are such a creative group of absolute babes and chatting to y'all

always brightens my day. You are the best little community I have ever been a part of.

To Alyx: life can be rough, but it always feels a little easier when I'm with you. Thank you for being a safe space to be myself, even when I'm a little bit weird. Love you big. Please give Max a kissie for me.

Finally, to my family, in particular my mum and dad: thank you so much for letting me move back home and looking after me while I wrote this thesis. It was always going to be a bit of a slog, but you lifted some of that burden for me and I am so grateful. You may have teased me about getting a proper job sometimes, but your support throughout my PhD has been unwavering. I hope I've done you proud. All my love.

Trans rights are human rights.

Declarations

I declare that the work included in this thesis is my own unless stated otherwise, and the specific contributions to the work by collaborators are declared at the start of each chapter. The availability of the data used in this thesis is also declared at the start of each chapter. This thesis has not been submitted to this, or any other academic institution, for any other degree or qualification. Work in this thesis has been published in peer reviewed journals, including Chapter 3, which is a reproduction of Osborn et al. (2021a), and Chapter 5, which is a reproduction of Osborn et al. (2023).

Abstract

In the scant few decades since the discovery of the first exoplanet, the existence of over 5000 exoplanets has been confirmed. However, there are still many open questions about the formation, migration, and evolution pathways that sculpt the known population. Intriguing population features, such as the Neptunian desert and the radius valley, aid us in unravelling these scenarios. However, studying these requires many exoplanet discoveries with precisely determined radii and masses. In this thesis, I primarily use photometry from the Transiting Exoplanet Survey Satellite (*TESS*) and follow-up spectroscopic observations from the High Accuracy Radial velocity Planet Searcher (HARPS) to characterise transiting exoplanets that are in or near these two population features, and thus contribute to our understanding of the mechanisms behind their existence.

I first present the discovery and characterisation of the three planet system around TOI-431 using multiple sources of photometry and spectroscopy, included but not limited to *TESS*, *Spitzer*, HARPS and HIRES. I carefully disentangle the stellar activity signal from the radial velocity data to do so. The planets are, in order of increasing distance from the star, an ultra-short period super-Earth, a non-transiting likely super-Earth, and a warm sub-Neptune. The two transiting planets straddle the radius gap and provide an interesting case study for differing atmospheric evolution above and below the gap.

In order to make inferences about a population of planets, they need to be studied in a homogeneous and statistically significant way. As a first step towards this goal, I create a sample of *TESS* Objects of Interest within the Neptunian desert. The sample totals 73 planet candidates, 26 of which do not have high precision radial velocity observations. I perform these observations within a large HARPS programme, dubbed “Nomads”, and then perform preliminary analysis of the RV data to determine whether or not the candidates are planetary. I finally examine the sample within the context of the Neptunian desert and by their compositions.

A discovery from the Nomads programme, I present the characterisation of an unusually massive planet with an ultra-short period, TOI-332 b. It is one of the densest Neptunes discovered thus far, and joins the very small handful planets discovered within the “deep” Neptunian desert. With a likely negligible hydrogen-helium envelope despite its size, it requires additional explanation beyond the core-accretion theory of planetary formation.

Abbreviations

Units

M_{\oplus}	Earth mass, 5.972×10^{24} kg	au	Astronomical unit, 150×10^6 km
R_{\oplus}	Earth radius, 6 378 km	BJD(-TDB)	Barycentric Julian Date (-Barycentric Dynamical Time)
M_J	Jupiter mass, 1.898×10^{27} kg	K	Kelvin
R_J	Jupiter radius, 71 492 km	mas	Milliarcsecond
M_{\odot}	Solar mass, 1.989×10^{30} kg	pc	Parsec
R_{\odot}	Solar radius, 695 700 km		

Telescopes, instruments, facilities

CTIO	Cerro Tololo Inter-American Observatory	LCOGT	Las Cumbres Observatory Global Tele- scope Network
ESO	European Southern Observatory	NGTS	Next Generation Transit Survey
ESPRESSO	Echelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations	PEST	Perth Exoplanet Survey Telescope
HARPS	High Accuracy Radial velocity Planet Searcher	SAAO	South African Astronomical Observatory
JWST	JW Space Telescope	TESS	Transiting Exoplanet Survey Satellite
		WASP	Wide Angle Search for Planets

Data and analysis

CCF Cross Correlation Function	PSF Point Spread Function
FAP False Alarm Probability	SNR Signal-to-Noise Ratio
FFI Full Frame Image	TIC <i>TESS</i> Input Catalogue
HMC Hamiltonian Monte Carlo	TOI <i>TESS</i> Object of Interest
MCMC Markov-Chain Monte Carlo	

Chapter 1

Introduction

1.1 A brief history of exoplanets

We have speculated about planets outside our own Solar System as early as the sixteenth century: Giordano Bruno, an Italian philosopher, posited that stars in the sky might be like our own Sun, with their own planets (Bruno, 1584). Isaac Newton expresses his religious views in his 1726 General Scholium; within this he thinks about how other stars may be the centres of other systems of planets and comets like ours, and how this connects to the existence of God (Newton, 1726).

It took until the twentieth century for methods to be proposed to detect these planets around stars other than our Sun (Struve, 1952) and it is only within the past three decades or so that any have actually been discovered. This led to the coining of the word “exoplanet” to describe an extra-solar planet.

The first confirmed exoplanet discovery was made by Wolszczan and Frail (1992): two planets were discovered orbiting the millisecond pulsar PSR B1257+12, with a third found shortly after (Wolszczan, 1994); the planets were detected via the timing shift they caused in its pulses.

Prior to this, another candidate had been proposed and was later confirmed. Campbell et al. (1988) observed a radial velocity signal from Gamma Cephei which they suggested was due to a Jupiter-sized companion, which was subsequently doubted (Walker et al., 1992), but finally confirmed to be due to a planetary companion in 2003 (Hatzes et al., 2003). Interestingly, there was evidence of exoplanets even earlier than this, from calcium absorption in the spectra of the white dwarf Van Maanen 2 (van Maanen, 1917) implying that planetary material was falling onto its surface, but this was not realised until 2016 (Farihi, 2016).

Notably, the first exoplanet discovered to orbit a main-sequence star (51 Pegasi) was found by Mayor and Queloz (1995), and was the first confirmed planet to be found by the

radial velocity technique. Mayor & Queloz were awarded the Nobel Prize in 2019 for this work.

Since these early discoveries, the field of exoplanet science has seen an almost exponential growth of discoveries (Fig. 1.1), surpassing 5000 in early 2022, and the accompanying science on how they form and evolve, what they're made of, and whether they could host life, has similarly expanded in scope. This introduction will provide a brief overview on what can, and has, filled several textbooks (see e.g., Haswell, 2010; Perryman, 2018; Deeg and Belmonte, 2018).

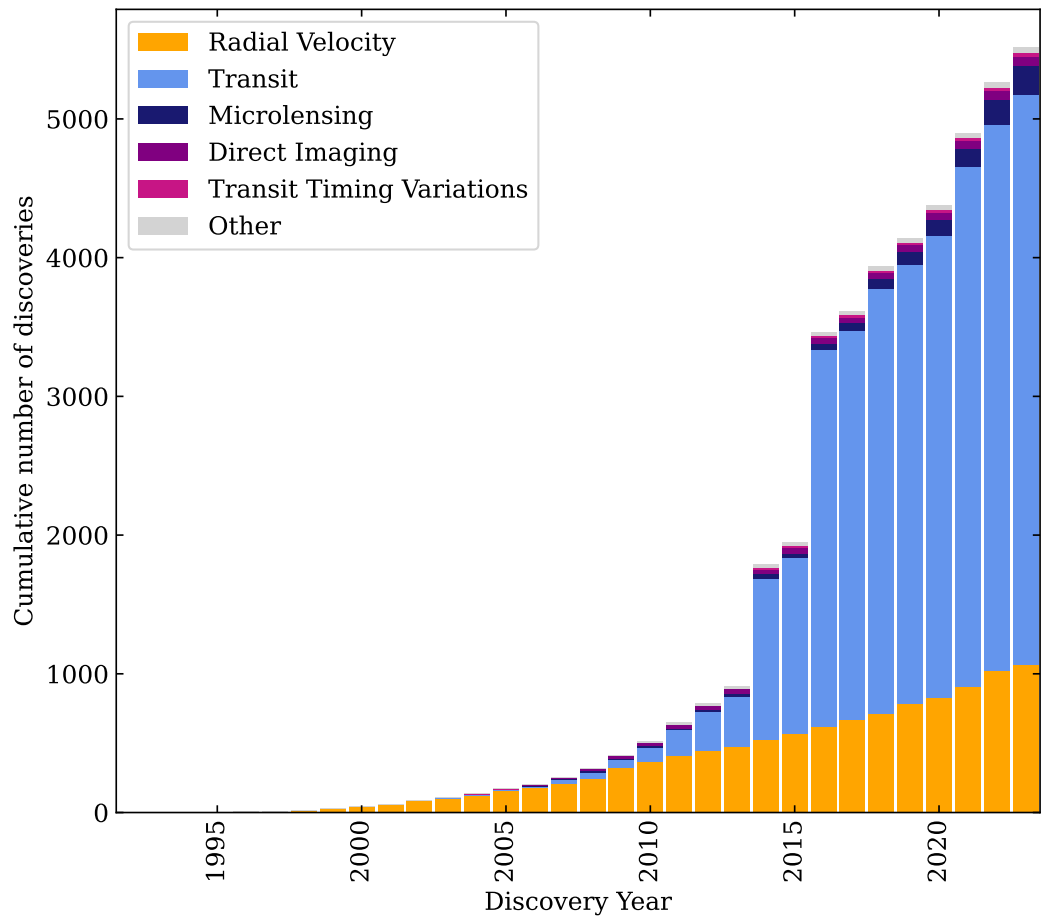


Figure 1.1: The cumulative number of exoplanet detections per year since the first confirmed detection in 1992, colour-coded by detection method. The jumps in 2014 and 2016 are due to large numbers of exoplanets from the *Kepler* mission becoming statistically validated by Rowe et al. (2014) and Morton et al. (2016), respectively. Data taken from the NASA exoplanet archive (<https://exoplanetarchive.ipac.caltech.edu/>) on 16 Sep 2023.

1.2 Exoplanet detection

Due to the difficulty associated with imaging an exoplanet directly (see Section 1.2.3), most exoplanet detections have been made through “indirect” methods, inferring the existence of a planet from observable effects it induces on its host star.

To date, the two most prolific methods by far have been the transit and radial velocity methods. They are both utilised extensively in this thesis, and here I will describe in detail their underlying principles, and briefly summarise several other detection methods. These methods can be thought of as complementary; they are each sensitive to a different region of exoplanetary parameter space (i.e., they find planets over different ranges of period, radius, and mass) as shown in Fig. 1.2, allowing exoplanet astronomers to find a huge variety of exoplanetary systems (see Section 1.3.3).

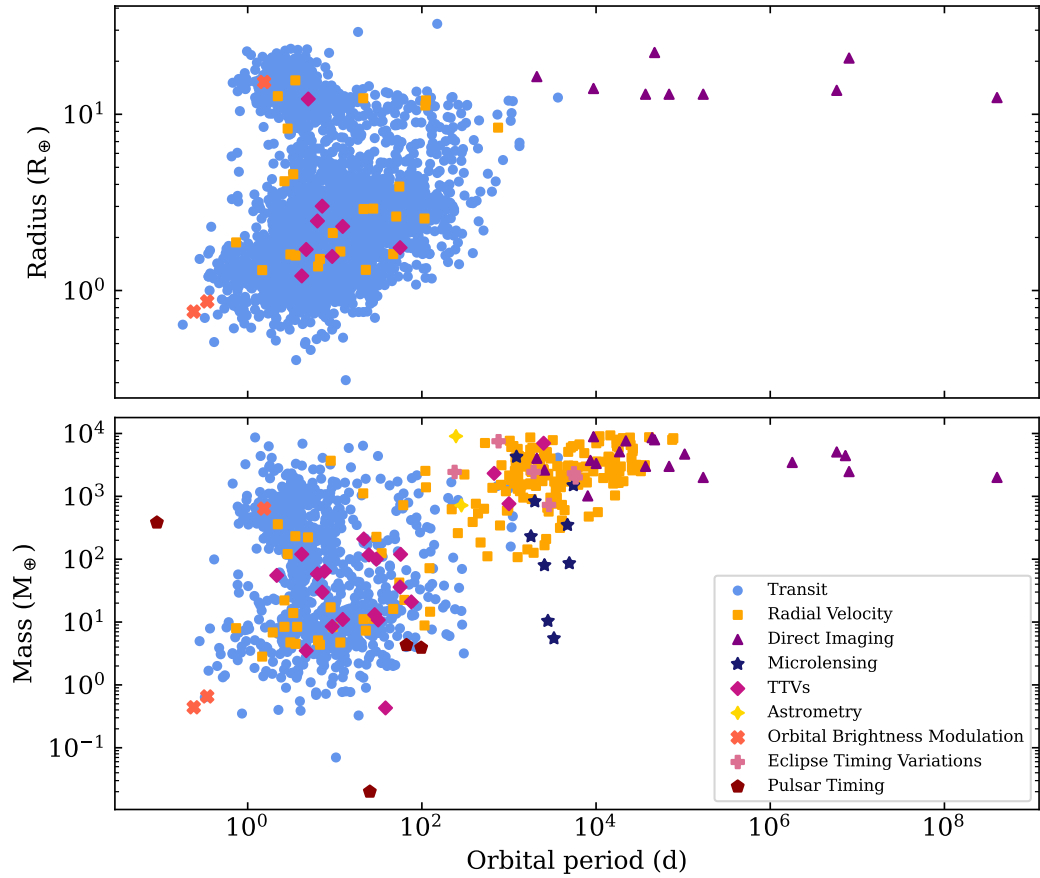


Figure 1.2: Known exoplanet orbital periods plotted against their radii (top) and masses (bottom), where the markers represent their discovery methods, showing the regions the detection methods are sensitive to. Data taken from the NASA exoplanet archive (<https://exoplanetarchive.ipac.caltech.edu/>) on 16 Sep 2023.

1.2.1 The transit method

The theory

The transit method relies on a favourable alignment geometry: that is, the orbital plane of the planet must align with the line of sight of the viewer to the star. Then, at a particular time in its orbit, the planet will pass through the viewer’s line of sight, blocking some of the stellar light that they would otherwise see and causing a dip in the apparent brightness of the host star. This effect will repeat every orbital period.

By obtaining a light curve, a measure of the brightness of a star over time, at the time of the occultation of the star by the planet, a substantial amount can be inferred about the orbital configuration of the system and the properties of the planet. This is illustrated in Fig. 1.3, and explained below for the simple case of a circular orbit. Fig. 1.4 through 1.6 and Equ. 1.1 through 1.10 have been adapted from Haswell (2010).

First, the orbital period of the planet, P , can be determined from the time between multiple transits. This can then be related to the semi-major axis of the orbit, a , through Kepler’s third law:

$$P^2 = \frac{4\pi^2 a^3}{G(M_\star + M_p)}. \quad (1.1)$$

This can be further simplified under the approximation that $M_\star + M_p \approx M_\star$. Use of this relation requires prior information about the star, its mass, which can be derived from stellar models (see Sections 3.2.4 and 5.3 for some methods in which this can be determined). When using stellar models, the planetary parameters are affected by uncertainties in the models. There are some alternative methods to derive stellar parameters, such as via *Gaia* data (Fouesneau et al., 2023) and asteroseismology (e.g., Rodrigues et al., 2017).

Next, the depth of the transit, δ , can be determined from the observed change in flux, ΔF , relative to the out-of-transit flux, F_0 . Assuming a stellar disc of uniform brightness, and neglecting flux from the planet, the depth of the transit can then be geometrically related to the relative radii of the planet, R_p , and star, R_\star , as follows:

$$\frac{\Delta F}{F_0} \equiv \delta = \left(\frac{R_p}{R_\star} \right)^2. \quad (1.2)$$

The depth will be in the range $0 < \delta < 1$: $\delta = 0$ is equivalent to no transit (a 0 per cent transit depth); $\delta = 1$ is equivalent to a transit that blocks out the whole star (a 100 per cent transit depth); and, more realistically, $\delta \approx 0.01$ is equivalent to an approximately Jupiter-sized planet transiting a Sun-like star (a ~ 1 per cent transit depth). To find an Earth-sized planet around a Sun-like star, a transit depth of ~ 0.008 per cent needs to be detectable.

Again, this requires prior information about the star; here, its radius, which can

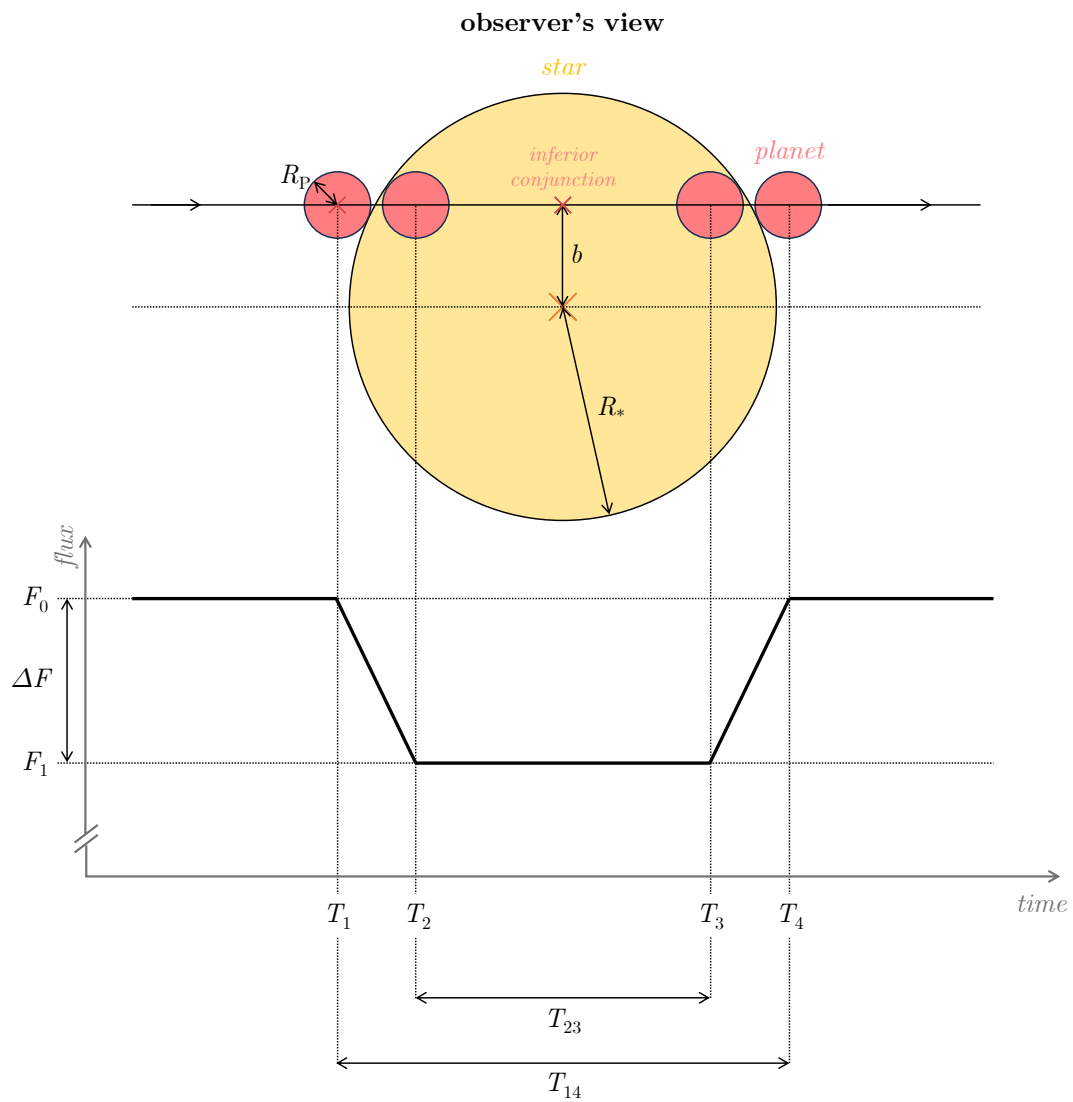


Figure 1.3: An idealised schematic of a transit from the observer's view (top) and the resultant light curve (bottom).

be determined via the methods described above for stellar mass, and I again emphasise that uncertainties from models will propagate forward into the determination of planetary parameters.

Another quantity that can be obtained from the light curve is the transit duration. As illustrated in Fig. 1.3, a transit has four chronological contact points which correspond to observable features in the transit light curve. First contact, T_1 , occurs when the limb (edge) of the planet first “touches” the limb of the star (from the viewpoint of the observer). Second contact, T_2 , occurs the moment the entire planetary disc is overlapping the stellar disc. The planet is then transiting across the stellar disc, and hits third contact, T_3 , the last instant where the entire planetary disc is overlapping the stellar disc. Fourth contact, T_4 , happens the moment the planetary disc no longer overlaps with the stellar disc. As such, two kinds of transit duration can be defined: the time the planet takes to move from contact point 1 to 4 known as the “total transit duration”, T_{14} , and the time from contact point 2 to 3 known as the “full-transit duration”, T_{23} (not confusing at all). T_{14} is used below for calculation of planetary parameters, and is also useful to know when scheduling transit observations; T_{23} is utilised less, but can, for example, be used in calculation of the signal-to-noise ratio of a transit (see Kipping, 2023).

The transit duration depends on the impact parameter, b , the vertical distance between the centre of the planet and the centre of the star at mid-transit (inferior conjunction) as viewed by the observer (illustrated in Fig. 1.4). Via trigonometry, the impact parameter can be related to the semi major axis and the inclination of the system, i , the angle between the planet’s orbital plane and the normal to the observer on Earth ($i = 0$ deg for a face on orbit, for which a transit would not be seen; $i = 90$ deg for an edge-on orbit):

$$b = a \cos i. \quad (1.3)$$

The geometry of the system can be examined as the planet moves from T_1 to T_4 (illustrated in Fig. 1.5). A right-angled triangle can be made out of the impact parameter, the sum of the stellar and planetary radii, and a third side of length l . From Pythagoras’ theorem, it can be said that:

$$l = \sqrt{(R_\star + R_p)^2 - (a \cos i)^2}. \quad (1.4)$$

Finally, the length, l , can be related to the total transition duration, T_{14} . As the planet moves from T_1 to T_4 , it sweeps an angle of α around the centre of the star (see Fig. 1.6). Making a triangle from T_1 to T_4 (noting that this is equivalent to $2l$) to the centre of the star, again trigonometry can be used to determine that:

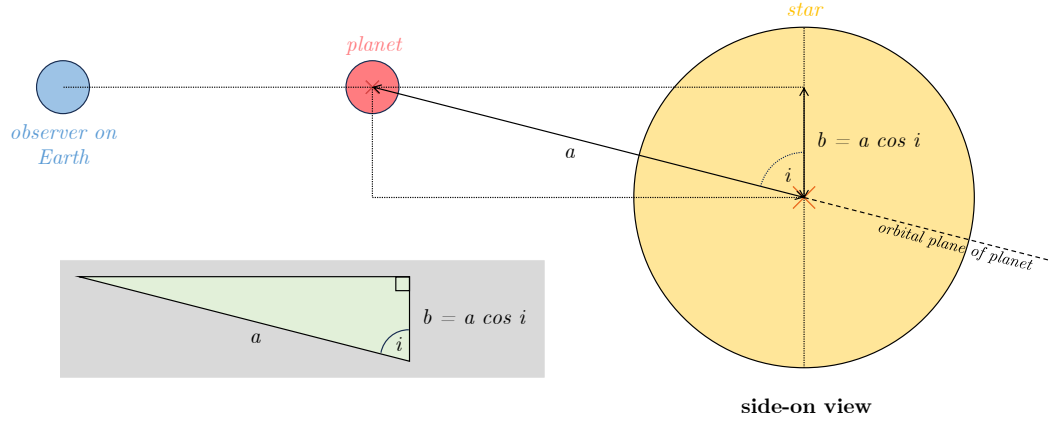


Figure 1.4: An edge-on schematic showing the trigonometry that determines the impact parameter, Equ. 1.3.

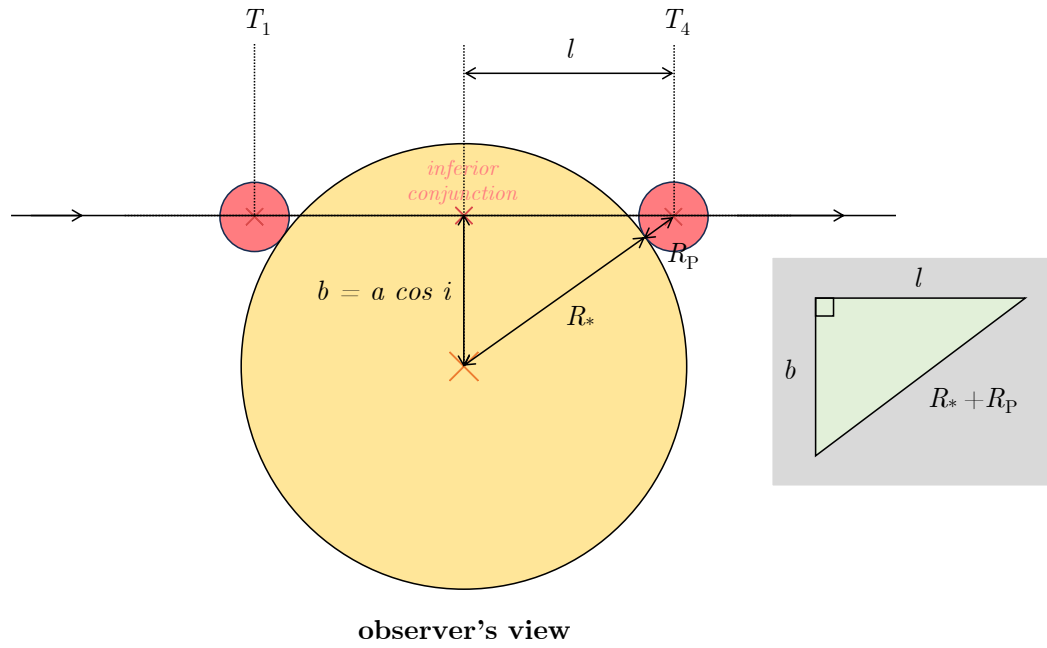


Figure 1.5: An observer's view schematic of the geometry created between the planet at T_1 and T_4 , from which the length, l , can be derived (Equ. 1.4).

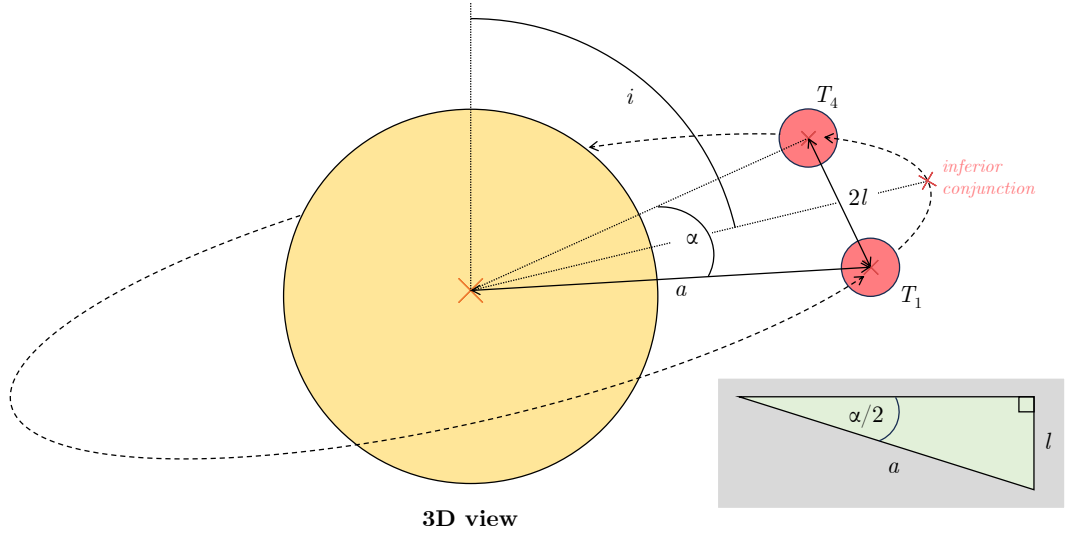


Figure 1.6: A 3D schematic of a planetary orbit showing the angle swept out by the planet moving from T_1 to T_4 and how this relates to l (Equ. 1.5).

$$\sin(\alpha/2) = \frac{l}{a}. \quad (1.5)$$

For a circular orbit, with α in radians, the total transit duration is:

$$T_{14} = P \frac{\alpha}{2\pi}. \quad (1.6)$$

Substituting α from Equ. 1.5 results in:

$$T_{14} = \frac{P}{\pi} \sin^{-1} \left(\frac{l}{a} \right). \quad (1.7)$$

Finally, substituting l from Equ. 1.4:

$$T_{14} = \frac{P}{\pi} \sin^{-1} \left(\frac{\sqrt{(R_{\star} + R_p)^2 - (a \cos i)^2}}{a} \right). \quad (1.8)$$

As T_{14} can be measured from the light curve, and how to determine P , R_p , and a is known from earlier, Equ. 1.3 and Equ. 1.8 can be used to determine b and i .

Unfortunately, reality is not quite so simple as the model derivations above. So far, circular orbits have been assumed, but planets are perfectly capable of having eccentric (elliptical) orbits. A more complex treatment including eccentricity can be found in Winn (2010).

Additionally, the stellar disc does not have uniform brightness, as assumed. Instead, it is brighter at the centre and becomes smoothly darker and redder towards the limb, an

effect called “limb darkening”, a consequence of the optical depth increasing towards the limb. Limb darkening results in a transit light curve having a rounded, rather than flat, bottom, and the effect it has varies depending on the wavelength a star is observed at. This is demonstrated in Fig. 1.7, where transits of HD 209458 b were observed by the Hubble Space Telescope at many different wavelengths (Knutson et al., 2007). The rounding of the bottom of the transit becomes more pronounced (more “u” shaped) the further towards blue wavelengths one goes, and the transit appears deeper. Consequently, it is important to model the effect of limb darkening when fitting transit light curves, as otherwise the depth of the transit and thus the radius of the planet might be misrepresented (Espinoza and Jordán, 2015). Limb darkening can be modelled with a range of “limb darkening laws”, including but not limited to linear and quadratic relations (Csizmadia, 2018). A frequently employed method of doing so is by using the parameterisation of Kipping (2013), which models physically realistic limb-darkening without requiring prior knowledge of the star.

Whether or not a transit will be seen in the first place is, as stated at the beginning of this section, reliant on favourable alignment geometry: the orbital plane of the planet must align with the line of sight of the viewer to the star, such that the disc of the planet transits the disc of the star. When the planet is closest to the observer (“inferior conjunction”), it is also as close as it will get to the centre of the stellar disc, and the distance between the centres of the two discs is simply the impact parameter, $a \cos i$ (see Fig. 1.4). Therefore, for the disc of the planet to transit the disc of the star, the inclination must satisfy:

$$a \cos i \leq R_{\star} + R_{\text{P}}. \quad (1.9)$$

In the case where $R_{\star} - R_{\text{P}} < a \cos i \leq R_{\star} + R_{\text{P}}$, a “grazing” transit is seen; that is, the disc of the planet only partially covers the disc of the star, and the depth of the transit will be less than expected from Equ. 1.2.

Equ. 1.9 can be simplified to infer a probability of a transit to be:

$$P_{\text{trans}} = \frac{R_{\star} + R_{\text{P}}}{a} \approx \frac{R_{\star}}{a}. \quad (1.10)$$

The probability of a transit occurring increases with decreasing semi-major axis, i.e. decreasing orbital period. Coupled with it being easier to infer the presence of a planet with multiple detected transits, this strongly biases the transit method towards close-in planets. Indeed, a detection is not claimed based on single (“mono”) transit, and even with two transits (a “duotransit”) the planet periodicity is not secure. Three or more transits are generally needed to determine planetary period, and work has been done to follow-up both mono and duotransit candidates (e.g., Gill et al., 2020; Lendl et al., 2020; Osborn, 2022; Osborn et al., 2022; Hawthorn et al., 2023a).

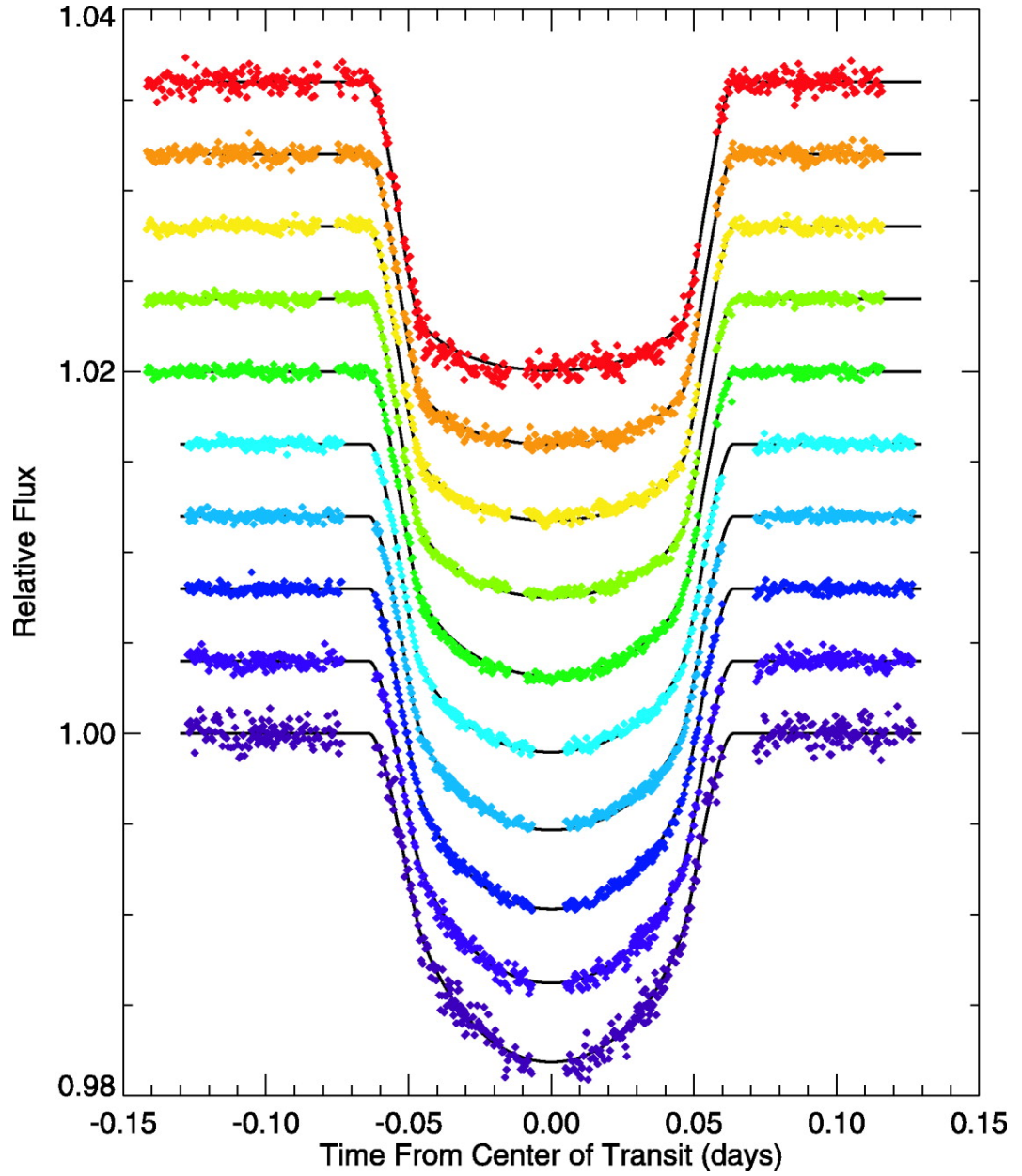


Figure 1.7: A demonstration of the effect of limb-darkening on the transit shape: HD 209458 b observed by the Hubble Space Telescope in 10 different bandpasses over a total wavelength range of 290 to 1030 nm. The bottom of the transit becomes noticeably more “u” shaped as the wavelength becomes bluer. Reproduced from Knutson et al. (2007).

Overall, the derivations above show that the transit method is a powerful tool, allowing the inference of many properties of an exoplanet detected with it, e.g., its period, radius, inclination, and impact parameter. Importantly, this is the only method that allows a direct determination of a planet’s radius. However, the major drawbacks are not being able to determine a planet’s mass (excepting the scenario where transit timing variations are present, see Section 1.2.3), which is necessary to confirm whether a signal is truly planetary in origin, and the low probability that a planet will transit from our viewpoint. There are also many astrophysical false positives that can mimic a transit signal (e.g., blended eclipsing binaries and grazing eclipsing binaries), so further information is often needed to rule these possibilities out.

Facilities and discoveries

The first planet found to transit was HD 209458 b, which had both radial velocity measurements and transits (Charbonneau et al., 2000; Henry et al., 2000), where the transits are shown in Fig. 1.8. The transits confirmed the radial velocity variations were planetary in origin. The first planet to be *discovered* via the transit method was OGLE-TR-56 b (see Fig. 1.9), which was confirmed as planetary with radial velocity measurements (Konacki et al., 2003). Both planets are hot Jupiters, giant planets on close-in orbits to their host stars (see Section 1.3.3).

In the two decades since the first transit discoveries, a little over 4000 planets have been discovered via the transit method¹. Transit detection facilities can be naturally split by whether they are ground-based or space-based.

Ground based facilities did the initial legwork in confirming the transit method was viable and continue to make new discoveries to this day, alongside the important work of following up candidates from space-based observatories. Early facilities include: the Hungarian Automated Telescope (HAT, Bakos et al., 2002); the Trans-Atlantic Exoplanet Survey (TrES, Alonso et al., 2004); XO (McCullough et al., 2005); the Wide Angle Search for Planets (WASP, Pollacco et al., 2006); the Kilodegree Extremely Little Telescope (KELT, Pepper et al., 2007); and HAT-South (Bakos et al., 2009). More recent facilities include: MEarth (Irwin et al., 2009); the Antarctic Search for Transiting ExoPlanets (ASTEP, Crouzet et al., 2010); the Transiting Planets and Planetesimals Small Telescope (TRAPPIST, Jehin et al., 2011); the Search for habitable Planets EClipping ULtra-cOOl Stars (SPECULOOS, Burdanov et al., 2018); and the Next Generation Transit Survey (NGTS, Wheatley et al., 2018). Of these, WASP has made the most discoveries, finding close to 200 planets¹ with its large field-of-view and short cadence. There are, however, significant challenges

¹As recorded by the NASA Exoplanet Archive, <https://exoplanetarchive.ipac.caltech.edu>, accessed 7 September 2023. All discovery numbers in the text in Section 1.2 refer to this footnote.

when observing transits from the ground, including weather, noise produced by the Earth’s atmosphere, and the limited night-time windows in which transits can be observed – and these can be solved by moving our telescopes to space.

The Convection Rotation Transits (CoRoT, Baglin et al., 2006) satellite was the first space-based transit detection facility; launched in 2006, it found over 30 transiting planets over its 7 years of operation¹. Then came *Kepler*, performing the initial *Kepler* mission (Borucki et al., 2010) and then the subsequent *K2* mission (Howell et al., 2014) after the loss of two of the spacecraft’s reaction wheels. It was groundbreaking; between both missions, over 3000 transiting exoplanets have been found¹ with candidate follow-up still ongoing to this day. Perhaps the biggest drawback of *Kepler/K2* is that the observed stars are often too dim for radial velocity follow-up, resulting in large numbers of planets without mass measurements, their planetary nature being “validated” rather than fully confirmed. Its successor, the Transiting Exoplanet Survey Satellite (TESS, Ricker et al., 2015) was designed to target bright stars more amenable to follow-up across the whole sky, and has discovered verging on 400 planets (with over 4000 candidates yet to be investigated)¹. The mission is currently ongoing.

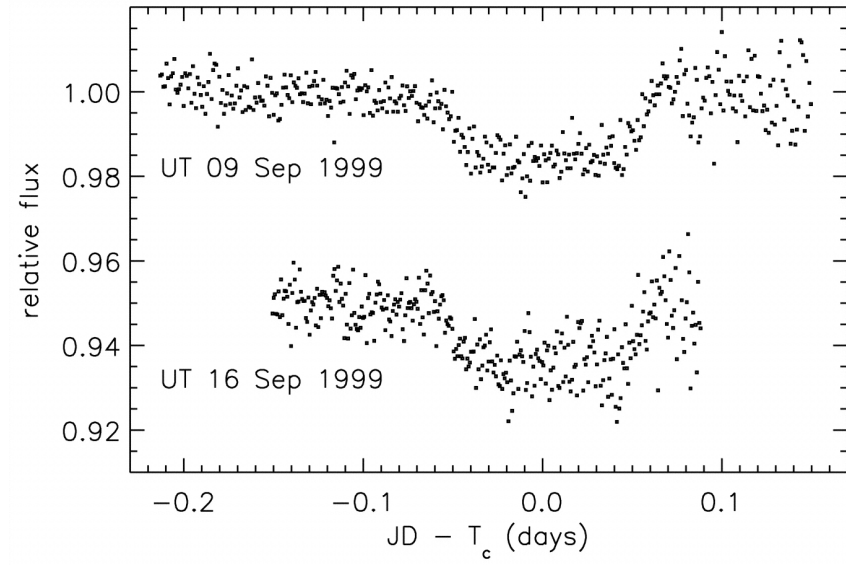


Figure 1.8: The first light curves taken encompassing the full transit of the planet HD 209458 b, reproduced from Charbonneau et al. (2000).

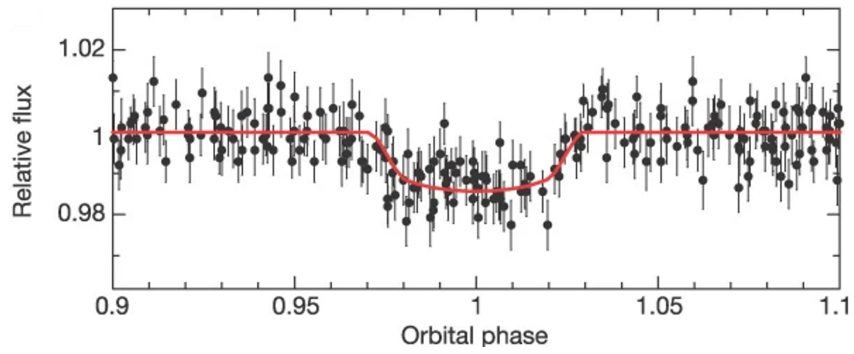


Figure 1.9: The light curve of the first planet to be discovered through the transit method, OGLE-TR-56 b, reproduced from Konacki et al. (2003).

1.2.2 The radial velocity method

The theory

In a star-planet system, the bodies orbit the “barycentre”, their common centre of mass. The mass of the planet is most likely small compared to the star, and thus the star has a relatively small orbit around the barycentre, but this motion still results in a periodic change to observable qualities of the star. A star’s radial velocity (RV), its velocity directly towards or away from a viewer, changes as it orbits the barycentre.

Stellar spectra can be taken multiple times over the full orbital period of the planet, and the position of the many absorption lines in the spectra will exhibit a Doppler shift depending on whether the star is moving away (red-shifted) or towards (blue-shifted) the viewer. Each spectra is cross-correlated against a template spectrum (where the spectrum is “at rest”) and the wavelength shift is converted to a velocity via:

$$v = \frac{\Delta\lambda}{\lambda}c, \quad (1.11)$$

where v is the line-of-sight velocity of the star, λ is the rest wavelength, $\Delta\lambda$ is the wavelength shift due to the Doppler effect, and c is the speed of light (Equ. 1.11 through 1.13 and derivations have been adapted from Perryman (2018)). Assuming a circular orbit, plotting velocities from several spectra covering a full planetary orbit will result in a sinusoid shape (see Fig. 1.10).

The semi-amplitude of this velocity sinusoid, K , can then be related to system parameters as follows (of which a full derivation can be found in Perryman (2018) as it is beyond the scope of this introduction):

$$K = \left(\frac{2\pi G}{P} \right)^{1/3} \frac{M_P \sin i}{(M_\star + M_P)^{2/3}} \frac{1}{(1 - e^2)^{1/2}}. \quad (1.12)$$

A value of K can be obtained from radial velocity data as shown in Fig. 1.10. An eccentricity value can also be determined from the radial velocity data; Fig. 1.11 shows how the eccentricity, e , and argument of periastron, ω , affect the shape of the obtained radial velocity curve.

Knowing K and e , and having a pre-determined value for M_\star , Equ. 1.12 shows that only a minimum mass, $M_P \sin i$, can be obtained, unless a value for inclination is known, which cannot be determined from radial velocity data (but can from transit data).

It can be seen that K will be larger when the planetary mass is larger (relative to the stellar mass), and with shorter orbital periods.

To estimate the radial velocity precision that is needed to reach to detect certain kinds of planets, one can assume a circular orbit where $M_P \ll M_\star$ and utilise useful units

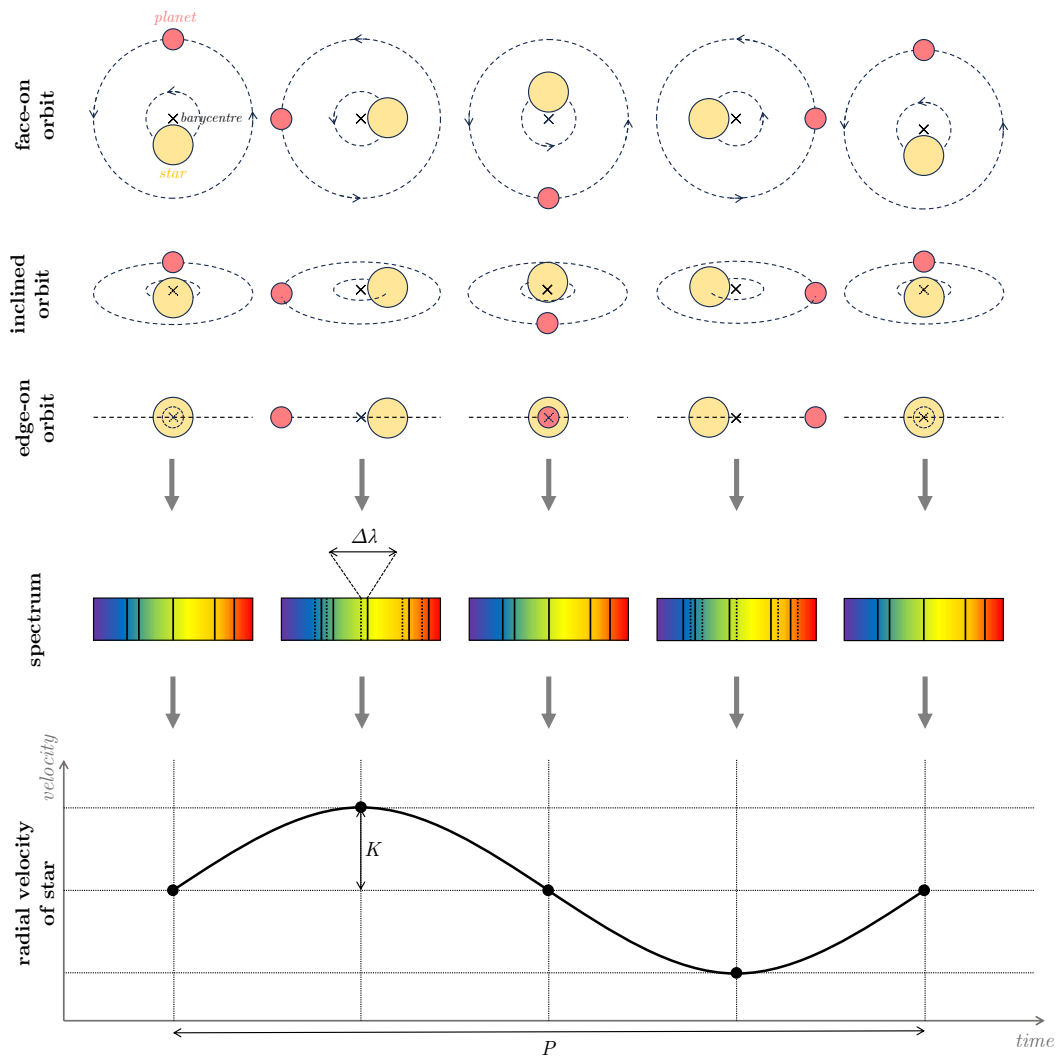


Figure 1.10: An idealised schematic of the radial velocity method, picturing one orbit of the planet going from left to right. Top: a planet-star system orbits its barycentre, and this motion on the sky is shown for different orbital orientations (labelled), where the observer is looking into the page. Middle: the velocity of the star towards and away from the observer causes spectral lines to be red- or blue-shifted compared to the at-rest spectrum. Bottom: the wavelength shift can be converted to a velocity. The wavelength shift and semi-amplitude will be maximum for an edge-on orbit, and zero for a face-on orbit. Reproduced into static format from a .gif created by Alysa Obertas at https://commons.wikimedia.org/wiki/File:Radial_velocity_doppler_spectroscopy.gif – the author encourages readers to view this .gif for a perhaps more intuitive view.

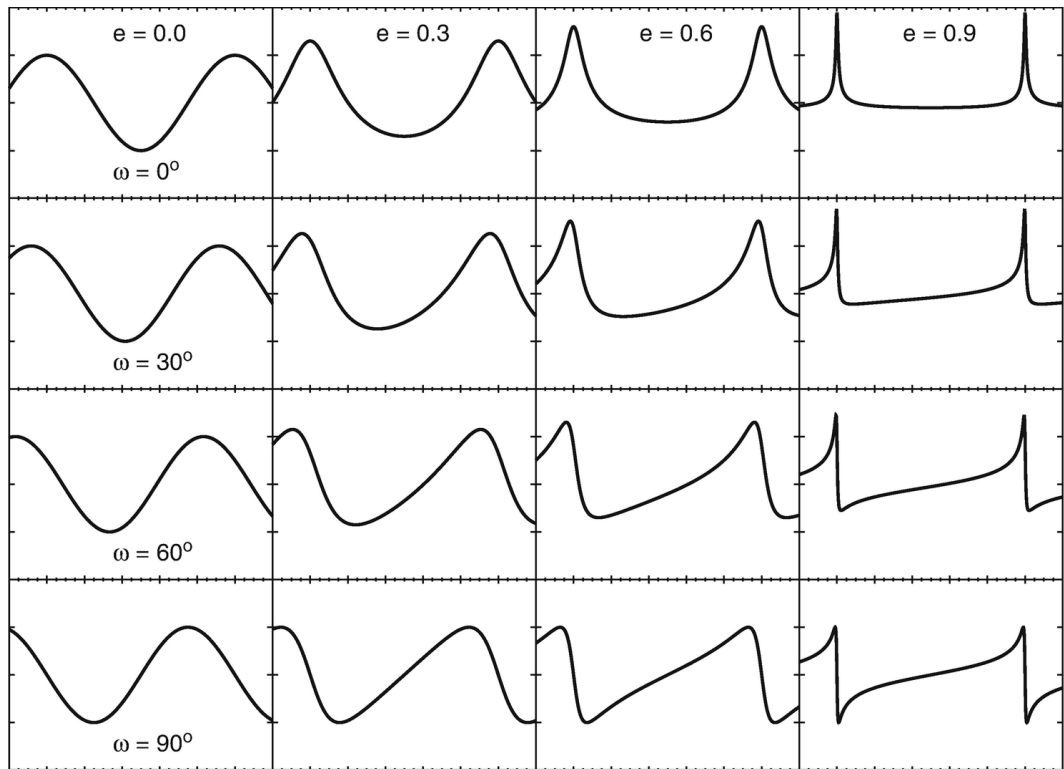


Figure 1.11: The effect of e and ω on the shape of a radial velocity curve, with each column showing constant e and each column showing constant ω . Reproduced from Wright and Gaudi (2013).

to simplify Equ. 1.12:

$$K = 28.4 \text{ms}^{-1} \left(\frac{P}{1 \text{yr}} \right)^{-1/3} \left(\frac{M_P \sin i}{M_J} \right) \left(\frac{M_\star}{M_\odot} \right)^{-2/3}, \quad (1.13)$$

where for a Jupiter around a Sun-like star, $K = 12.5 \text{ m s}^{-1}$, and for Earth, $K = 0.09 \text{ m s}^{-1}$. So, to detect Earth-like planets around Sun-like stars, a $\sim 10 \text{ cm}$ -level precision is needed in radial velocity data.

Overall, the radial velocity method allows determination of a planet's period, eccentricity, and minimum mass (when coupled with an inclination measurement from another method, the true mass can be obtained). It also allows exoplanets to be found at a wider range of inclinations than the transit method, though the signal will be strongest at $i = 90^\circ$. However, a planet's radius cannot be determined from this method – the transit and radial velocity methods can be seen now to be truly complimentary. When taken together, all important system properties can be determined, which allows inferences to be made about, for example, a planet's composition from its bulk density. Additionally, there are again false positives that can confuse the nature of a radial velocity signal, the biggest issue being stellar activity (see Section 2.2.3), so further information is again beneficial here.

Facilities and discoveries

The first discovery made via the radial velocity method was 51 Pegasi b (Mayor and Queloz, 1995), shown in Fig. 1.12; as mentioned earlier in Section 1.1, it was also the first planet discovered around a Sun-like star. Like the early transit discoveries, this is yet another hot Jupiter. In the years since, about 1000 planets have been discovered by the radial velocity method¹.

The CORrelation-RAdial-VELOCities instrument (CORAVEL, Baranne et al., 1979) was the precursor and basis of a generation of planet-hunting spectrographs (Pepe et al., 2018): ELODIE (Baranne et al., 1996), the instrument used to discover 51 Pegasi b; CORALIE (Queloz et al., 2000); the High Accuracy Radial velocity Planet Searcher (HARPS, Pepe et al., 2002); and the Spectrographe pour l'Observation des Phénomènes des Intérieurs stellaires et des Exoplanètes (SOPHIE, Perruchot et al., 2008). HARPS has achieved a short-term precision as low as 0.2 m s^{-1} , and long-term from $0.3\text{-}0.6 \text{ m s}^{-1}$ (Pepe and Lovis, 2008), successfully detecting many small super-Earths and Neptunes. The Echelle Spectrograph for Rocky Exoplanet- and Stable Spectroscopic Observations (ESPRESSO, Pepe et al., 2014) leads the newest generation of spectrographs, with the aim of achieving 10 cm s^{-1} precision. It confirmed the presence of Proxima b with a precision of 26 cm s^{-1} (Suárez Mascareño et al., 2020).

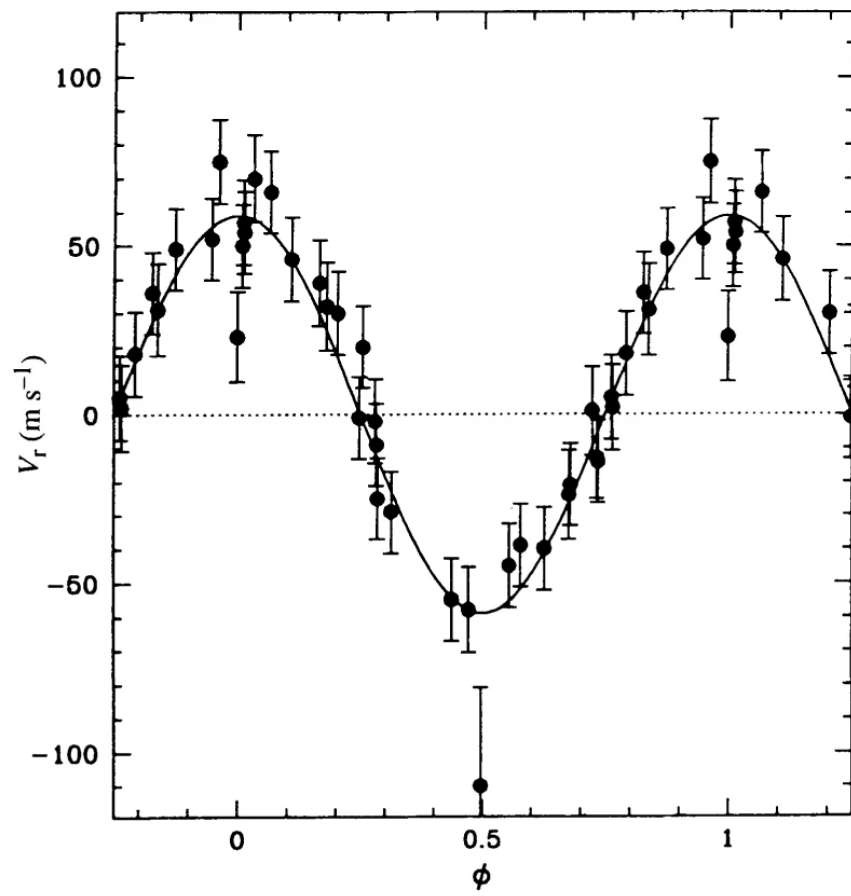


Figure 1.12: The phase-folded radial velocity curve of 51 Pegasi b. Reproduced from Mayor and Queloz (1995).

1.2.3 Further methods

Transit timing variations

Transit Timing Variations (TTVs) are an unusual method of exoplanet detection, as they can only be investigated when an orbiting body in a system (planetary or larger) has already been found to transit. A two-body Newtonian solution is predictable, the planet orbiting on a constant period and thus having predictable future transit times. However, some planets have been found that do *not* transit to schedule, instead with transits occurring before or after the expected time (see Fig. 1.13). This can be caused by the gravitational force exerted by other additional bodies in the system, making it accelerate or decelerate its orbit. Precise measurements of TTVs over time and modelling of the deviations of measured transit times compared to the expected times can be used to obtain planet masses (Lithwick et al., 2012).

TTV signals are larger in the presence of larger mass planets, which exert more gravitational force, and also in systems with resonant orbits (Nesvorný and Vokrouhlický, 2016). The TTV signal is also sensitive to inclined and eccentric systems (Nesvorný, 2009).

The first system detected to have TTVs was Kepler-9, where Kepler-9 b and c both transit and show TTVs due to their interaction with each other (Holman et al. (2010)). This system has been the subject of much study since (e.g., Torres et al., 2011; Borsato et al., 2014; Hadden and Lithwick, 2014; Berger et al., 2018; Freudenthal et al., 2018; Borsato et al., 2019), with the longest baseline of TTVs shown in Freudenthal et al. (2018), see Fig. 1.13. Another notable system for which TTVs are present is TRAPPIST-1, a seven-planet system where the planet masses derived from TTVs are determined to a precision of 3-5 per cent, which, due to their small mass values, would be equivalent to a RV precision of 2.5 cm s^{-1} (Agol et al., 2021). This is more precise than our current RV capabilities, described in Section 1.2.2. There have been some 24 TTV discoveries of non-transiting planets, where the TTVs are seen in their transiting companions¹.

Direct imaging

Methods so far have relied upon observing the effect a planet has on other bodies in its system. It is, however, possible to image planets directly (see e.g., Currie et al., 2023b). The point-source image of an exoplanet can be detected through the reflected light from its star (visible wavelengths), or through its own thermal emission (infrared wavelengths). Unfortunately, a star is many orders of magnitude brighter than its planet; the ratio of planetary flux to stellar flux is very small. At visual wavelengths at a distance of 10 pc, this ratio for Jupiter and the Sun is $\sim 10^{-9}$; for the Earth and the Sun, it is $\sim 10^{-10}$. This contrast can be improved by moving to infrared wavelengths, which simultaneously reduces the brightness of the star and increases the brightness of a planet from its thermal emission.

A coronagraph is also employed to block as much stellar light as possible.

This method is more sensitive to larger planets on wider orbits, as a wider angular separation ensures the light from the star does not overwhelm the light from the planet. It is also more sensitive to young, warm and as such self-luminous planets (due to their thermal emission), which are found around young stars. Finally, nearby stars are preferable; a large, shorter period planet might be detectable as its angular separation will be increased due to proximity.

The first planet detected via direct imaging was 2MASS J12073346-3932539 b (Chauvin et al., 2004), a heavy Jupiter orbiting a brown dwarf. There have been some 35 planets detected via direct imaging¹, all of them having multiple Jupiter masses. A notable system with multiple planets detected via direct imaging is HR 8799 (see Fig. 1.14, Marois et al. (2008, 2010)). Facilities used for direct imaging of exoplanets include the Spectro-Polarimetric High-contrast Exoplanet REsearch instrument (SPHERE, Beuzit et al., 2008), the Gemini Planet Imager (GPI, Macintosh et al., 2008), the High Contrast coronagraphic Imager for Adaptive Optics (HiCIAO, e.g., Kuzuhara et al., 2013).

Gravitational microlensing

Gravitational microlensing (see e.g., Gaudi, 2012) exploits general relativity, namely that massive bodies will bend light around them. When the light from a distant background source is bent around a massive foreground lens, for example a galaxy, it produces images of the source which are distorted, multiple, and significantly amplified. This relies on the the alignment of the source, lens, and observer. Also, the relative motion of the source, lens, and observer leads to the amplification varying over time. In the microlensing scheme, where a lens is much less massive (e.g., a star or planet), the images are not resolved, but the amplification of the source is still seen. An exoplanetary system may act as a double lens if the alignment is favourable. The stellar lens will produce a relatively large peak in the brightness of the source, and the planet will produce a second, relatively small and shorter duration (compared to the peak caused by the stellar lens), peak (see Fig. 1.15), and from this one can infer the planet's presence and mass.

Microlensing probes a particularly unique region of parameter space (see Fig. 1.2). It is most sensitive to planets orbiting in the lensing zone, typically several AU, which is further out than those typically discovered by the transit or radial velocity methods. The amplitude of the secondary peak is largely independent of mass, allowing detection of even Earth-mass planets. There are, of course, disadvantages to this method. An event will only happen once. Because the probability of an event happening over a background star (due to the alignment needed) is low even looking towards the Galactic bulge, on the order of 10^{-8} , it requires many stars to be monitored. The timescale of the secondary peak event is

also short (hours to days), requiring a decent cadence of observation to record peak features. This is somewhat similar to the transit method.

The first planet discovery via microlensing was OGLE 2003-BLG-235 b/MOA 2003-BLG-53 b, a 1.5 Jupiter mass planet on a 3 AU orbit (Bond et al., 2004). There have been close to 200 planets discovered via microlensing since then¹, with the majority of discoveries having been made by the Optical Gravitational Lens Experiment (OGLE, Udalski, 2003), the Microlensing Observations in Astrophysics collaboration (MOA, Bond et al., 2001), and the Korea Microlensing Telescope NETwork (KMTNet, Kim et al., 2016).

Astrometry

Astrometry (see e.g., Perryman, 2018) is closely related to the radial velocity method: both are concerned with the motion of a star due to the gravitational influence of its planet(s). While the radial velocity method measures this motion in spectra, astrometry measures it from the actual motion of the star on the sky. For most star-planet systems, this motion is going to be tiny, requiring sub-milliarcsecond precision, and this is the main limiting factor of the method. The astrometric signature α is proportional to the planetary mass M_P (compared to the stellar mass M_\star) and the semi-major axis a , and inversely proportional to the distance d from the observer:

$$\alpha = \left(\frac{M_P}{M_\star} \right) \left(\frac{a}{1 \text{ AU}} \right) \left(\frac{d}{1 \text{ pc}} \right)^{-1} \text{ arcsec.} \quad (1.14)$$

So, astrometry is sensitive to more massive planets on long period orbits – a 5 Jupiter-mass planet on a 3 year orbit around a Sun-like star 10 parsecs away will produce a larger, \sim milliarcsecond astrometric signature.

There have been so few astrometric detections that they can be counted on one hand. The first was not a planet at 28 Jupiter masses, but a brown dwarf (Sahlmann et al., 2013), where the astrometric motion is shown in Fig. 1.16. A candidate Saturn-mass planet was then found orbiting an ultracool dwarf, TVLM 513-46546 (Curiel et al., 2020). The first confirmed discovery was GJ 896 A b, a Jupiter-like planet orbiting the main component of a low-mass stellar binary (Curiel et al., 2022). Finally, there was a joint astrometric and direct imaging detection of a large, perhaps not planetary (its mass being between 14 and 16 Jupiter masses) companion orbiting HIP 99770 (Currie et al., 2023a).

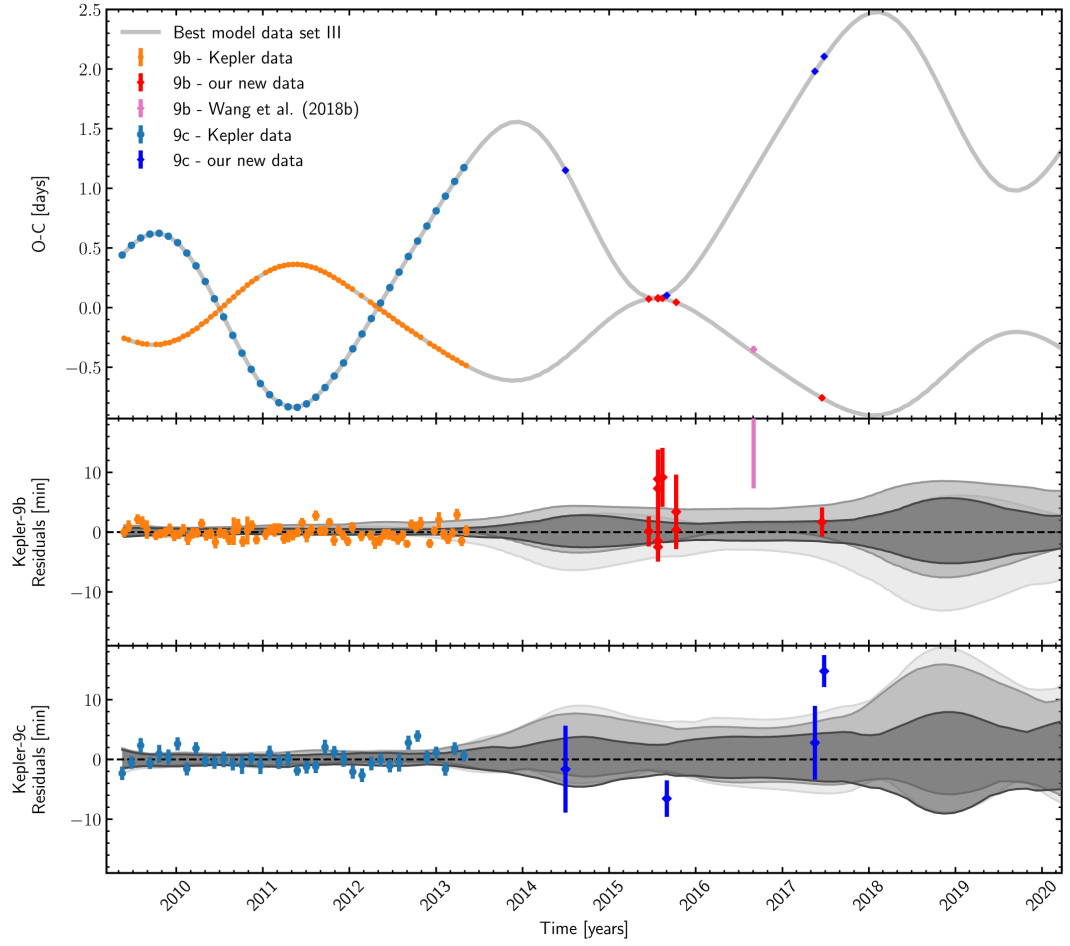


Figure 1.13: The transit timing variations of Kepler-9 b and c, shown as O-C transit times, O being the observed transit times, and C being the calculated transit times from a linear ephemeris. The residuals show the deviations from the photodynamical model of the system. Reproduced from Freudenthal et al. (2018).

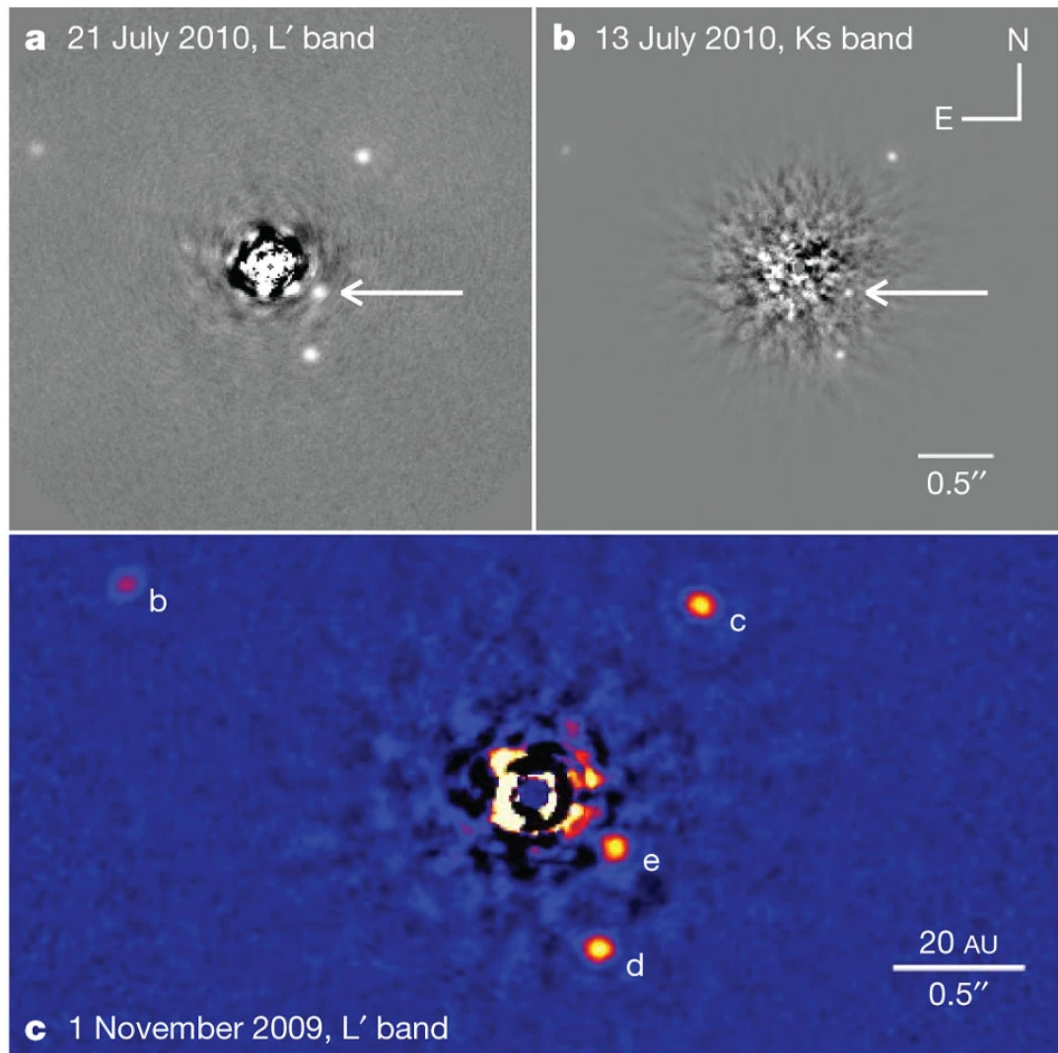


Figure 1.14: Direct imaging of the HR 8799 system, with the four planets, b, c, d, and e, labelled in the bottom panel. Reproduced from Marois et al. (2010).

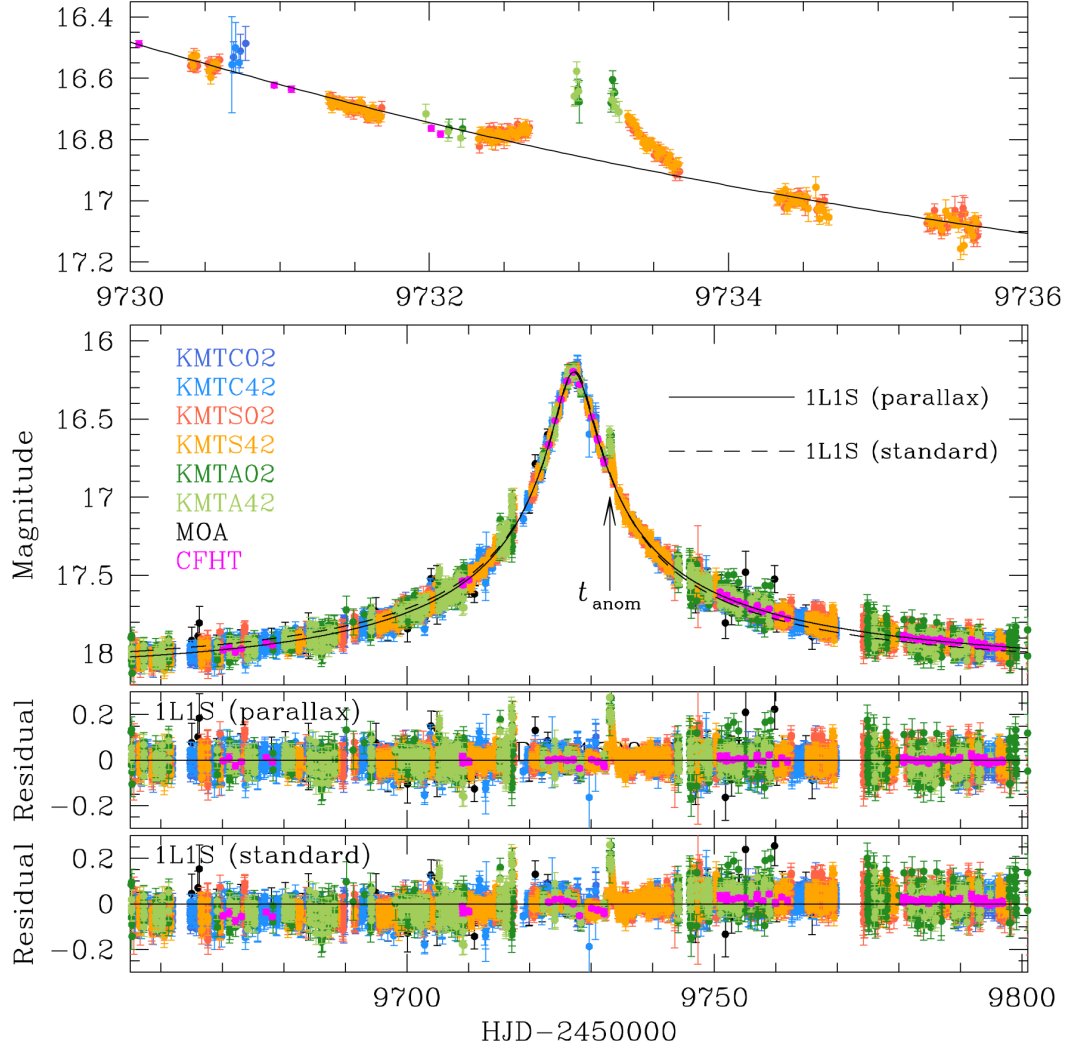


Figure 1.15: The light curve of the microlensing event MOA-2022-BLG-249. The colours of the symbols refer to different observation facilities, which are all observing the same microlensing event. The arrow marked t_{anom} points out the secondary peak produced by a super-Earth planet. Reproduced from Han et al. (2023).

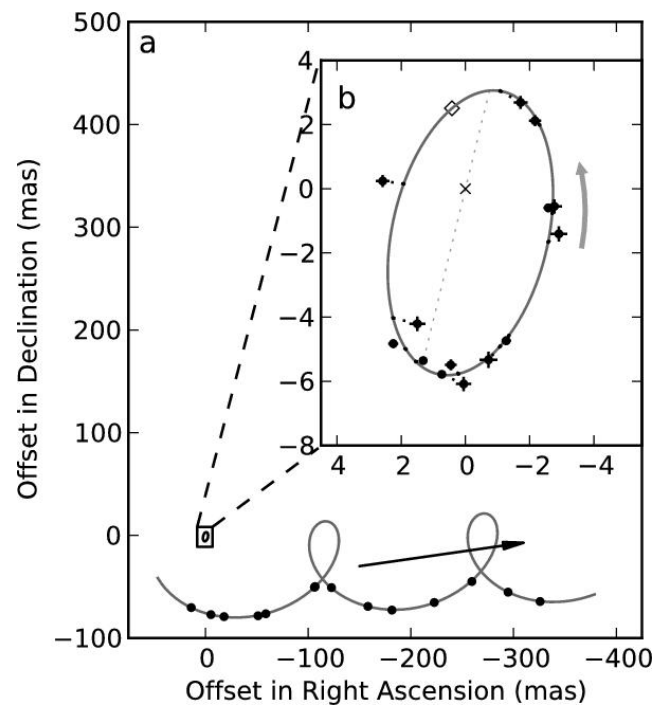


Figure 1.16: The astrometric motion of DE0823-49 caused by a brown dwarf, with the inset showing its barycentric orbit. Reproduced from Sahlmann et al. (2013).

1.3 Exoplanet characterisation

Beyond the initial detection of a planet, there is its characterisation, attempting to answer questions such as (but not limited to): how do we classify the planet based on its size and mass? What is the planet's internal structure? Is it habitable? How could it have formed and evolved?

1.3.1 Determination of system parameters

Here I will summarise the system parameters can be determined from a combination of the transit and radial velocity methods, as these are the focus of the thesis.

From transit light curves (Fig. 1.3), an orbital period, P , a transit depth, δ , and a transit duration, T_{14} , can be determined. From P and prior knowledge of the stellar mass, M_\star , the semi-major axis, a , can be calculated (Equ. 1.1). From δ and prior knowledge of the stellar radius, R_\star , the planetary radius, R_P , can be calculated (Equ. 1.2). From T_{14} and P , R_\star , R_P , and a , the inclination, i (Equ. 1.8), and impact parameter, b , can be calculated (Equ. 1.3).

From radial velocity curves (Figs. 1.10 and 1.11), a determination can be made of an orbital period, P , an RV semi-amplitude, K , and the eccentricity, e , and argument of periastron, ω . From K , e , and P , and prior knowledge of the stellar mass, M_\star , a minimum planetary mass, $M_P \sin i$, can be calculated. Combining this with i found from transit light curves, one can obtain a true planetary mass, M_P .

Combining these two methods, where possible, gives a full basic set of system and planetary parameters: P , a , R_P , i , b , e , ω , and M_P . These can be used, alongside other techniques based on the above methods, to infer further planetary characteristics.

1.3.2 Inference of further characteristics

There is information beyond that which can be ascertained by planet detection methods that is important for the further characterisation of discovered planets. Here I will explore a few ways in which measured parameters can be used to make further inferences, and how follow-up observations can complement knowledge from detections, all of which are utilised or mentioned later in the thesis.

Composition via bulk density

When there is a radius and mass measurement for a planet, its bulk density, ρ , can be inferred, where $\rho \propto M_P/R_P^3$.

Radius, mass and density alone cannot be used to find the definitive composition of a planet. Neither can one glean the presence of a core, and whether it is structurally differentiated or mixed. One can, however, study the planets in our Solar System, for example using seismology to study the core of the Earth, or a probe to sample the atmospheres of our gas giants. The main constituents are known to be iron, rock, ices, and H/He (see Section 1.3.3). Combining these observations with thermal equilibrium calculations to predict what species might be present, and their consequent equations of state under internal and external conditions, allows inference of plausible interior models.

It is common to use mass-radius plots of known exoplanets, shown in Fig. 1.17, to compare their masses and radii to computed interior models (e.g., Sotin et al., 2007; Otegi et al., 2020a). However, there is an issue of degeneracy: different interior models can fit the same mass and radius measurement of a planet. In order to draw meaningful conclusions about the structure of a planet, this must be accounted for (see Dorn et al., 2017).

One way to reduce degeneracies is with exoplanet host star observations. There is a correlation between stellar and planetary relative bulk abundances, allowing us to put an abundance constraint on what the planet is made from, which improves constraints on core size and mantle composition for rocky planets (Dorn et al., 2015).

Planets are often modelled as layered spheres: an iron core, a silicate mantle, a water layer, and an atmosphere. The atmosphere has often been modelled as pure H/He, but it is known that there is compositional diversity in atmospheres that might effect a planet's radius (Vazan et al., 2015), and the modelling of Dorn et al. (2017) accounts for this.

Characterisation via equilibrium temperature

The closer a planet is to its host star, the more flux it receives, and the hotter it is. This can be quantified in a simple way via the equilibrium temperature, T_{eq} , i.e. the temperature the planet would be if the energy received from the host star per unit time is equal to the thermal energy radiated by the planet per unit time.

The planet receives stellar flux, F_{\star} , over the area of the disc it presents towards the star, reflects a fraction A of the intercepted light where A is its albedo, and an assumption is made that it emits as a blackbody, thus:

$$T_{\text{eq}} = \left(\frac{(1 - A)F_{\star}}{4\sigma} \right)^{1/4}, \quad (1.15)$$

where σ is the Stefan-Boltzmann constant.

Liquid water is often defined as a prerequisite for a planet to harbour life, and so knowing the equilibrium temperature of a planet should allow a determination whether it would be within a star's "habitable zone" - indeed, the equilibrium temperature was used

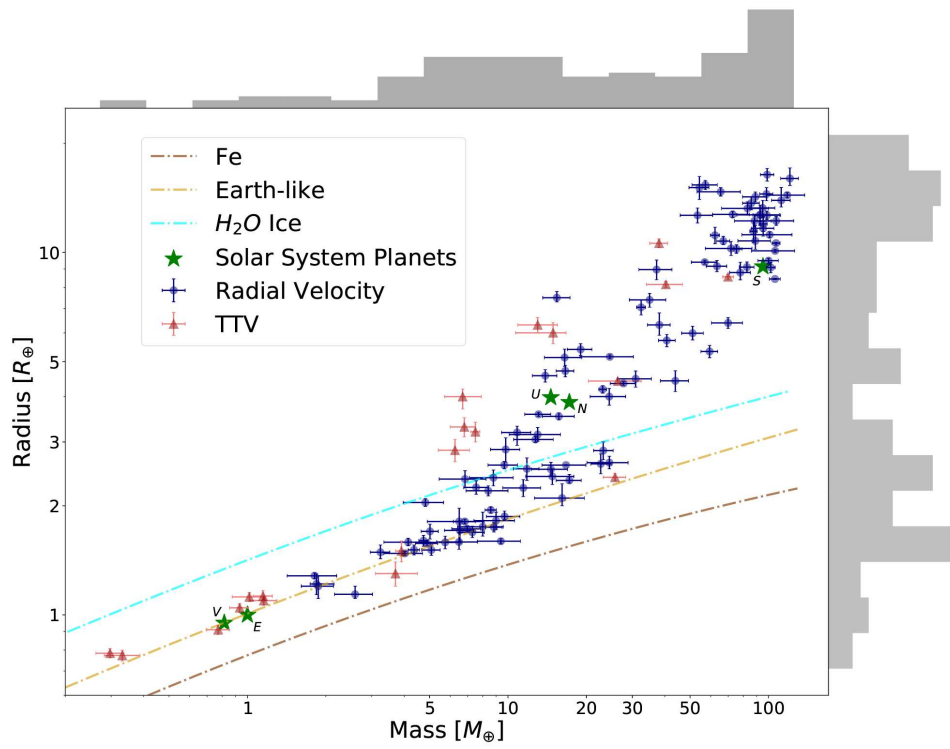


Figure 1.17: Mass-radius plot showing only planets with robust mass and radius measurements. Composition lines of pure-iron, Earth-like, and water-ice from Dorn et al. (2015) are shown. The grey histograms represent the distribution of mass and radii. Reproduced from Otegi et al. (2020a).

in early calculations of the habitable zone (e.g., Kasting et al., 1993). However, there is a dependence on albedo, the calculation of which varies depending on stellar effective temperature and atmospheric characteristics like clouds. So, the preferred measure is instead in terms of the incident stellar flux at the top of a planet’s atmosphere (Kopparapu et al., 2013).

The equilibrium temperature of a planet can also be useful for classifying it into categories like cold, warm, hot and ultra-hot (see Section 1.3.3), as certain physical processes occur at specific temperatures, like phase transitions. Equilibrium temperature is also used in calculation of mass-loss rates under photoevaporation (e.g., Kubyskhina et al., 2018), which is discussed in Sections 3.4 and 5.5.3.

Spin-orbit angle via the Rossiter-McLaughlin effect

The Rossiter-McLaughlin (RM) effect (Rossiter, 1924; McLaughlin, 1924; Triaud, 2018) can be seen when a planetary transit is observed spectroscopically. A rotating star appears to the observer to have one hemisphere blue-shifted, the other red-shifted. When a planet transits, it covers different parts of the star sequentially. As it covers the blue-shifted portion, the flux seen by the observer will have a redward offset, which creates a positive shift in the measured radial velocity of the star, deviating from the usual Doppler reflex motion (see Section 1.2.2). If there are enough radial velocity measurements throughout the transit, the effect shown in Fig. 1.18 can be seen.

The size and shape of the effect contains information relevant to planetary characterisation. The semi-amplitude, A_{RM} , scales with the size of the planet (via transit depth, δ), the impact parameter of the planet, b , and the stellar rotation velocity at the equator, v , which is modified by $\sin i_{\star}$, the inclination on the sky of the stellar rotation axis (Triaud, 2018):

$$A_{\text{RM}} \approx \frac{2}{3} \delta v \sin i_{\star} \sqrt{1 - b^2}. \quad (1.16)$$

For a typical hot Jupiter around a typical star, the semi-amplitude is of order $\sim 20 \text{ m s}^{-1}$, with a few hour duration and ingress/egress of a half-hour. This necessitates a cadence of around 10-15 mins when trying to observe the Rossiter-McLaughlin effect.

From the shape of the effect, it can be seen how long a planet spends over one hemisphere versus the other, from which the sky-projected angle between the planet’s orbital path and the stellar equator, known as the sky-projected spin-orbit angle or obliquity, λ , can be determined. There are various ways to model λ , including the classical RM effect used in the original studies, Doppler tomography (Collier Cameron et al., 2010), the RM

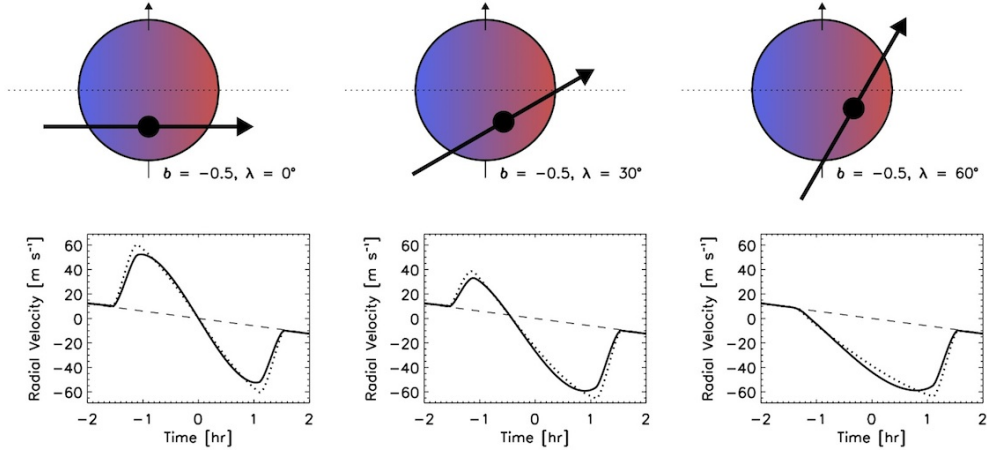


Figure 1.18: Three transiting planet trajectories with the same impact parameter and different spin-orbit angles, which produces different Rossiter-McLaughlin effect shapes. The dotted lines represent a case with no limb-darkening. Reproduced from Gaudi and Winn (2007), coloured.

effect reloaded (Cegla et al., 2016), and an update to the latter, the RM effect revolutions (Bourrier et al., 2021).

Finding the *true* spin-orbit angle, Ψ is the goal. Fabrycky and Winn (2009) derive statistical constraints on Ψ , but to determine it empirically, a measurement of i_\star is needed. This is hard to do accurately and has often relied on transforming a measurement of rotation period into a stellar equatorial velocity, v . The most effective method is via asteroseismology, where oscillation modes are split by rotation. Ψ can also be determined by the RM effect reloaded technique, which utilises the differential rotation of the star to do so (Cegla et al., 2016).

Measuring the RM effect allows inferences to be made about the formation and evolution histories of exoplanet systems. Disc-driven migration is expected to conserve alignment between the spin of a star and the orbital plane of its planet(s) (e.g., Winn and Fabrycky, 2015); however, a variety of formation and migration scenarios can cause misalignments, for example a primordial tilt (e.g., Lai et al., 2011), a tidal torque on the protoplanetary disc caused by a nearby star (Batygin, 2012), or scattering caused by an outer planetary or stellar companion (e.g., Wu and Murray, 2003). Mutual inclinations between planets in a system and with their star can distinguish somewhat amongst these.

1.3.3 Types of exoplanets

Various attempts have been made to classify the thousands of planets that have been discovered. This subsection will begin by defining what a planet actually is (which is not as simple as it sounds), then will explore several categories of planet, several of which feature later in this thesis.

The definition of an exoplanet

The International Astronomical Union’s (IAU) working definition of an exoplanet has changed several times over the past few decades as the knowledge on exoplanets has grown, and the current (2018) working definition of an exoplanet is as follows (Lecavelier des Etangs and Lissauer, 2022):

1. Exoplanets are objects that:
 - (a) have true masses below the limiting mass for thermonuclear fusion of deuterium ($\sim 13M_J$ for solar metallicity);
 - (b) orbit stars, brown dwarfs, or stellar remnants;
 - (c) have a mass ratio with the central object of $\lesssim 1/25$, i.e. a criterion required for linear stability of the Lagrangian points L_4 and L_5 , regardless of how they formed;
 - (d) have a minimum mass and size the same as that used to define objects as planetary in our Solar System, i.e. they must have mass sufficient for a (nearly) spherical hydrostatic equilibrium shape due to self-gravity and for clearing the neighbourhood around their orbits.
2. Substellar objects with masses above the limiting mass for thermonuclear fusion of deuterium are brown dwarfs, regardless of how they formed or where they are located.
3. Free-floating objects in young stellar clusters with masses below the limiting mass for thermonuclear fusion of deuterium are not planets; they are “sub-brown dwarfs”.

Under this definition, exoplanets span a vast range of orbital periods, radii, and mass, and thus have a wide diversity in their characteristics. Based on these, loose categories of planets can be defined, the boundaries between which are debated – approximate radius and mass limits are given in Table 1.1. The radius bins are taken from Borucki et al. (2011a) and the mass bins from Stevens and Gaudi (2013). It is also useful to consider categories based on planetary composition, discussed in Section 1.3.2 and below.

Planet class	Radius (R_{\oplus})	Mass (M_{\oplus})	Solar system	Radius (R_{\oplus})	Mass (M_{\oplus})
Earths	< 1.25	$0.1 - 2$	Venus	0.949	0.815
			Earth	1.00	1.00
			Mars	0.53	0.107
super-Earths	$1.25 - 2$	$2 - 10$	-		
Neptunes	$2 - 6$	$10 - 100$	Uranus	4.01	14.5
			Neptune	3.88	17.1
Jupiters	$6 - 15$	$100 - 4100^*$	Saturn	9.45	95.2
			Jupiter	11.2	318

Table 1.1: Planet categories, named after their Solar System counterparts (right), defined by radii and mass ranges. These boundaries are not universally adopted; here, boundaries defined by Borucki et al. (2011a); Stevens and Gaudi (2013) are used. $^*4100M_{\oplus} \approx 13M_J$, the deuterium burning limit.

Terrestrial planets - Earths and super-Earths

Earths and super-Earths are terrestrial (a.k.a. telluric or rocky) planets.

The classical view is that they are composed of a silicate rock mantle with a differentiated iron core, based on the terrestrial planets in our Solar System, Mercury, Venus, Earth, and Mars. However, due to the degeneracies involved in composition modelling given a specific mass and radius (see Section 1.3.2), and considering the composition of the material available for terrestrial planet formation, it is possible that terrestrial planets of different compositions may exist (e.g., Seager et al., 2007; Bond et al., 2010; Rogers and Seager, 2010). Coreless terrestrial planets are possible when metallic iron in the accreting material is oxidised, binding it into the silicate mantle (Elkins-Tanton and Seager, 2008). Around stars with a low carbon-to-oxygen ratio, like 55 Cancri, carbon- and silicon-dominated terrestrial planets may exist (e.g., Kuchner and Seager, 2005; Madhusudhan et al., 2012; Teske et al., 2013; Wilson and Militzer, 2014). The inner disc of a brown dwarf may also be an example of a low oxygen-carbon ratio environment that could produce carbon planets (Pascucci et al., 2013).

Terrestrial planets also likely form enriched secondary atmospheres through outgassing from their interiors (Dorn and Heng, 2018; Thompson et al., 2021), volatile delivery from cometary impacts (Kral et al., 2018), or from accretion in late gas discs composed of mainly CO, carbon, and oxygen (Kral et al., 2020). Indeed, the Earth has both a secondary and tertiary atmosphere, the latter caused by the presence of life.

The discovery of many terrestrial planets has been one of the major successes of planet-finding missions like *Kepler* (Borucki et al., 2010) and *TESS* (Ricker et al., 2015). The first confirmed terrestrial planet discovery with both mass and radius determined was Kepler-10 b (Batalha et al., 2011), following discoveries of potential terrestrial planets; for example, multiple planets in the GJ-581 system (Udry et al., 2007; Mayor et al., 2009) and

CoRoT-7 b (Léger et al., 2009). These are all “super-Earths”, a surprising discovery.

Super-Earths are terrestrial planets that are significantly larger than the Earth. They have no Solar System analogue. Some definitions include planets of up to $4R_{\oplus}$ (Morbidei and Raymond, 2016), but from radial velocity monitoring of *Kepler* super-Earths, the division between high-density and rocky super-Earths and lower density mini-Neptunes with large volatile envelopes likely lies around $1.5 - 2.0R_{\oplus}$ (e.g., Lopez and Fortney, 2014; Rogers, 2015; Chen and Kipping, 2017). The lower boundary was defined by the *Kepler* team (Borucki et al., 2010), separating Earth-analogues from their larger counterparts.

A subset of terrestrial planets are lava worlds; they are theorised to have surfaces covered mostly or entirely by molten lava (Léger et al., 2011; Rouan et al., 2011). Their very short orbital periods ($P \lesssim 1$ d, $a \lesssim 0.02$ AU) result in a tidally locked planet with a day-side always facing the star, and extreme heating at the substellar point, producing temperatures up to ~ 3000 K. At temperatures this high, the silicate rock surface turns into a lava ocean that covers most of the dayside. The nightside is conversely expected to be very cold. Known lava worlds are CoRoT-7 b (Léger et al., 2009, 2011), Kepler-10 b (Batalha et al., 2011; Rouan et al., 2011), and Kepler-78 b (Sanchis-Ojeda et al., 2013).

There are also theorised water worlds, where large terrestrials could have compositions dominated by volatiles like water ice (Kuchner, 2003; Léger et al., 2004). They are expected to have an outer water envelope, and, depending on orbital period, this could be entirely gaseous (Léger et al., 2004), liquid with a gaseous atmosphere (Kuchner, 2003), a liquid ocean which would have an eventual bottom ice layer at the pressure required for liquid to solidify (Léger et al., 2004), or completely icy, perhaps with a subsurface ocean (Fu et al., 2010; Ueta and Sasaki, 2013). There could also be water-cycling between the ocean and mantle, allowing for exposed continents (Cowan and Abbot, 2014). Candidate ocean worlds include GJ 1214 b (Charbonneau et al., 2009), TOI-1452 b (Cadieux et al., 2022), and Kepler-138 c and d (Piaulet et al., 2023).

Neptunian planets - Neptunes and sub-Neptunes

Neptunian planets have gaseous envelopes, typically hydrogen and helium, surrounding a core of rock and heavier metals.

As stated previously, the division between super-Earths and Neptunes comes down to Neptunian planets having a volatile envelope, reducing their overall density. Due to its low density, even a relatively minor amount of H/He of ~ 1 per cent has a large impact on planet’s radius; for planets of sub-Neptune size or larger, the H/He envelope will dominate the planet’s size.

Neptune and Uranus are the two Neptunian planets in our Solar System, and they are both ice giants. They have H-He envelopes of $\sim 10 - 20$ per cent of their total mass,

but also contain water, methane and ammonia ices above their small rocky cores (Podolak et al., 1995). However, their internal structures are not well understood (Helled et al., 2011; Podolak and Helled, 2012; Helled and Fortney, 2020). Most Neptunian planets that have been found have orbits much closer in, and are unlikely to be ice giants - Neptune and Uranus probably formed beyond the snow line (D’Angelo et al., 2010). Microlensing has been able to find a potential Uranus-like exoplanet (Poleski et al., 2014).

A distinction is often drawn between Neptunes (those of Neptune radius and larger) and sub-Neptunes (those smaller than Neptune’s radius, but which can be more massive), where sub-Neptunes lie in the $1.7 - 3.5R_{\oplus}$ range. The composition of sub-Neptunes is a degenerate problem (see Section 1.3.2): they could have a massive layer of water without requiring a large envelope to explain their radii (e.g., Venturini et al., 2020); they could be rocky with a large envelope (e.g., Owen and Wu, 2017); or they could be a mix of the two. Small atmospheres versus larger ones could be distinguished through the transmission spectra of these planets (Hu et al., 2021) – and the first JWST transmission spectrum of K2-18 b, one of the archetypal planets illustrating this problem (Cloutier et al., 2017), was serendipitously published the day I wrote this section (Madhusudhan et al., 2023). These observations are consistent with a Hycean world, defined as a planet composed of a water-rich interior and massive ocean under a H_2 rich atmosphere (Madhusudhan et al., 2021).

The *Kepler* mission revealed an interesting class of Neptune dubbed the “super-puff”. They have masses comparable to super-Earths but radii more like gas-giants, giving them densities of $< 0.1 \text{ g cm}^{-3}$. Few are known: all 3 planets in the Kepler-51 system (Masuda, 2014); Kepler-79 d (Jontof-Hutter et al., 2014; Yoffe et al., 2021); and K2-141 c likely completes this list, albeit with only an upper mass estimate (Bonomo et al., 2023). The transmission spectra of Kepler-51 b and c are featureless, with no significant water absorption, and probably the result of a high-altitude aerosol layer (Libby-Roberts et al., 2020).

There are two important exoplanet population features that affect Neptunian planets: the radius valley; and the Neptunian desert (see Section 1.4.1). The radius valley separates the population of super-Earths from sub-Neptunes, with a paucity of planets in-between. The Neptunian desert is a distinct lack of Neptune-sized planets on short-orbital periods. Both point to photoevaporation of low-density atmospheres sculpting the planet size distribution (Fulton and Petigura, 2018).

Gas giants - Jupiters and hot Jupiters

Jupiter-sized planets, like Saturn and Jupiter in our Solar System, are dubbed gas-giants as they are large and composed mostly of gas (hydrogen and helium). Only gas giants can achieve radii above $7R_{\oplus}$, due to electron degeneracy pressure in terrestrial materials

(D’Angelo and Lissauer, 2018) – for $M_P \lesssim 500M_\oplus$, Coulomb forces balance gravity in hydrostatic equilibrium, but for larger masses, gravity is instead balanced by the degeneracy pressure of free electrons, and as more mass is added, the planet shrinks (Seager et al., 2007).

Upon gravitational compression, hydrogen becomes a metallic fluid, i.e. protons within a sea of degenerate electrons (Wigner and Huntington, 1935). Treatments of the composition of Jupiter have often consisted of three layers (e.g., Guillot, 2005; Nettelmann et al., 2012; Miguel et al., 2016): a core (a central region of heavy element concentration); a mantle (the aforementioned metallic hydrogen, enriched with helium and heavy elements); and a molecular envelope (hydrogen and helium, again enriched). At the mantle-envelope boundary, helium rain is thought to form. However, the problem with this three-layer model is in the sharp boundaries; instead, the core was theorised “fuzzy” and not necessarily distinct from the envelope with some hydrogen and helium, and a gradual heavy-element structure in the deep interior (Helled and Stevenson, 2017a). Observations from Juno (Bolton et al., 2017) have solidified (pun intended) our understanding of Jupiter and its interior (Debras and Chabrier, 2019; Stevenson, 2020): it is *not* fully mixed (i.e. it does not have a heavy-element mass fraction that increases towards the core); instead, it has a “dilute” or “fuzzy” core which extends to up to half of its radius (Helled et al., 2022), challenging ideas about gas giant formation (see Section 1.4.3). Whether this applies to gas giants in exoplanetary systems is an open question.

The first exoplanet discovery was a gas giant, but a surprising one with an orbital period of 4 days (Mayor and Queloz, 1995), resulting in a new sub-class of Jupiters being defined. Hot Jupiters are gas giants on orbital periods < 10 days. Like super-Earths, these hot Jupiters have no Solar System analogue. It is common for hot Jupiters to have equilibrium temperatures of ~ 1500 K; KELT-9 b is over 4000 K, making it as hot as a K-dwarf star (Gaudi et al., 2017). This makes it an “ultra hot Jupiter”, a hot Jupiter with an equilibrium temperature over 2000 K (e.g., Baxter et al., 2020). The atmospheres of hot Jupiters are relatively easy to study with transmission spectroscopy, and many atomic and molecular species have been seen in them (e.g., Charbonneau et al., 2002; Swain et al., 2009; Kreidberg et al., 2014; Evans et al., 2016; Lothringer et al., 2022); recently, HD 149026 b was found to have high atmospheric metal enrichment (Bean et al., 2023) and vanadium oxide was detected on WASP-76 b (Pelletier et al., 2023). Due to the relative ease with which hot Jupiters can be detected via transit and radial velocity observations (being giant planets on short orbits), many hot Jupiters have been discovered despite their relatively low occurrence rate (see Section 1.4.2), especially during the early years of exoplanet detection (e.g., Mayor and Queloz (1995); Konacki et al. (2003); and the many WASP discoveries).

Comparatively, the gas giants in our Solar System are cold, with long orbital periods:

~ 12 and ~ 29 years for Jupiter and Saturn, respectively. Another boundary can therefore be defined: between hot and cold Jupiters lie warm Jupiters, which have boundary conditions defined differently depending on who you ask: orbital periods of 10-200 days (Dawson and Johnson, 2018; Jackson et al., 2021; Dong et al., 2021) or 10-100 days (e.g., Wu et al., 2018), a planet-star separation of $10 < a/R_\star < 100$ (e.g., Dong et al., 2023), or boundaries determined by how period affects the occurrence rate of giant planets (Santerne et al., 2016). Warm Jupiters have eccentricities distribution ranging from circular to highly eccentric (Dong et al., 2021), and have more companion planets than hot Jupiters (Huang et al., 2016), which may have implications for potential formation pathways (see Section 1.4.3).

Other boundaries and types of Jupiter have also been posited. For example, a boundary for “temperate” Jupiters has been defined as those having an equilibrium temperature between 350-500 K, where they are relatively free of condensates, and have a atmosphere/spectrum dominated by CH_4 , NH_3 , and H_2O , and Jupiters under this definition have been found from periods between 100-300 days (Encarnaz et al., 2018).

Regardless of the exact boundary definition, planets on longer periods are hard to detect with the transit and radial velocity methods, due to their biases towards shorter period planets. There are less than 50 planets found with longer periods (> 100 days) that have measured masses and radii. The longest period confirmed transiting gas giant discovered by *TESS*, TOI-4600 c, has a period of 480 days (Mireles et al., 2023), joining a handful of long period giant planets that bridge the gap between warm Jupiters and our Solar System gas giants. There are two *Kepler* gas giants with masses and radii that orbit on periods greater than 1000 days: Kepler 167 e (Chachan et al., 2022), and Kepler-1647 b (Kostov et al., 2016) – the other discoveries in this regime are from direct imaging (see Section 1.2.3).

1.4 The exoplanet population

With a large number of exoplanet discoveries comes the ability to look at the planetary population as a whole. This can be through looking at features in the distributions of exoplanet properties, or in the overall occurrence rates of planets, and both are discussed in this section. I also discuss the formation and evolution of planets, as this will sculpt the overall distribution of planets.

1.4.1 Population features

When looking at the overall population of exoplanet discoveries, for example in period-radius and period-mass space, or by distributions of particular properties like radii, it can be seen that their distribution is not uniform. Some of this will be due to the biases in detection methods, but others may be an imprint of particular formation and/or evolution processes that affect what exoplanets can exist. Here, I will describe two of these features: the Neptune desert, and the radius valley, which are both explored later in this thesis.

The Neptune desert

In both period-radius and period-mass space, there is a conspicuous empty space (Fig. 1.19), a lack of intermediate sized planets from approximately $2 - 4R_{\oplus}$ and $6 - 250M_{\oplus}$ on short orbital periods of $\lesssim 3$ days. This should not be one of the aforementioned observational biases; both the transit and radial velocity method should readily find close-in, relatively large planets.

This desert was first found by Szabó and Kiss (2011) when there were just 106 planets with precise masses, radii, and orbital periods. Its existence was confirmed multiple times in different papers over the next few years: the “sub-Jovian Pampas” (Beaugé and Nesvorný, 2013); a “hot-super-Earth desert” (Lundkvist et al., 2016); and the “short-period Neptunian desert” (Mazeh et al., 2016). Mazeh et al. (2016) drew a triangular shape for this region and derived empirical relations for the slopes of the upper and lower boundaries in period-radius and period-mass space, using a technique that enhanced the planetary “density” (occurrence rate) contrast above and below the boundaries.

Turning to what could cause the desert, it was posited that short-period exoplanets may lose their envelopes through atmospheric escape caused by the X-ray and extreme ultra-violet (XUV) radiation of their host stars (Kurokawa and Nakamoto, 2014). Matsakos and Königl (2016) instead interpret the desert boundaries as being a consequence of the orbital circularisation of planets arriving near their host stars due to high-eccentricity migration. Considering both ideas, Owen and Lai (2018) found that the lower boundary could be accounted for by the photoevaporation of highly irradiated sub-Neptunes. However, they also

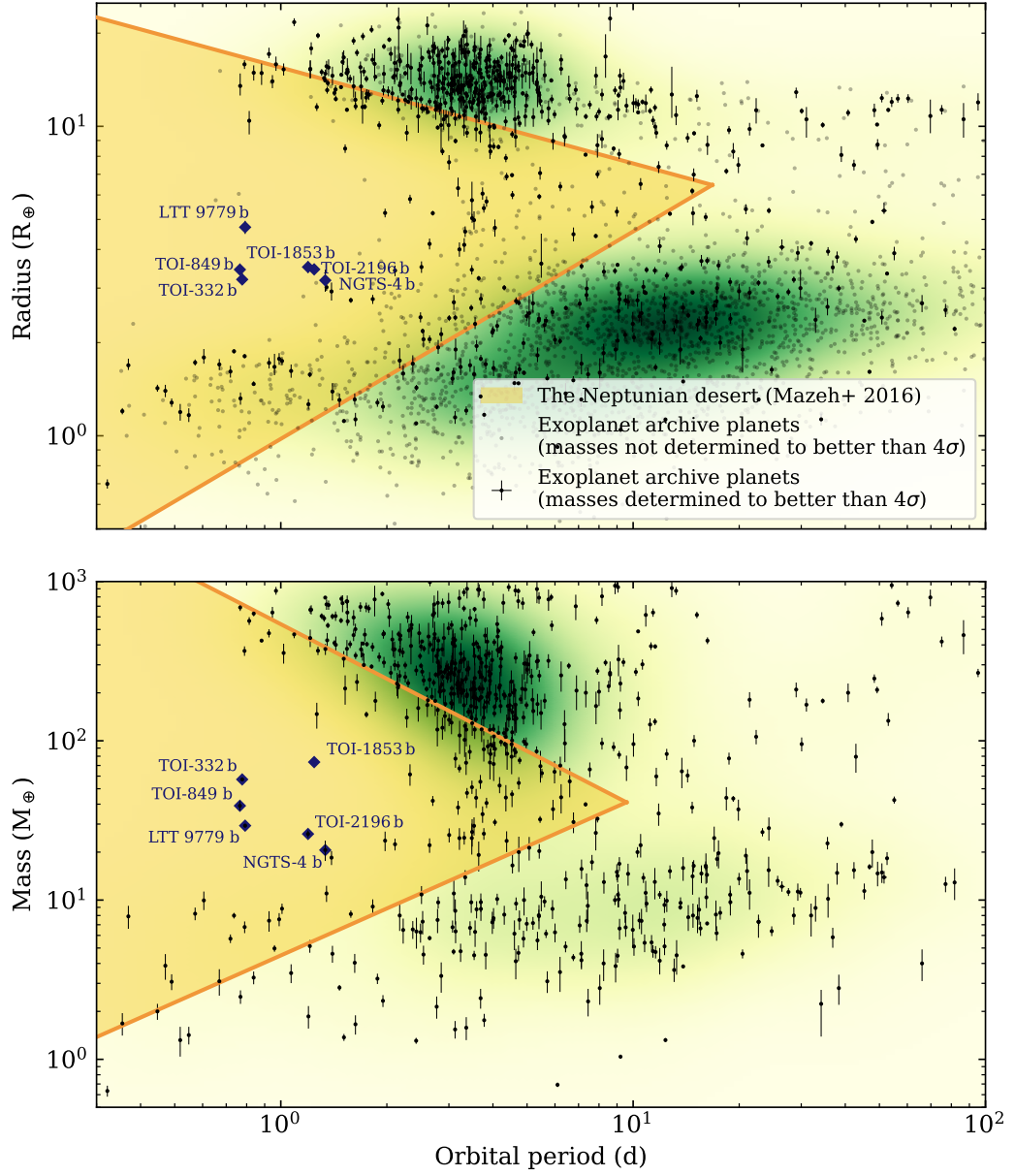


Figure 1.19: The Neptunean desert in period-radius (top) and period-mass space (bottom). The desert boundaries from Mazeh et al. (2016) are plotted as solid lines, with the enclosed desert area shaded in yellow. Known exoplanets were sourced from the NASA exoplanet archive (<https://exoplanetarchive.ipac.caltech.edu/>) on 16 Sep 2023. Planets without mass determinations or mass determinations worse than 4σ are plotted as grey dots in the top panel only; those with mass determinations better than 4σ are plotted as black dots. Specific planets found within the desert are labelled (dark blue diamonds).

determined that the upper boundary may extend too high in radius and mass for photoevaporation to be the sole mechanism responsible for sculpting the desert, a conclusion also drawn by Ionov et al. (2018), and in direct contrast to Kurokawa and Nakamoto (2014). They suggest an alternate explanation of a “tidal disruption barrier”. Planets migrating inwards under high-eccentricity migration need to be able to circularise and stabilise to not end up inside their host star, and only planets above a certain mass and radius at a given orbital period can do this successfully. They find this can also contribute to the lower boundary, in addition to photoevaporation. Vissapragada et al. (2022) searched for strong outflow features in gas giant planets, and find that outflow efficiencies are too small for photoevaporation to carve the upper boundary of the desert, further supporting that it may be caused by migration. Further studies like this are necessary to constrain mass-loss rates around the boundaries of the desert (e.g. Guilluy et al., 2023).

Other properties have been correlated with the desert boundaries. Owen and Lai (2018) made two more predictions: that the boundaries depend on the core mass of the planet (planets with larger core masses can exist further into the period-boundary of the desert), and the metallicity of the planet (planets with higher metallicity can exist at shorter periods). More planets have been found around higher-metallicity stars at the lower boundary of the desert, implying metallicity may have an effect on photoevaporation (Petigura et al., 2018; Owen and Murray-Clay, 2018). Hallatt and Lee (2022) predicts that, around M-dwarf stars, the sub-Jovian desert should be present at orbital periods of $\lesssim 0.7$ days, as opposed to the 3 days seen around solar-mass stars. Szabó and Kálmán (2019); Szabó et al. (2023) explore planet occurrence near the boundaries as a function of a number of system parameters, finding that it depends on stellar parameters including, in order of decreasing significance, temperature, metallicity, $\log g$, and maybe stellar mass, again invoking the effect of metallicity on photoevaporation.

In the past few years, there have been an increasing number of planets discovered within the desert. This has led to the desert becoming more of a “savanna”, populated by these “giraffe” planets (Kálmán et al., 2023). This includes NGTS-4 b (West et al., 2019), LTT-9779 b (Jenkins et al., 2020), TOI-849 b (Armstrong et al., 2020), TOI-2196 b (Persson et al., 2022), TOI-332 b (Osborn et al. (2023), see Section 5), and TOI-1853 b (Naponiello et al., 2023), and all are labelled in Fig. 1.19. The latter two are massive, dense, bare cores, likely having no envelopes, which challenges conventional ideas of planet formation and evolution. TOI-849 b may have half the mass of TOI-332 b and TOI-1853 b, but is still confusingly massive, with a H-He envelope of probably no more than 4 per cent of its mass. Photoevaporation is not expected to account for the mass-loss required to strip cores this big of the Jupiter-like envelope they would be expected to accrete. LTT-9779 b has a probable 9 per cent H-He envelope, but its ultra-hot nature raises questions about how it may have

managed to retain such a large envelope – one possible explanation is that the star has had an unusually dim XUV history, as hinted at by recent XMM-Newton observations (Fernández Fernández et al., in review). Current and future study of the planets within the desert is important evidence towards the formation and evolution mechanisms that sculpt it (e.g., Kálmán et al., 2023), and is the focus of my work on Nomads in Section 4.

The radius valley

The California-*Kepler* survey (CKS, Petigura et al., 2017; Johnson et al., 2017) used Keck/HIRES (Vogt et al., 1994) to perform homogeneous spectroscopic characterisation of a stellar sample to obtain stellar parameters including their radii, and subsequently self-consistently recalculated the radii of planets orbiting these stars. As explained in Section 1.2.1, the radius of a planet obtained from the transit method is dependent on a prior determination of the stellar radius, and uncertainties in the stellar radius propagate into the planetary radius. CKS allowed Fulton et al. (2017) to build on previous work (Owen and Wu, 2013) by updating the planet radii and confirming the existence of a gap in the distribution of planetary radii between 1.5 and 2.0 R_{\oplus} , shown in Fig. 1.20. A follow up study utilising *Gaia* DR2 (Gaia Collaboration et al., 2018) to improve the stellar and planetary radii uncertainties allowed them to discover that host star mass sculpts the gap – for lower mass stars, the gap shifts to the left, explained by smaller stars creating smaller planetary cores (Fulton and Petigura, 2018). The orbital period also affects the gap, and this, alongside host star mass, relates to a planet’s XUV-irradiation history.

Indeed, the gap was interpreted by Owen and Wu (2017) to be due to photoevaporation of planets which “herds” them into two regimes: they either become bare cores of $\sim 1.3R_{\oplus}$, or retain a H/He envelope and have larger radii $\sim 2.6R_{\oplus}$. Lehmer and Catling (2017) also found that rocky planets have an upper limit on radius of 1.5 – 2 R_{\oplus} , again due to XUV emission from young stars during their first 100 Myr.

Another sample was then analysed by Van Eylen et al. (2018): 117 planets with even more precise radii, thanks to stellar parameters derived by asteroseismology. As well as reproducing the bi-modality, they characterised the slope of the valley in period-radius space, shown in Fig. 1.22. The slope of the valley - whether it is positive or negative - allows us to constrain some formation/evolution scenarios. Lopez and Rice (2018) proposed late, gas-poor formation to be a potential cause of the valley, where short period super-Earths are a population of planets that never had envelopes in the first place, rather than losing them due to photoevaporation. In this situation, the location of the valley would be a function of available solid material for accretion in collisions, and cause a positive slope. This is in direct contrast to photoevaporation, where larger cores are most resistant to photoevaporation, causing a negative slope. Van Eylen et al. (2018) find a negative slope, inconsistent with late, gas-

poor formation. They also draw inferences about core composition: the clear gap they see implies homogeneous core composition (Owen and Wu, 2017); and the location of the valley is consistent with cores made of a large fraction of iron, rather than ice (Owen and Wu, 2017; Jin and Mordasini, 2018). Martinez et al. (2019) reanalyse the CKS sample and find a slope consistent with Van Eylen et al. (2018).

Another alternative scenario for the origin of the gap has been recently explored. Ginzburg et al. (2018) investigate core-powered mass loss, where the luminosity of a rocky core can boil off a light envelope but not a heavy one. They show it can naturally reproduce the gap, but cannot dismiss simultaneous photoevaporation. Gupta and Schlichting (2019) came to a similar conclusion. This has led to efforts being made to distinguish which is the dominant mechanism. Berger et al. (2018) find evidence for core-powered mass-loss being the dominant mechanism. Rogers et al. (2021) simulated what kind of planet survey could distinguish between the mechanisms via analysing the radius gap in 3D parameter space, finding that a survey of $\gtrsim 5000$ planets with a wide range of stellar mass and with measurement uncertainties of $\lesssim 5$ per cent is needed, requirements which neither the CKS or *Gaia-Kepler* Survey fulfil. Ho and Van Eylen (2023) refit a sample of planets using *Kepler* 1 min short cadence observations, showing again that precise and homogeneously determined planetary parameters are very important for radius valley studies, and they find that the location of the valley strongly depends on orbital period and stellar mass (and weakly depends on stellar age), and has a slight preference for core-powered mass-loss.

Owen and Schlichting (2023) use the Bondi radius (the sonic point of core-powered outflow) and the penetration depth of XUV photons to map out where in parameter space photoevaporation or core-powered mass-loss dominates atmospheric escape. They find several points of interest: core-powered mass-loss dominates for lower-gravity planets with high equilibrium temperature, and the opposite for photoevaporation; planets can transition between the two mass-loss methods, but a planet stripped by core-powered mass-loss probably never experienced photoevaporation, while a planet stripped by photoevaporation could have experience early core-powered mass-loss; and photoevaporation may be responsible for the final form of the radius valley. Affolter et al. (2023) find the two methods work together to produce a slightly shallower slope that agrees well with observations.

There are yet further proposals for how the radius gap comes to be which do not require mass loss. Zeng et al. (2019) and Venturini et al. (2020) interpret it as a compositional dichotomy in planetary cores, where planets below the gap are rocky terrestrial cores, and planets above the gap are a mixture of rock and water. This is, however, difficult to square against the position of the gap changing depending on orbital period and stellar mass (Berger et al., 2020). It could also be due to primordial gas accretion: planets below the gap are light cores that cannot accrete a sufficient amount of gas and so remain as rocky cores due

to isothermal cooling limits; planets above the gap are cores that have assembled late when the gas in the disk is depleted and undergo late-stage (gas-poor) accretion (Lee et al., 2022). This agrees with the shape of the valley in Van Eylen et al. (2018), reproduces the rise in gap radius with stellar mass in (Berger et al., 2020), and does not preclude post-formation mass loss.

Further work is needed to determine if there even is a dominant mechanism behind the gap, or whether it is sculpted by several working in tandem.

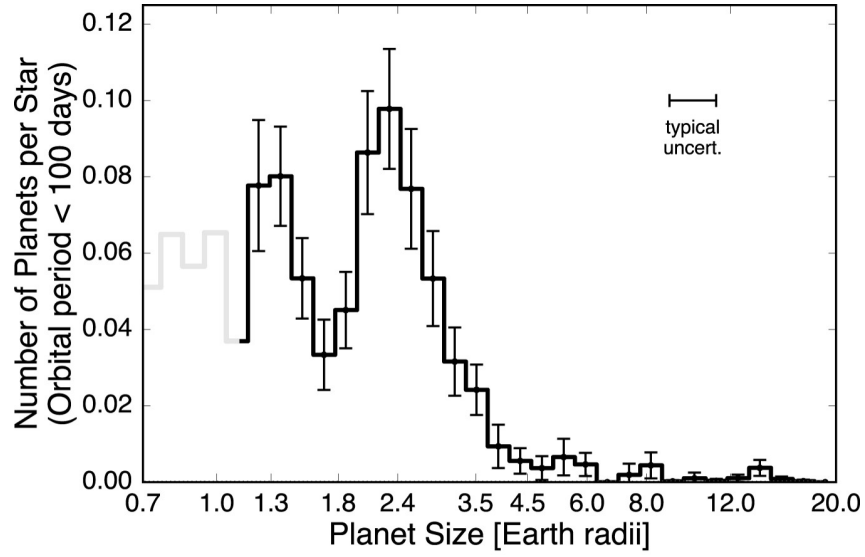


Figure 1.20: A completeness-corrected histogram of planets from the CKS survey, showing the distribution of planet radii for planets with orbital periods less than 100 days. There is a clear valley around $\sim 1.8R_{\oplus}$. Reproduced from Fulton et al. (2017).

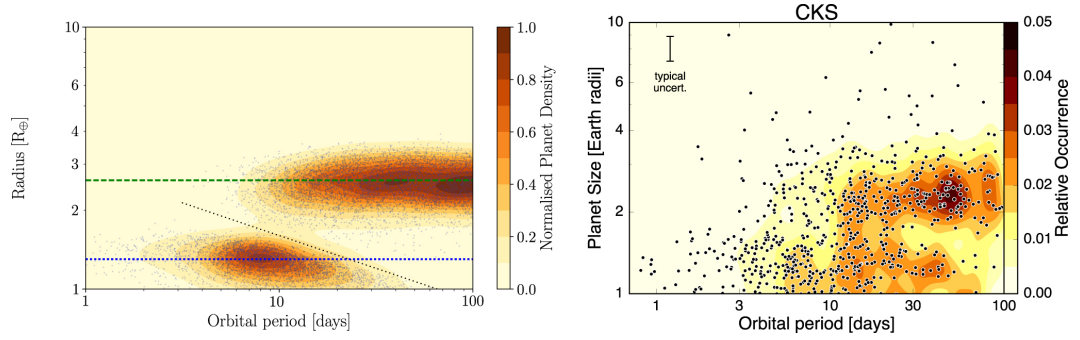


Figure 1.21: Left: the period-radius distribution of planets predicted from the evolution models of Owen and Wu (2017). The dashed horizontal lines show the positions of the two peaks in the radii distribution. The dotted black line demarcates the lower boundary of the valley, showing their analytic result for the largest planet as a function of period that can be stripped bare by photoevaporation. Right: the observed planet distribution from Fulton et al. (2017), which broadly matches the model on the left. Reproduced from Owen and Wu (2017).

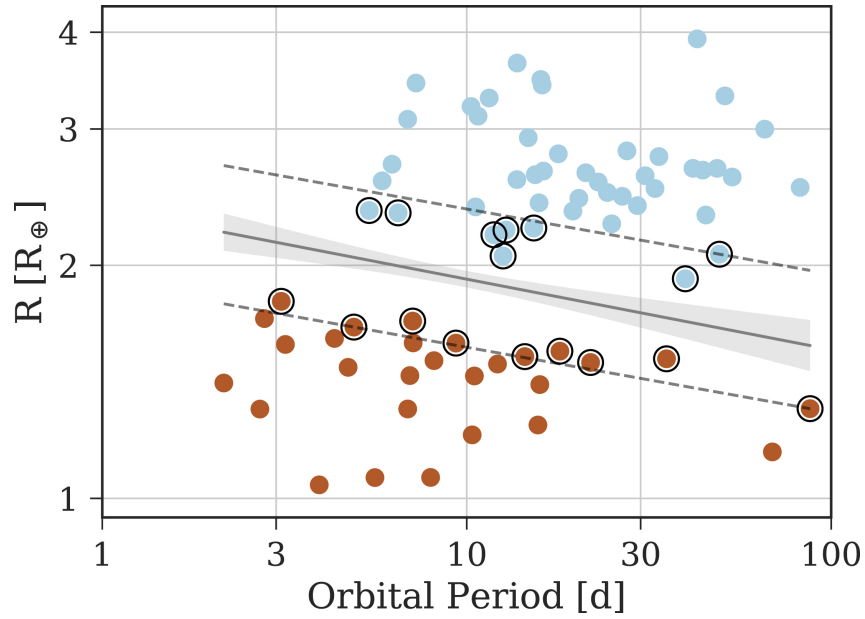


Figure 1.22: The slope of the radius valley using a sample where stellar parameters have been determined from asteroseismology. The super-Earths below the valley are coloured red; the sub-Neptunes above the valley are coloured blue. Reproduced from Van Eylen et al. (2018).

1.4.2 Occurrence rates

Similarly to the discussion of population features in Section 1.4.1, looking at the occurrence rates of planets can be used to infer formation and evolution channels.

At its basis, the average number of planets per star is the occurrence rate. But it is also a much more nuanced subject than that. Occurrence rates for planets have been found to differ based on their size, as have the occurrence rates of planets around different kinds of stars. Exoplanet scientists have long been interested in “eta Earth” (η_{\oplus}), the mean number per star of Earth-sized rocky planets within the optimistic habitable zone (Haghighipour, 2015).

In order to calculate occurrence rates, first you must have observed enough stars, then you must search the observations consistently for planets. Then, corrections for biases and completeness need to be made. Exoplanet survey catalogues are not complete, as not every planet will be detected. As discussed in Section 1.2, our various detection methods each suffer from biases; for example, the transit method only finds planets near an inclination of 90 deg. Nor are the catalogues perfectly reliable, due to false positives. A brief overview of occurrence rate studies for different planet surveys are discussed below.

The first occurrence rate studies came from early radial velocity surveys. For example, Mayor et al. (2011) used a volume-limited sample of 376 stars observed over an 8 year HARPS and CORALIE program. They found 50 per cent of solar-type stars host planets $< 30M_{\oplus}$ with periods less than 100 days, and that gas giant planets are significantly less common and their occurrence rate correlates with host star metallicity.

Then came *Kepler*, with its primary goal being to study exoplanet populations (Borucki et al., 2010). Some early analysis was performed by (Fressin et al., 2013): they quantified false positive rates, and the distribution and rate of planet occurrence down to Earth-size. They find a global false positive rate of 9.4 per cent, and that 16.5 per cent of FGK stars have at least one Earth-sized planet with orbital periods ≤ 85 days. Their occurrence rate of planets of different sizes around FGKM stars is shown in Fig. 1.23. Hsu et al. (2019) updated this with the full *Kepler* DR25 dataset and *Gaia* DR2 data (Gaia Collaboration et al., 2018), and their results are shown in Fig. 1.24.

Looking at a different orbital period regime, Poleski et al. (2021) analyses 20 years of OGLE microlensing survey data to find the occurrence rate of wide-orbit ice-giants. They find that, on average, every microlensing star hosts 1.4 ice giants, implying they are very common.

Zhou et al. (2019) performed the first occurrence rate study on the *TESS* data, specifically looking at hot Jupiters around A, F, and G stars. They find good agreement with the *Kepler* hot Jupiter occurrence rate, and that from Mayor et al. (2011), all around 0.4 per cent, and they find no dependence on stellar mass. Gan et al. (2023) also look at hot

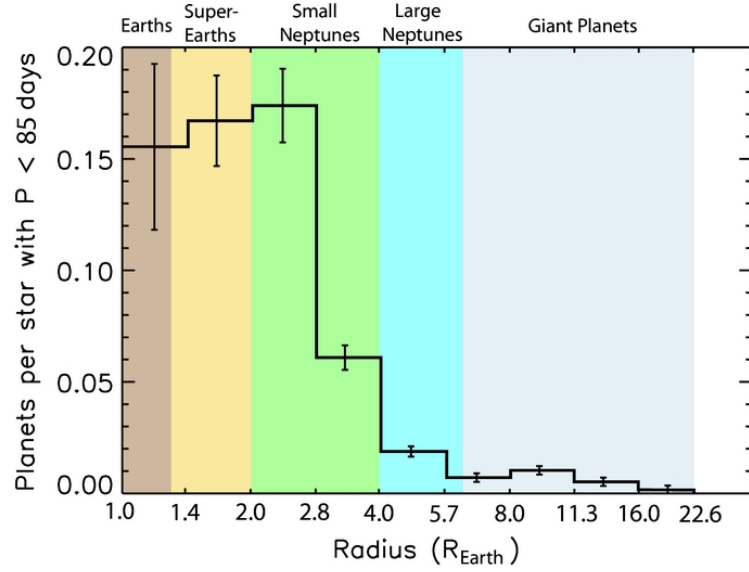


Figure 1.23: The average number of planets of increasing radii around main-sequence FGKM stars from *Kepler*, corrected for false positives and incompleteness. Reproduced from Fressin et al. (2013).

Jupiters from *TESS*, but around early M-dwarfs, finding an occurrence rate of 0.27 per cent, smaller than measurements for FGK stars but consistent within a few sigma. More recently, Bryant et al. (2023) goes to even lower-mass stars, looking at giant planet occurrence and finding a rate of 0.19 per cent for the full sample, and surprisingly finding the occurrence rate around stars less than $0.4M_{\odot}$ is non-zero. There is still plenty of work to be done on occurrence rates from *TESS* as it progresses through its extended missions.

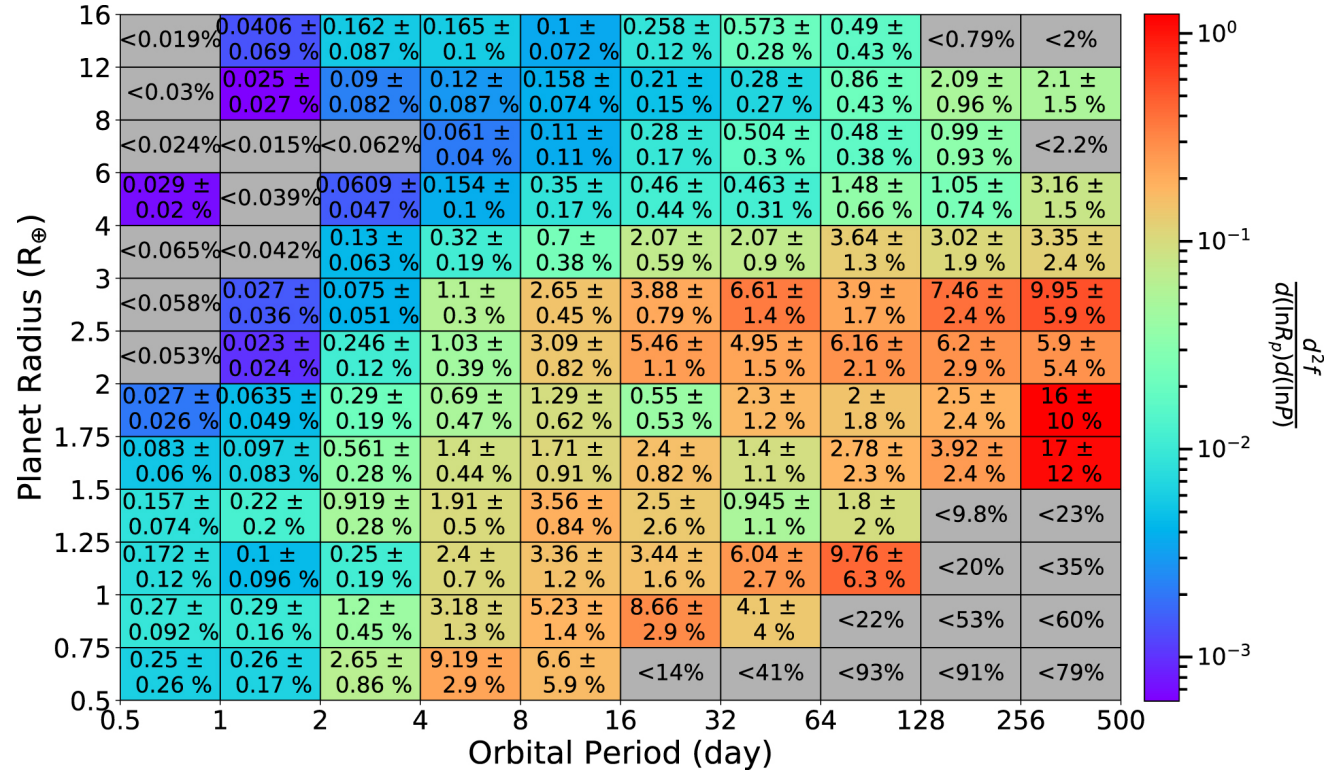


Figure 1.24: Occurrence rates in period-radius bins for the *Kepler* DR25 planets around FGK stars, corrected for detection and vetting efficiency. Reproduced from Hsu et al. (2019).

1.4.3 Formation and evolution

It is important to understand theories of planet formation and evolution when studying planetary population features like the Neptunian desert, as these population features may be reflective of these mechanisms influencing what exoplanets can exist. There are also occasional discoveries of unusual planets, such as TOI-332 b (Section 5), which challenge these theories.

The protoplanetary disc

A molecular cloud of dust and gas can collapse under its own gravity to form a protostar spinning with the angular momentum of the nebula (Shu et al., 1987). As the material around the protostar contracts it flattens and forms a protostellar disc, due to conservation of angular momentum (Safronov, 1972). When the protostar reaches a surface temperature similar to that of a main sequence star and becomes visible, it has become a T Tauri star, a pre-main sequence star in the process of contracting onto the main sequence (Herbig, 1962). The disc is now a protoplanetary disc, made of 99 per cent gas and 1 per cent dust grains and ice particles (Williams and Cieza, 2011). It absorbs and re-emits the radiation from its star; the gas component is hard to observe due to it only emitting at certain wavelengths, but the dust emits in infrared and sub-millimetre wavelengths and scatters light which can both be imaged, showing the dust in a rotating, disc-like structure (e.g., Smith and Terrile, 1984), with a “flaring” vertical height that increases with radius (Kenyon and Hartmann, 1987).

This is where planets are born. Solids orbit the star at Keplerian speed, but gas orbits slightly slower due to a radial gradient in gas pressure that supports it against the central stellar gravity; stellar gravity also causes dust grains to settle into the disc mid-plane (Weidenschilling, 1980). Material in the disc is accreted onto the star due to angular momentum transport (Papaloizou and Lin, 1995), and the disc thus loses mass due to this and also via photoevaporation (Gorti and Hollenbach, 2009) over perhaps 10 Myr (Hillenbrand, 2008) until it is depleted. Indeed, most discs do not show infrared excess after 3 Myr (Mamajek, 2009). The lifetime of the disc is the principle constraint on planet formation within it.

Creating planets

There are two theories for planet formation in the protoplanetary disc: gravitational instability (Boss, 1997), and core accretion (Pollack et al., 1996).

In the gravitational instability model, the solar nebula is fragmented due to gravitational instabilities. These clumps of gas and dust then contract to eventually form giant planets. In pre-collapse, the extended, cold protoplanet contracts quasi-statically, the speed

of which depends on the initial mass, i.e. more massive protoplanets evolve faster (Helled and Bodenheimer, 2011). When molecular hydrogen starts to dissociate at the centre, hydrodynamic collapse starts (Bodenheimer et al., 1980). The initial composition would be the same as the young disc, therefore stellar, but the planet may be subsequently enriched in heavy elements via planetesimal capture (Helled and Schubert, 2008). Newly formed protoplanets would have no cores, but they could form by sedimentation of silicate grains (Helled et al., 2008). Final core mass depends on available solids in the planetary envelope, and if enough solids can be accreted before the temperature becomes hot enough to evaporate them, several Earth mass cores can be made (Helled and Schubert, 2008). Ice and inorganic material entering the envelope could also settle into the core.

In core accretion, dust grains accrete, growing from grains to dust aggregates to pebbles, eventually forming solid planetesimals that then merge into a solid core composed of rocks and ices. Coalescence is faster beyond the snow-line, where there is more solid material (Kennedy and Kenyon, 2008). Solid accretion slows when planetesimals in the “feeding zone” of the forming planet are depleted. The core can then go on to accrete gas, forming an envelope; if the core is massive enough (10-20 Earth masses), it undergoes runaway gas accretion and becomes a gas giant with a large envelope (Pollack et al., 1996; Lambrechts et al., 2014; Piso et al., 2015). Gas accretion stops when the nebular gas dissipates, or when it opens a gap in the disc (Lissauer et al., 2009), isolating the planet from the disc, and then it contracts quasi-statically and cools over $\sim 10^9$ years (Lissauer and Stevenson, 2007). Final core mass is proportional to orbital distance (isolation mass increases with distance), and the availability of solids (Lissauer, 1987; Helled and Schubert, 2009) (stellar metallicity is correlated with presence of giant planets, see e.g., Fischer and Valenti (2005); Adibekyan (2019)). The composition depends on the disc properties, and in the case of gas giants, the dust within the accreted gas, and any planetesimals accreted after rapid gas accretion.

The timescale of core accretion was thought to be at odds with the formation of gas giants; the timescale of formation of 10 Earth-mass cores was calculated to be on the order of 10^6 years (Lissauer, 1987), which is the time at which a young solar-type star would be losing its gas disc (Shu et al., 1993). If the gas disc is already dissipating, then Uranus-like planets would be produced, rather than gas giants. This, plus interior modelling leading to predicted core masses of Saturn and Jupiter being smaller than expected (Chabrier et al., 1992), led to the theory of gravitational instability to be posited. However, increased understanding of the timescales of core accretion has now resolved this issue (Lambrechts and Johansen, 2012).

Furthermore, core accretion and gravitational instability do not need to be in tension with each other. Gravitational instability can form gas giants on wide orbits, like in the

HR 8799 system (Dodson-Robinson et al., 2009), whereas core accretion is the favoured scenario for terrestrial planets (while also relying on the gas giants in the same system, see e.g., Morbidelli et al. (2012); Izidoro and Raymond (2018)).

Migration

There are some planets, however, that defy explanation by the formation methods above. Extremely close-in planets, like hot Jupiters, would not have had enough planet-forming material to have formed “in-situ”. Planets have also been found with unusual orbits, e.g., highly eccentric (Wu and Murray, 2003), and retrograde (Winn et al., 2009) orbits. This all implies that, between their formation and now, planets must have undergone migration.

There are many migration pathways, which can be broadly categorised first based on whether or not they occur when the protoplanetary disc still exists. The former is disc migration, the latter is perturbation from another body.

Disc migration can be split into two categories, dependent upon the mass of the planet. Type I migration (Ward, 1986) occurs for planets less than 0.1 Jupiter masses, when Lindblad resonances excite spiral density waves in the surrounding gas. The wave exterior to the planet’s orbit exerts a greater torque than the one interior, and the planet migrates towards its star. Type II migration (Lin and Papaloizou, 1986) occurs when a planet is massive enough to deplete the region around its orbit and open a gap in the disc. As the gas accretes onto the star, the gap and the planet embedded within it are carried inwards. A transition from Type I to Type II can occur if the planet grows large enough. Type II migration, however, may not be realistic – it has been found that, unless there are extreme thermal or viscosity conditions, gas flows through the gap (Lubow and D’Angelo, 2006). Therefore, torques on the planet are more akin to Type I torques, and it may be called modified Type I migration (D’Angelo and Lubow, 2008). Additionally, Type II migration can cause eccentricity growth, even if an orbit starts circular (D’Angelo et al., 2006).

Migration can alternatively occur when angular motion is exchanged with a perturbing body. A classic example of this is the Kozai-Lidov effect (Kozai, 1962; Lidov, 1962). Interaction with a massive perturber exterior to an inner, lower-mass planet’s orbit, like a gas giant or star, causes the inner planet to exchange eccentricity and inclination, increasing the planet’s eccentricity and shrinking its perihelion enough to create tides between it and its star. The orbit will shrink, eventually enough to exit the influence of the Kozai-Lidov effect, and will then become tidally circularised. In this process, an orbit may also become retrograde (Naoz et al., 2011).

As discussed in Section 1.18, observations, like of the Rossiter-McLaughlin effect, may be able to differentiate between these scenarios. Disc-driven migration is thought to preserve alignment with the stellar spin axis and retain circularity, while perturbation can

cause inclined, eccentric orbits (unless the planet is subsequently tidally circularised).

Finding unusual planets that defy expectations from formation and migration mechanisms, like in Section 5, can help us test these scenarios and propose new mechanisms that may be responsible for their existence.

1.5 Thesis outline

This thesis presents work characterising exoplanetary systems primarily with the Transiting Exoplanet Survey Satellite, *TESS*, and the High Accuracy Radial velocity Planet Searcher, HARPS. This begins in Chapter 2 by describing how the instruments work and how to analyse the data they produce. This is then put into practise in the science Chapters 3 through 5. Chapter 3 presents the detailed analysis and characterisation of the three-planet system TOI-431. Chapter 4 introduces the Nomads HARPS observing programme, with the goal of obtaining precise masses for many planets within the Neptunian desert. It details the sample construction, first-look analysis of the targets observed, and places them in the context of the desert. Chapter 5 presents the detailed analysis and characterisation of a particularly intriguing planet that was observed as part of the Nomads programme, TOI-332 b. The thesis concludes with a summary and future outlook in Chapter 6.

Chapter 2

Methods

In this Chapter, I describe the two main observation facilities used in this thesis, the Transiting Exoplanet Survey Satellite (*TESS*), and the High Accuracy Radial velocity Planet Explorer (HARPS), describing their science goals, design, and data. As the focus of Chapter 4 is performing a large follow-up programme on HARPS, I then go into the details of scheduling observations and analysing radial velocity data. Finally, I give an overview of joint-fit modelling methods for exoplanetary systems, as this is the basis of both Chapters 3 and 5.

2.1 Transit detection with *TESS*

The planets explored in Chapters 3, 4, and 5 are all *TESS* Objects of Interest. Here, I describe the mission which discovered them.

2.1.1 About *TESS*

Primary goals

Transiting exoplanets are valuable discoveries – as discussed in Sections 1.2.1 and 1.3.1, from transits we can determine important system parameters like orbital period and planetary radius, and further follow-up allows us to then determine planetary mass, and interior and atmospheric compositions. Still, relatively little is known about planets intermediate to the Earth and Neptune. One, because they have no Solar System analogues. Two, because while *Kepler* was revolutionary to exoplanet science and found many planets within this regime, it has been limited by the majority of its observed stars being too faint for detailed followup like precise RVs and ground-based transit detections.

Enter the Transiting Exoplanet Survey Satellite, *TESS*: its primary mission goal is to find hundreds of transiting planets smaller than Neptune ($R \leq 4 R_{\oplus}$) whose host stars are

bright enough for spectroscopic follow-up. The *TESS* mission primary science requirements are to (Stassun et al., 2019):

1. Search $> 200\,000$ stars for planets with $P < 10$ days and $R < 2.5R_{\oplus}$;
2. Search the 10 000 stars in the ecliptic pole regions for planets with $P \leq 120$ days and $R < 2.5R_{\oplus}$;
3. Determine masses for ≥ 50 planets with $R < 4R_{\oplus}$.

Thus the design and strategy of the mission needs to fulfil these goals.

Design considerations

The *Kepler* transit detection rate for super-Earths with periods less than 10 days was 0.2 per cent (Borucki et al., 2011b). So, to detect one super-Earth, one would need to monitor around 500 stars. In order for *TESS* to detect hundreds of these, it needs to monitor $\geq 10^5$ stars – and these must be bright, in order to enable follow-up. Such is the need for an all-sky survey.

TESS focuses on main-sequence dwarfs of spectral types F5 to M5. Large stars are harder to detect transiting planets around, and early spectral types have rapid rotation which broadens spectral lines and makes precise radial velocity follow-up harder. M-dwarfs present an interesting science case: they are small, making transits easier to detect; they are very abundant; and they did not feature heavily in the *Kepler* target list. However, stars with a later type than M5 are optically faint.

The aperture size of each *TESS* camera is 10 cm, chosen to ensure the precision needed for super-Earth detection was met. The bandpass is from 600-1000 nm, chosen to increase sensitivity in the red for observations of the small red stars that make exoplanet detection easier, and with a wide range to reduce photon-counting noise.

The choice of maximum observable orbital period, P_{\max} , determines the length of the mission, and therefore has an excessive influence on the cost. Additionally, as discussed in Section 1.2.1, the transit method is already strongly biased to short period planets. Having sectors of 27.4 days allows two detections of a 10 d period planet; while a third is generally needed to secure the orbital period, this is either likely to happen in a subsequent reobserving of the sector, or via ground-based follow-up. Having a $P_{\max} \sim 10$ days does, however, mean that *TESS* is not optimal for detect transiting planets in the habitable zones of Sun-like stars. Fortunately, with a $P_{\max} \sim 40$ days, habitable zone planets around mid M-dwarfs become accessible, and this is possible at the areas of more continuous viewing like the ecliptic poles (Ricker et al., 2015). Extending the mission beyond the nominal 2 year prime mission also

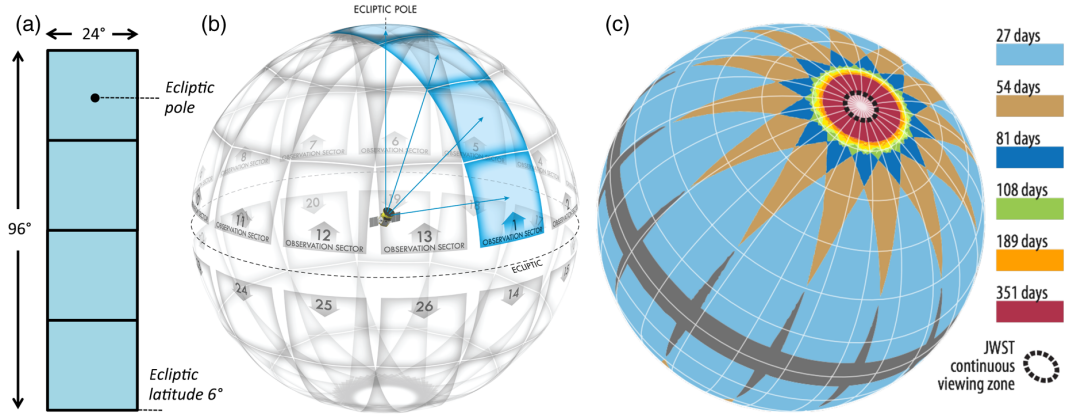


Figure 2.1: (a) The field of view of the four *TESS* cameras. (b) The *TESS* sectors projected onto the celestial sphere, 13 per hemisphere. (c) the duration of observations on the celestial sphere in one full cycle, where regions with sector overlap are observed for longer. The celestial poles are a continuous viewing zone, aligning with the *JWST* continuous viewing zone (black dotted circle). Reproduced from Ricker et al. (2015).

increases P_{\max} , and indeed, the mission has had two extensions thus far and stands a good chance of a third (see below on the primary and extended missions).

Observations

TESS has a simple stare and step strategy. The four square fields (of 24 deg \times 24 deg), one per camera, combine together to make a 1 \times 4 array, or a sector (covering 24 deg \times 96 deg, or 2300 square degrees). An observation orbit, where the camera array is pointed in the anti-solar direction, is ~ 13.7 d; directly after there is a gap where the data is downlinked. Each sector is observed for two observation orbits, totalling ~ 27.4 days. The telescope then rotates eastward to move on to the next sector, and observes almost a whole hemisphere on the sky in a total of 13 sectors, equal to 26 orbits or close to a year. The whole-sky survey therefore takes 2 years to complete, the length of the nominal mission. There is overlap between adjacent sectors, resulting in certain stars being observed for much longer than the nominal length of one sector, with the ecliptic poles being continuously viewed (aside from downlink time). The *JWST* continuous viewing zones are also located at the ecliptic poles. The viewing geometry is represented in Fig. 2.1.

To have such a wide field-of-view, the *TESS* cameras have a large pixel size of 21 arcsec per pixel. This causes blending of nearby stars, which has to be corrected in the reduction pipeline using knowledge of stellar positions from *Gaia* DR2. However, this does not account for issues caused by eclipsing binaries, see Section 2.1.2.

Primary and extended missions

The primary mission ran for the first two years (cycle 1 and 2), ending in July 2020. Cycle 1 (sectors 1-13) observed the southern hemisphere. Cycle 2 (sectors 14-26) aimed to observe the northern hemisphere; however, the pointing had to be shifted during sectors 14-16 and 24-26 to avoid over-contamination by moonlight. Then *TESS* entered its first extended mission, lasting another two years (cycles 3 and 4) and completed in September 2022. This and subsequent missions aim to fill in the gaps that were not observed, e.g., between the CCDs and along the ecliptic. Cycle 3 (sectors 27-39) re-observed the southern ecliptic hemisphere. Cycle 4 (sectors 40-55) changed tack, observing 16 sectors rather than 13 (lasting 15 months) which were a mix of northern hemisphere and ecliptic pointings. *TESS* is now in its second extended mission, set to last 3 years (cycles 5, 6, and 7). Cycle 5 (sectors 56-69) will complete the second scan of the northern hemisphere, started in cycle 4, and start a new scan of the southern hemisphere. The pointings for cycles 6 and 7 are not yet confirmed. A visual representation of all of the sectors until the end of cycle 5 is shown in Fig. 2.2

Data

From all the sources that *TESS* could observe which populate the *TESS* Input Catalog (TIC), a smaller catalogue of stars (the Candidate Target List, or CTL) is composed from bright stars that are likely to be dwarfs within the temperature range of interest, taking extra care to include M-dwarfs. From within the CTL, stars with a smaller radius, better expected photometric precision, and that are seen in more sectors are prioritised (Stassun et al., 2019). These are the stars that are amenable for detection of small planets, observations of which are more likely to fulfil the primary science requirements. In the primary mission, $\sim 15\,000$ of the high priority stars were observed at a high cadence of 2 mins per sector. In the first extended mission, an additional ~ 600 stars per sector were observed at 20 second cadence as well as $\sim 15\,000$ 2 min cadence targets. In the second extended mission, it is a split of $\sim 2\,000$ stars at 20 second cadence, and $\sim 8\,000$ at 2 mins. The data is downlinked as a postage stamp, or Target Pixel File (TPF), around each target star. The TPF is the pixels contained within a predefined mask around a star, and these are packaged as a time series of images in a binary FITS table, and used to create the data found in the light curve files.

The data reduction and light curve production is performed by the *TESS* Science Processing Operations Centre (SPOC, Jenkins et al., 2016). This includes several steps. First, they perform calibration to remove instrumental effects and other polluting artefacts. Then they find optimal apertures for each target star's postage stamp to use for extracting photometry. The photometric analysis is then performed, including removing background

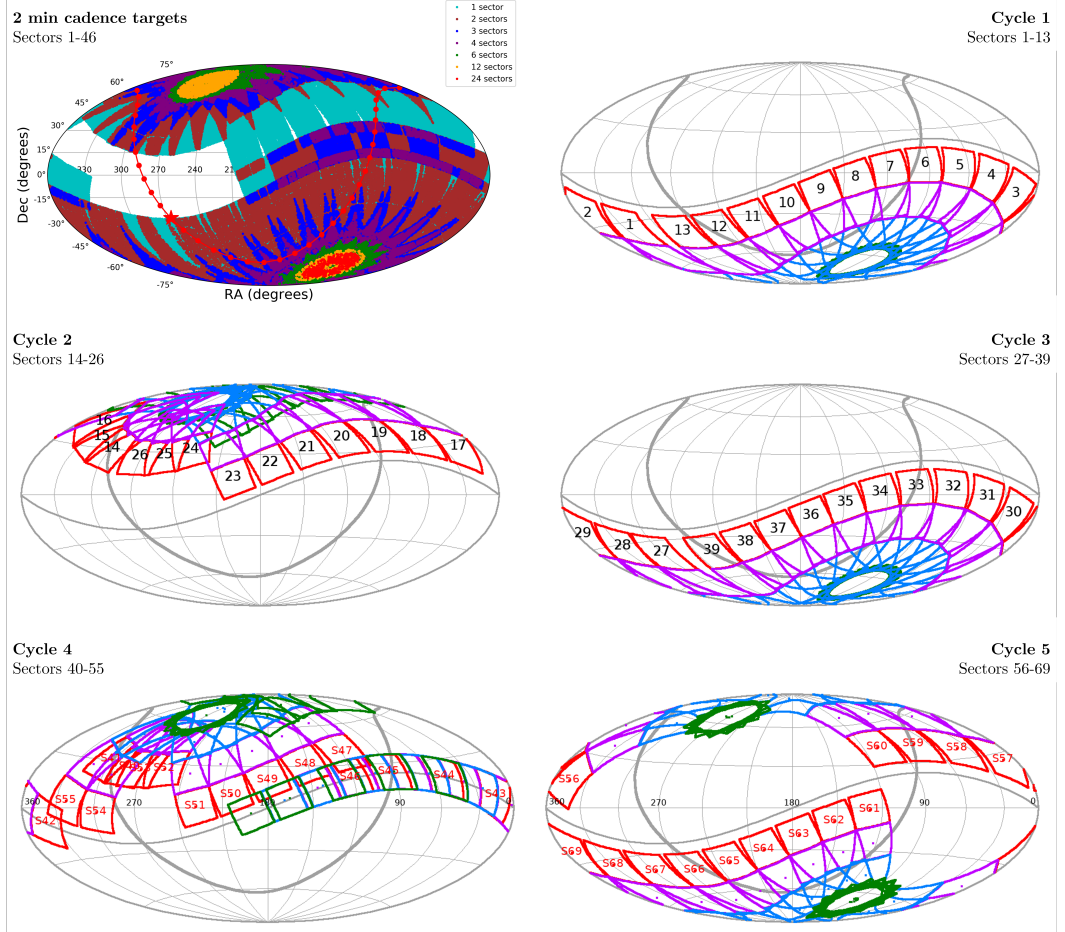


Figure 2.2: Sectors observed by *TESS* over the first five cycles of its operation, shown in ecliptic coordinates. The top left picture shows the total viewing time for all 2 min cadence targets within sectors 1-46. Reproduced from the MIT *TESS* website at <https://tess.mit.edu/observations/>, accessed on 20 Sep 2023.

flux and cosmic rays, measuring the centroid of each target star, and PSF fitting to measure the pointing and focus of the cameras. At this stage, Simple Aperture Photometry (SAP) light curves are produced, one of the final data products. Next, Presearch Data Conditioning (PDC) performs important corrections to the SAP light curves, removing instrumental effects and discontinuities from radiation events, removing isolated outliers, and finally correcting the flux time series for crowding and flux lost outside the photometric aperture. PDC_SAP light curves are produced, another final data product. These are subject to a transit search, and Threshold Crossing Events (TCEs) are fit with a limb-darkened transit light curve model and passed on to the *TESS* Science Office (TSO), where they are vetted and released as *TESS* Objects of Interest (TOIs) (Guerrero et al., 2021). Finally, SPOC also releases Data Validation Reports (DVRs), which contain the results of diagnostic tests on the transiting planet signatures that, for example, compare even and odd transits, and check difference image centroiding to rule out background sources.

Full Frame Images (FFIs), which contain the full set of science and collateral pixels across all the CCDs of a given camera, are also downlinked. The cadence of the FFIs was 30 mins in the primary mission, 10 mins in the first extended mission, and now 200 seconds in the second extended mission. Several pipelines exist to produce light curves from the FFIs, e.g., the MIT Quick Look Pipeline (QLP, Huang et al., 2020), and *TESS*-SPOC (Caldwell et al., 2020). While the goal of *TESS* is not to perform population studies, as the targets are not selected in the FFIs like they are for the higher cadence observations (rather, all sources are observed), a uniform search of light curves generated from these would enable such studies to happen.

2.1.2 Confirmation of transit signals

Follow-up of planet candidates is an integral part of the *TESS* mission. To this end, the *TESS* Follow-up Observing Program (TFOP, Collins et al., 2018) has been set up, a global network of people and follow-up facilities. TFOP is split into five Sub Groups (SGs). SG1-3 are focused on identifying false positives; SG4-5 are for further characterising strong candidates.

1. **SG1: ground-based photometry.** This aims to confirm *TESS* transit candidates as “real” by confirming whether the transit signal occurs on target, or whether another star is responsible for the signal. This is especially important given the aforementioned large pixel scale, which causes source blending. It is also used to obtain further transits of a system, which can be useful for constraining the planetary period and radius.
2. **SG2: reconnaissance spectroscopy.** This is to identify whether the signal is caused by a binary star system, in which case the radial velocity variation will be much

larger than expected from a planet or drift. Precision RVs are not needed for this work; identifying stellar binaries before targets are observed with high precision saves valuable time on expensive facilities. Additionally, a stellar rotation rate can be obtained, which is important to estimate stellar activity and determine feasibility of precision RVs.

3. **SG3: high-resolution imaging.** This is used to search for nearby contaminating sources that can contaminate the photometry or be the source of astrophysical false positives, such as background eclipsing binaries.
4. **SG4: precision radial velocities.** These confirm the planetary nature of a candidate, and provide a mass measurement. They can also be used to pin down stellar activity. The Nomads programme (Section 4) is part of SG4, and precision radial velocity follow-up with HARPS is described in the subsequent Section 2.2.
5. **SG5: space-based photometry.** This provides improved light curves, used to bolster ephemerides and radius determinations and monitor TTVs.

2.2 Radial velocity follow-up with HARPS

We follow up all the planets in Chapters 3, 4, and 5 with the HARPS spectrograph in order to determine their masses. Here, I describe how the spectrograph works, how to schedule observations with it, and how to analyse the data obtained from it.

2.2.1 About HARPS

The instrument

The High Accuracy Radial velocity Planet Searcher (HARPS) is located at La Silla Observatory, on the edge of the Atacama desert in Chile, at an elevation of 2 400 m. It is mounted on the cassegrain focus of the 3.6 m telescope, and operated by the European Southern Observatory (ESO). It is a spectrograph designed to achieve a high precision of 1 m s^{-1} (Pepe et al., 2002, 2018; La Silla Observatory, 2019).

HARPS is fibre-fed, which eliminates the problem of illumination of a spectral slit and collimator. The fibre is also smaller than a slit, with a circular field of view with a diameter 1 arcsec on the sky. It uses a simultaneous reference via a second fibre to track instrumental drift, where one fibre is pointed at the target star, and the other is on the reference (Baranne et al., 1996). Differential and varying illumination of the two fibres is mitigated using a double scrambler (see e.g., Halverson et al., 2015), and is also reduced by using a small fibre and stable guiding.

It is a cross-dispersed echelle spectrograph, which uses two steps to increase spectral resolution. First, it passes light onto an echelle grating, a type of diffraction grating which has increased dispersion of spectral features. Then, it uses a grism (grating prism) at 90 degrees to the echelle grating to separate the overlapping orders. This results in a 2D spectrum with increased wavelength separation, which is beamed onto a mosaic of two CCDs. The full wavelength range is 380-690 nm.

The instrument has a resolving power $R = 115\,000$, a dimensionless quality that defines the ability of a spectrograph to distinguish between two spectral lines with minimal separation, $\Delta\lambda$, at a certain wavelength, λ , i.e. $R = \lambda/\Delta\lambda$. Spectral resolution is exceedingly important for a high-precision spectrograph (e.g., Bouchy et al., 2001). Calibration uses emission and absorption lines – the photocenter of a line is determined with a precision that increases with increasing resolving power, and the line depth (or signal) also increases with resolving power.

The intrinsic stability of the spectrograph is tackled on several fronts: the whole spectrograph is inside a vacuum in order to avoid air pressure changes; the temperature is kept as stable as possible to avoid thermomechanical effects; and the mechanical design avoids moving parts and uses optimal materials.

The data reduction pipeline

The HARPS Data Reduction Software (DRS) implements the scientific reduction of the raw data, and extracts radial velocities (Bouchy and Queloz, 2003). As soon as the exposure is complete, the DRS runs automatically. Here I focus on the case of classical fibre spectroscopy, which is used for all observations in Sections 3, 4, and 5. The DRS applies corrections for detector bias, dark, flat field, and cosmic ray removal. Sky-background correction is left to the user, as is correcting for instrumental drift. It does spectrum extraction and calculates the RVs according to the method presented in Baranne et al. (1996), using a template spectrum out of a choice of G2V, K5V, or M2V (La Silla Observatory, 2019). Spectra, the cross-correlation function, and RVs are provided as data products to the end user, accessible via the ESO archive¹. Raw data is also available², on which users may wish to use their own reduction pipelines. Data is proprietary for one year after the observation is taken.

2.2.2 Scheduling a HARPS follow-up program

Pointing restrictions

To explain where HARPS can point on the sky, first the equatorial coordinate system needs to be defined, shown in Fig. 2.3. Objects in the sky are on the celestial sphere. Declination, δ , is the equivalent of latitude and is measured in degrees north or south of the celestial equator, i.e. from +90 deg at the north celestial pole to -90 deg at the south celestial pole. Right ascension, RA or α , is the equivalent of longitude and is measured in hours, minutes, and seconds eastward of the vernal equinox to its intersection with the object in question's hour circle, where the hour circle is the circle on the celestial sphere drawn through the object and both celestial poles. 24 hours of right ascension is 360 degrees; 1 hour is 15 degrees. Finally, the hour angle, HA or H , is the angle between a celestial object and the observer's meridian. The zenith is also defined, the point on the celestial sphere directly above the observer, and the zenith distance, ZD, the angular distance of a celestial body from the zenith. This is equivalent to 90 degrees minus the body's altitude above the horizon, and so can also be named the coaltitude.

The primary constraints on what HARPS can view are as follows³:

1. HARPS can point 5 h 30 m either side of its meridian, i.e. at objects that satisfy

¹The ESO archive Phase 3 reduced data is available at http://archive.eso.org/wdb/wdb/adp/phase3_main/form

²The ESO archive raw data is available at http://archive.eso.org/eso/eso_archive_main.html

³Detailed at <https://www.eso.org/sci/facilities/lasilla/telescopes/3p6/overview.html>, accessed 20 September 2023.

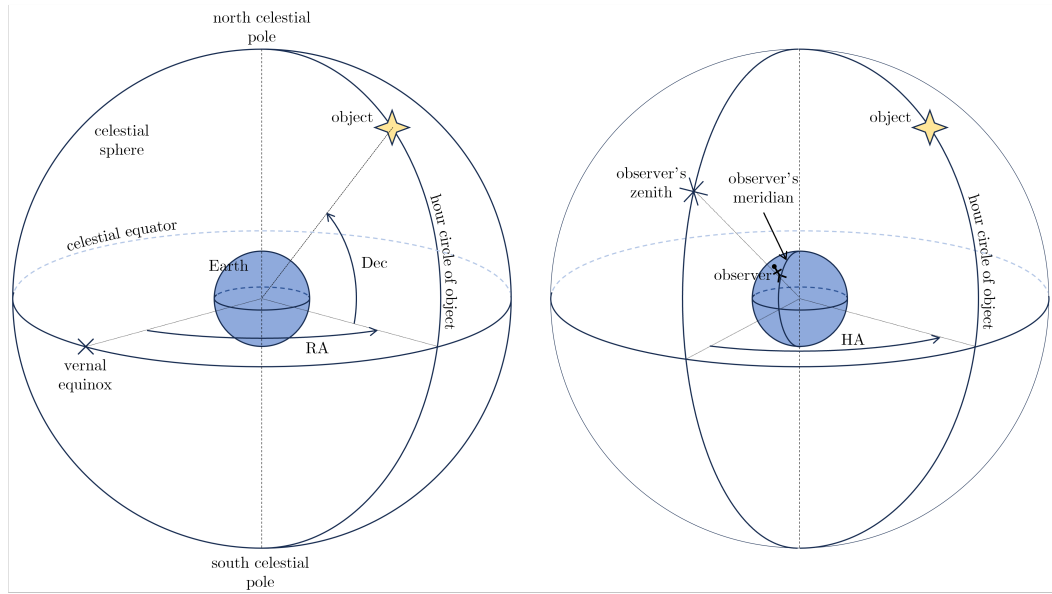


Figure 2.3: A diagram of the equatorial coordinate system. Left shows the Right Ascension (RA) and Declination (Dec) of an object in the sky, and right shows its Hour Angle (HA).

$-5 \text{ h } 30 \text{ m} < \text{HA} < +5 \text{ h } 30 \text{ m}$. So, when an object at zenith has an RA of 10 h, HARPS can observe stars with RAs between 4 h 30 m and 15 h 30 m.

2. HARPS can *theoretically* access declinations from -120 degrees to +29.5 degrees. The former represents an accessible sky region below the southern pole that is 12 h away in hour angle with respect to 1. In *practice*, this is limited by airmass, i.e. angular distance from the zenith (discussed below). Note that HARPS also cannot point within 3 degrees of the pole.
3. HARPS is limited to a $\text{ZD} < 70$ degrees. It can point at the zenith, due to its equatorial mount, but cannot point within 20 degrees of the horizon.

There are a few further concerns when choosing where to point the telescope. Air mass quantifies the amount of air along the line of sight to the object being observed when the observer is within the Earth's atmosphere. It is measured in relative terms – an object at zenith has a minimal air mass of 1, and this increases with zenith distance. At a zenith angle of 60 degrees, the air mass is ~ 2 . It is not advisable to observe at air mass > 1.8 , as the greater the air mass, the more the light of the object is attenuated (scattered and absorbed) by the atmosphere, which negatively affects the quality of observations.

The separation of the moon from the target is also critical when the moon is bright in the sky. Light from the moon can contaminate the observation, even more so if there is

scattered light due to clouds. As a general rule of thumb, if the target is within 30 degrees of the moon, it should not be observed.

Finally, the target must itself be of a suitable magnitude. A target of higher magnitude is going to have a lower signal-to-noise, and require a longer exposure time to achieve the same RV precision. This is further explained in Section 2.2.2.

Weather restrictions

For the safety of the telescope, the telescope dome is closed in adverse weather conditions. This includes⁴: a wind speed $> 20 \text{ m s}^{-1}$; humidity > 80 per cent; a temperature within 3 degrees of the dew point; when there is dew on the dome; and when the sky is overcast with a chance of rain. Cloud cover can also make it impossible to observe a target. If the adverse weather conditions are subsequently within operating limits for 30 minutes, the dome may reopen.

Turbulence in the atmosphere also degrades the image of an object, which can be seen as blurring, twinkling (of a point source like a star) or distortion, as the refractive index along the line of sight rapidly changes. The quality of the image at a particular location at a particular time is the seeing. The seeing is measured as the diameter (also known as the full width half maximum, FWHM) of the seeing disc of an object, a shape caused by the averaging out of the rapid distortions. This is measured in arcseconds, where a 1.0 arcsec seeing is good for the average astronomical site. Seeing at La Silla is usually between 0.5 and 1.5 arcsec (with an average of 0.9 arcsec (La Silla Observatory, 2019)), but can get down to as low as ~ 0.3 in good conditions⁵. A seeing above 1.5 is not ideal; a seeing above 2.0 will heavily impact the errors on RV measurements.

The ongoing weather conditions are publicly available on the La Silla MeteoMonitor page⁶.

Target details

Certain information about a target star needs to be provided in order to observe it. These are as follows (La Silla Observatory, 2019):

1. **Target RA and Dec (J2000).** Precession of the Earth's axis causes a very slow change of the RA and Dec of objects in the sky, and so it is necessary to define a

⁴Detailed at <https://www.eso.org/sci/facilities/lasilla/telescopes/3p6/overview.html>, accessed 20 September 2023.

⁵The seeing statistics at La Silla can be found at <https://www.eso.org/sci/facilities/lasilla/astclim/seeing.html>, accessed 21 Sep 2023.

⁶The La Silla MeteoMonitor can be found at <https://www.ls.eso.org/lasilla/dimm/>, accessed 21 Sep 2023.

specific epoch with them. J2000.0 is used by ESO, which is a reference time of noon at Greenwich, England (UT) on 1 Jan 2000.

2. **Target proper motion.** A star is not stationary in space, and its position on the sky may change. This is its proper motion. This should be provided as a change in RA and Dec in arcsec per year, and is measured for many stars by *Gaia*.
3. **Target radial velocity.** An initial guess of the RV of the star is required for optimum RV determination by the pipeline. Again, this is already measured for many stars by *Gaia*.
4. **Target spectral type.** The spectral type is matched to one of the template spectra (G2V, K5V, or M2V) which is then used to calculate the cross-correlation function (see Section 2.2.1). If spectral type is unknown but effective temperature is known, the relation of Pecaut and Mamajek (2013) can be used⁷.
5. **Acquisition mode.** HARPS has a number of instrument configurations. These are: HARPS_ech_acq_wavesimult for simultaneous Fabry-Perot or Thorium-Argon lamp exposures; HARPS_ech/eggs_acq_objA for fibre spectroscopy with the object on fibre A; HARPS_ech/eggs_acq_objAB for fibre spectroscopy with the object on fibre A and the sky on fibre B; HARPS_pol_acq_cir for circular spectro-polarimetry; and HARPS_pol_acq_lin for linear spectro-polarimetry.
For the purposes of the observations in this thesis, the sole acquisition mode used is HARPS_ech/eggs_acq_objAB for its high precision.
6. **Choice of fibre/readout mode.** There is a choice of two fibres to observe with. The default fibre type is the High-Accuracy Mode (HAM), with a readout mode of 416 kHz. A high-efficiency mode (EGGS) is also available, using a different, wider fibre, and a slower readout mode of 104 kHz. With EGGS, the photon collecting efficiency is increased by up to a factor of two to the detriment of the radial velocity accuracy. It is useful for faint stars (less than a visual magnitude of ~ 13.5) which only require an accuracy of $\geq 3 \text{ m s}^{-1}$.
7. **An exposure time for the observation.** ESO provides an exposure time and signal-to-noise calculator for HARPS (and NIRPS)⁸. As a general rule of thumb, to reach a precision of a few m s^{-1} , stars of a visual magnitude brighter than 11 need an exposure time of 1200 s or less; 11-12 need an exposure time of 1800 s; fainter than 12 need an

⁷This is accessible in digital table format at https://www.pas.rochester.edu/~emamajek/EEM_dwarf_UBVIJHK_colors_Teff.txt, accessed 21 Sep 2023.

⁸The ESO exposure time calculator (ETC) version 2.0 is available at <https://etc.eso.org/observing/etc/home>.

exposure time of 2400 s. Exposure times may need to be increased in worse air mass / seeing conditions.

Scheduling considerations

Here, I will describe observation scheduling for a target that has a candidate planet with a known orbital period, radius, and ephemeris, i.e. a *TESS* Object of Interest, as these are what are observed in Sections 3, 4, and 5. This will not be applicable to, for example, a blind search for planets.

HARPS observing is split into semesters, where a semester is 6 months long. There is a summer semester (April-September inclusive), and a winter semester (October-March). An observer will be allocated a certain number of nights/hours in a semester, over either the duration of one semester (a “regular” programme) or four semesters (a “large” programme). HARPS observing is run as a time-share programme. Instead of having, for example, 6 nights of observing in a one week block and that being that (as is usual for observing), an observer’s hours are split over the whole semester. This agreement is predicated on the nature of RV observations: for one target, a maximum of 1-2 exposures in a night is likely wanted, with a baseline of weeks to months of regular observations on that target, depending on the orbital period of the planet. This is in order to fill out its phase curve. It is worth noting that the timeshare is performed with the consent of the PIs - individual PIs can opt out of the timeshare if they wish.

A target will not be visible all the time. As the Earth orbits the Sun, different parts of the sky come in and out of view. A star will usually rise in the sky for a month, be observable for several more, then set over a month and become unobservable again, the cycle repeating at the same time the next year. As the star travels across the sky in the months it is visible, its period of minimum air mass will move from the end of the night to the start of the night, so its window of observability will change. It is important to know when targets are rising and setting, and when they can be observed within air mass limits for the entire needed lengths of exposures (see Table 4.6). There is the need to be able to obtain as many RV points as required to secure the mass within the timeframe of observing semesters. Additionally, if only one semester of observing is being undertaken, there will be some targets that are just not accessible.

The most efficient way to confirm that an RV signal is of planetary nature is to observe it “at quadrature” - i.e. an observation when the RV phase curve is predicted to peak, and an observation where it is predicted to trough. While the mass of a planet is not known prior to RV observations, it can be predicted through mass-radius relations like in Chen and Kipping (2017), and the subsequent semi-amplitude will likely be on the order of tens of m s^{-1} or less. An eclipsing binary, however, will have a semi amplitude of km s^{-1} ,

and it is very easy to distinguish with a few RV points.

Filling out the RV phase curve is the next important step – capturing enough RVs over the whole orbital period of the planet. The shape of the curve determines the eccentricity. A good fit around the peak and trough of the curve allows the semi-amplitude to be pinned down, and thus the mass of the planet. For composition studies of exoplanets, it is important to determine the mass to a high enough precision, which is here defined as an error of 20 per cent or better, where the higher precision the better.

There are, however, factors that can cause confusion in the RV signal of a planet. These are stellar activity, stellar drift, and/or additional planets in the system. These are discussed below.

2.2.3 Confirmation of radial velocity signals

As RV data is collected, it is important to be able to do first-look analysis to inform ongoing decisions on whether a target is showing a planetary signal or flatlining, and whether there are complications arising such as stellar activity and drift. Monitoring data as it is taken allows observing strategy to be adjusted on the fly, which can save valuable hours of observation time (e.g. ceasing observations of targets that are false positives) and ensures the data fulfils the goals of the observing programme (e.g. a mass value with errors determined to better than 20 per cent).

Setting up a full joint-fit model (as described in Section 2.3) is often too complicated and slow for analysis on the fly, especially if tens of targets are being observed. There are also specific diagnostic tools used to determine the validity of RV signals before they make it to fitting, for example periodograms and stellar activity correlations. A platform like the Data & Analysis Center for Exoplanets (DACE⁹) is a lightweight but powerful solution, equipped with both diagnostic and fitting tools. It also automatically pulls data from the ESO archive, where RV data from HARPS is stored. This was used extensively for the Nomads programme (Section 4).

Below, I describe some of the various methods employed to analyse RV data.

Simple Keplerian fit

With a few points of RV data, in DACE it is very simple to give priors for known orbital period and ephemeris (in this case from the TOI database) and fit a simple Keplerian. As well as checking the semi-amplitude and comparing it to a predicted semi-amplitude using e.g., Chen and Kipping (2017), it allows a check of whether the data points are in phase with the expected RV curve. If the maximum falls where there should be a minimum, and vice

⁹DACE is available at <https://dace.unige.ch/dashboard/>

versa, it is a first indicator that perhaps something is not as it should be. Phase-folding to the period allows a single period and where all the data points lie on that curve to be seen.

Periodograms

Any signal can be approximated as a sum of sine waves with varying amplitude and frequencies. When looking at RV data, there is a need to identify the important periods in the time-series data that arise from the presence of planets (or stellar activity). The Lomb-Scargle periodogram (Lomb, 1976; Scargle, 1982) is an algorithm for detecting periodicity in an unevenly sampled time-series, and as such is useful for analysing RV data. It estimates Fourier power as a function of period of oscillation, and peaks in the power spectrum correspond to periodicities that fit the time-series particularly well (VanderPlas, 2018).

There are a few practical considerations when using the Lomb-Scargle periodogram for RV data. First, the sampling rate (i.e. observations can only be made at night) is going to introduce alias peaks in the power spectrum around 1 day – the “daily aliases” (Dawson and Fabrycky, 2010). This makes distinguishing planets that have a period close to 1 day particularly difficult.

Second, a periodic signal will have harmonics at e.g. half or twice its value. Sometimes it can be hard to tell which value is the true periodicity – phase folding data to see whether a particular periodicity is sampled well by the data points can be useful here (and then obtaining further data to fill in any gaps), as well as prior knowledge of the expected period (from transits, or photometric monitoring of stellar activity). Additionally, to identify whether there are truly multiple signals in the data (rather than just aliasing), a fit to the prominent periodicity can be removed from the RV data and examine what is left over with another computation of the periodogram on the residuals.

How can it be known that a signal in the periodogram is real and not spurious? For this, the False Alarm Probability (FAP) is used, which measures the probability that a data set with no signal would coincidentally result in a peak of similar magnitude. There are several methods to compute the FAP; DACE uses the analytical method of Baluev (2008). The 10, 1 and 0.1 per cent FAP limits are usually used, in order of increasing confidence (i.e. a 0.1 per cent FAP corresponds to a 0.1 per cent change of the signal being spurious).

Finally, and perhaps most importantly, planets are not the only source of periodicity – stellar activity also introduces periodicity to the RV data. This is discussed below.

Stellar activity and stellar activity indicators

The stellar surface is not, unfortunately, uniform and quiet, and one of the main problems in finding and characterising exoplanets via RVs is the intrinsic variability of their host

stars. Stellar surface variations caused by granulation (e.g., Cegla et al., 2019), solar-like oscillations (e.g., Butler et al., 2004), and magnetic activity, the latter of which produces spots (e.g., Lagrange et al., 2010) and faculae (e.g., Meunier et al., 2010), cause signals in RV data.

A Sun-like star has an outer convective envelope in which hot bubbles of plasma rise up, cool, and then cooled plasma sinks back down in surrounding regions. These upwelling regions are known as granules, the sinking regions as intergranular lanes, and the overall effect is granulation, or stellar surface magnetoconvection, shown in Fig. 2.4. The upwelling and sinking regions cause blue and redshifts in the light of the star respectively. Over the surface of the Sun-like star, the granules cover more surface area than the intergranular lanes, and so there is a net convective blueshift. This can be seen in the bisector span of the CCF, explained below. Overall, the stellar velocity variations from granulation are on the order of tens of cm s^{-1} , the same scale as the Doppler shift caused by small planets (Cegla et al., 2019).

Solar-like oscillations are excited stochastically by turbulent convection in the convective envelopes of Sun-like stars. The convection excites acoustic waves, which can in turn set up internal standing waves, and result in resonant p-modes (pressure modes, named for the pressure gradients that are the restoring force) (Chaplin and Miglio, 2013). This can cause RV shifts of tens of cm s^{-1} to m s^{-1} (Chaplin et al., 2019).

The surface of a Sun-like star is complex, host to temporary active regions where magnetic activity is intense. Star spots are areas of magnetic activity inhibited convection in a star's photosphere, showing up as dark spots with a central dark umbra and, sometimes, a surrounding penumbra (Fig. 2.4). The umbra is where the magnetic field is strongest and normal to the photosphere. The penumbra is made of penumbral filaments and are areas of more inclined magnetic fields, and this will grow around a proto-spot (Schlichenmaier et al., 2010). Sometimes, a cluster of several umbra will be surrounded by one continuous penumbra. There are also bright regions known as plage, caused by suppression of convection (Fig. 2.4). They have less contrast in temperature compared to spots, but cover a much larger area. They are also not just present in active regions, but also in network structures outside of these. Overall, the contribution of spots and plages can cause RV shifts of tens of m s^{-1} (Meunier et al., 2010). Star spots emit less flux and appear as emission bumps in absorption lines; the opposite applies for plages. This changes the shapes of observed line profiles. Also, as a star rotates, these features come into and out of view, and they have variation over time as they grow and decay. This periodicity can both be mistaken for a planetary signal, or obfuscate one (e.g., Haywood et al., 2014; Rajpaul et al., 2016).

Longer-term magnetic signals can also add further periodicity which causes confusion when looking for long-period exoplanets (e.g., Lovis et al., 2011).

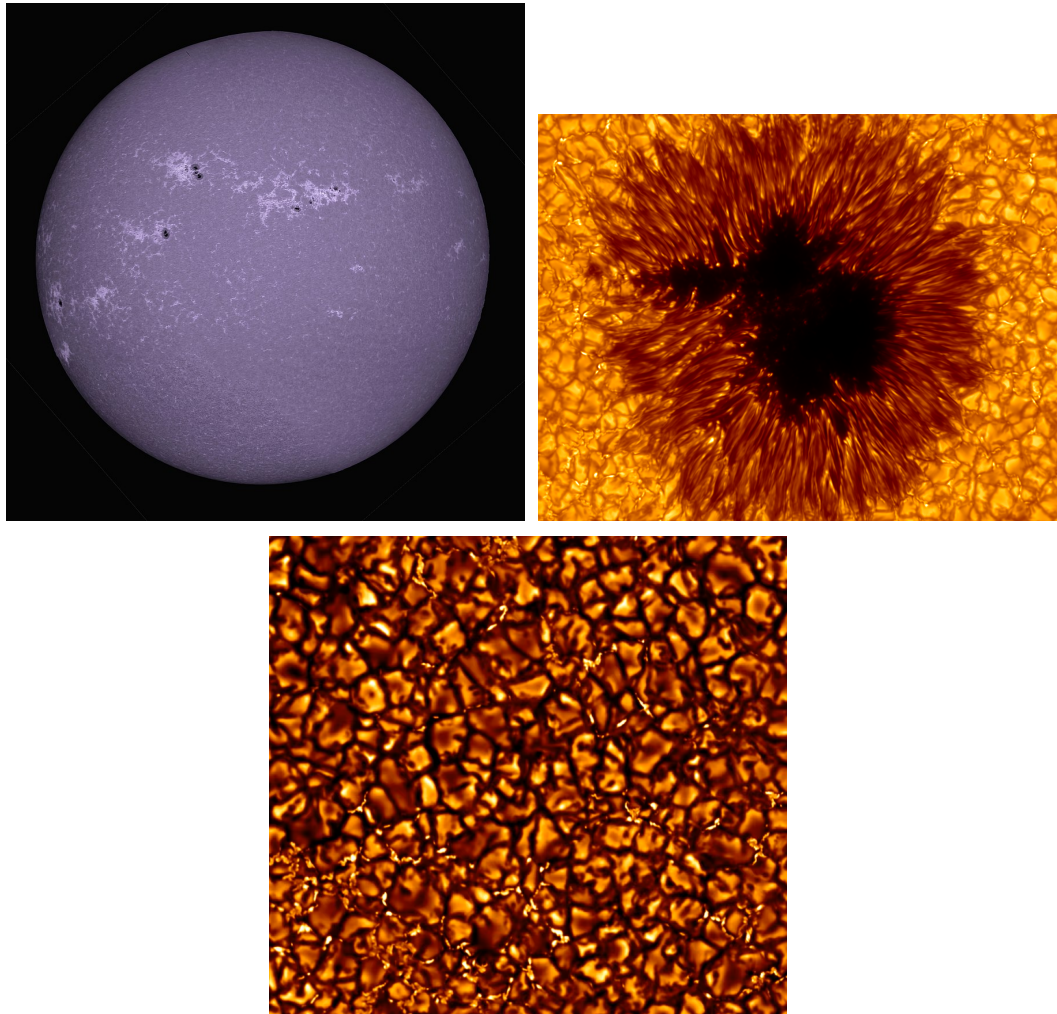


Figure 2.4: Observations of activity on the surface of the Sun.

Top left: the entire surface of the Sun in calcium K, which shows the dark spots and bright plages. Reproduced from Wikimedia Commons, [https://en.wikipedia.org/wiki/File:Calcium_K_Mosaic_2015_04_18_\(16569533153\).jpg](https://en.wikipedia.org/wiki/File:Calcium_K_Mosaic_2015_04_18_(16569533153).jpg).

Top right: a zoomed in view of a sunspot observed with The Swedish 1-m Solar Telescope, showing the umbra and penumbra. Granulation can also be seen on the surrounding stellar surface. Reproduced from Wikimedia Commons, https://en.wikipedia.org/wiki/File:Sunspot_SST_05Sep2016.png.

Bottom: a frame from video showing granulation motions on the Sun's surface, observed with The Swedish 1-m Solar Telescope. The hot granules and dark intergranular lanes can be clearly seen. Reproduced from Wikimedia Commons, where the video can also be viewed, https://en.wikipedia.org/wiki/File:Granulation_Quiet_Sun_SST_25May2017.webm.

Fortunately, there are a number of “indicators” present in the stellar spectrum and CCF that can inform us on the activity of the host star (depicted in Fig. 2.5). Several are computed as part of the HARPS data reduction pipeline, and are as follows:

- **S-index:** the flux ratio of two bandpasses centred on the H and Ca II emission cores and two continuum regions either side. This is caused by enhancement of coronal and chromospheric emission induced by magnetic activity (Egeland et al., 2017).
- $\log R'_{HK}$: a derivative of the S-index, where the contribution from the photosphere to the total bolometric emission of the star is excluded (Noyes et al., 1984). I note that this activity indicator is *not* available direct from the HARPS pipeline, but can instead be calculated independently from the HARPS spectra using, e.g., ACTIN2 (Gomes da Silva et al., 2018, 2021).
- **H α -index:** in active regions of the chromosphere, plages, H α photons are emitted, and so the depth of the H α absorption line is dependent on activity – it fills in with increasing activity. This activity indicator may be especially important for M-dwarfs, as they have lack of flux near the calcium lines used in the S-index (Robertson et al., 2013).
- **FWHM of the CCF:** changes can occur in the spectral line profiles of a star due to pulsation or an active atmosphere, causing the RV to change in a way that *is not* due to the velocity of the star induced by the presence of a planet. The easiest way to look for this effect is by checking the shape of the CCF: if it is constant in time, then the signal is likely due to a planet. If there is a change, for example the full-width at half maximum of the CCF changes in size, it indicates the stellar line profiles are changing in shape due to stellar atmosphere changes (Queloz et al., 2001; Lafarga et al., 2020), or the presence of a binary companion.
- **Bisector span of the CCF:** the bisector of the CCF displays a classic “C” shape due to granulation on the surface of the star – upwelling regions are blue-shifted, sinking regions are red-shifted, and a solar-like star will have a greater surface area of upwelling regions. The bisector span is a measure of the difference in radial velocity between the top and bottom of a bisector. If the bisector shape changes due to changing granulation, this will reflect in the bisector span value changing (Baştürk et al., 2011).
- **Contrast of the CCF:** this is similar in concept to the FWHM, but the contrast instead measures changes in the depth of the CCF, rather than the width (Lafarga et al., 2020).

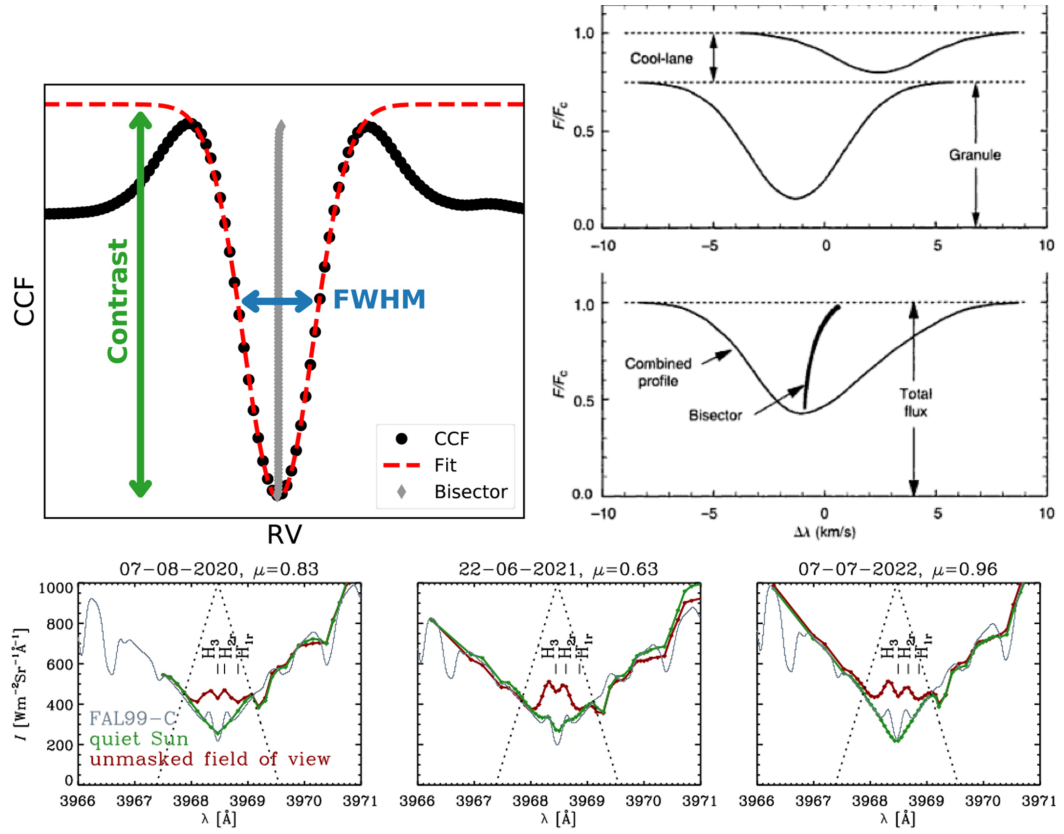


Figure 2.5: Diagrams showing how the activity indicators are derived.

Top left: a CCF profile for an M-dwarf is given in black, and the contrast and FWHM of the CCF are labelled. The bisector span is also shown in grey. Reproduced from (Lafarga et al., 2020).

Top right: the top panel shows the relative contributions to the CCF of the cool intergranular lanes and the hot granules. These combine to create the asymmetric profile in the bottom panel, and the resultant curved, “c”-shaped bisector is shown. Reproduced from (Gray, 2005).

Bottom: a zoom in of the H line used for the S-index as measured from the Sun. A comparison is made between what this line looks like for a quiet area of the Sun’s surface (green line), where there is no emission in the core, and for the whole field of view including an active region (maroon), which shows emission in the core. Reproduced from (Sownya et al., 2023).

There are two main ways to examine these: they may vary periodically as the strength of the activity waxes and wanes, and this signal will show up in a periodogram created from the activity indicator. The second is to plot RV against activity indicator and look for a correlation between the two, which indicates that the changes in the RVs are caused by the activity rather than a planetary signal. This can be quantified using the Pearson's R statistic, a measure of correlation strength (Pearson, 1895).

It is worth mentioning here that stellar activity also affects photometric data - for example, sunspots change the flux emitted from the star and so can show up in light curves (e.g., Almenara et al., 2022; Libby-Roberts et al., 2023). So, looking for periodic signals in long-term light curves (several sectors of TESS data, or photometric monitoring from a facility like WASP or NGTS) can also be used to corroborate periodicity found from the aforementioned indicators.

Stellar drift and long-term magnetic activity cycles

Sometimes, a baseline RV can be seen to very slowly increase or decrease as time passes in observations. This “drift” can be due to the influence of a large, outer body in the system, or a longer-term stellar magnetic cycle as discussed in the previous section. Because these effects are often on the scale of years, and observations of a single target do not usually have a baseline of that magnitude, only a small portion of the periodic signal is seen. Whether this signal shows up in a periodogram also depends on the baseline of observations. Depending on the shape of the drift that is seen, it is often fit with a linear or quadratic relation. This is easy to do in DACE, and by removing this fit the drift-free data can be examined for the often shorter-term planetary signals of interest.

2.3 Joint fit modelling of exoplanet systems

Once observational data is collected on an exoplanetary system using our various detection methods, it needs to be simultaneously fit to derive values for parameters of interest. This work is performed in Chapters 3 and 5.

A joint-fit of an exoplanetary system is a complex problem that hangs together data detrending techniques, Keplerian orbital motion, limb darkening laws, and a method for exploring the parameter space to find the best fit to all of these elements. There are many frameworks that have been developed specifically for modelling exoplanetary systems, `exoplanet` (Foreman-Mackey et al., 2021), `juliet` (Espinoza et al., 2019), and `allesfitter` (Günther and Daylan, 2021, 2019) to name a few, and each have their own goals and come with various advantages and disadvantages. In this work, I exclusively use `exoplanet` for the reasons outlined below.

2.3.1 Using `exoplanet`

`exoplanet` is a toolkit built upon the probabilistic Python library PyMC3 (Salvatier et al., 2016), and it allows the user to build models of planetary systems and fit them using MCMC methods (see Section 2.3.2). It is high-performance, important for models with a large number of parameters, and flexible. `exoplanet` offers many of the functions and distributions needed to fit a full range of exoplanetary data and exoplanet system architectures. This includes a solver for Kepler’s equation, incorporation of `starry` (Luger et al., 2019) for limb-darkened light curves, and incorporation of `celerite` for Gaussian Processes (Foreman-Mackey et al., 2017). Due to the simplicity of the framework, where necessary functionality is not included or covered in the many tutorials, it is easy for the user to build themselves. This was necessary in Section 3, where Pixel Level Decorrelation of *Spitzer* data and a custom GP kernel to detrend stellar activity in the RV data were included.

Sections 3 and 5 explain specific set-ups of `exoplanet` for two different exoplanetary systems. In short, the user defines prior distributions for each parameter being fit, using PyMC3 distributions¹⁰. Depending on the data available (i.e. transit data and RV data allow us to fit different parameters, see Section 1.3.1), this could be stellar radius and mass, orbital period, time of midtransit, eccentricity and argument of periastron, planet radius, quadratic limb darkening parameters, RV semi-amplitude, offsets and jitter on the data, and GP kernel parameters. The user builds these into a model framework, and then the prior distributions are sampled as outlined below.

¹⁰A full list of available PyMC3 distributions are available at <https://www.pymc.io/projects/docs/en/v3/api/distributions/continuous.html>, accessed 22 Sep 2023.

2.3.2 Bayes and MCMC

Joint fit modelling of exoplanetary systems utilises Markov-Chain Monte Carlo methodology to explore the parameter space of the variables in a model that is being used to fit a set of data. It finds the high probability areas, which allows for estimation of the parameter values and their uncertainties.

MCMC is built upon Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (2.1)$$

where:

- $P(A|B)$ is the “posterior”: the probability of A being true, given that B is true;
- $P(B|A)$ is the “likelihood”: the probability of B being true, given that A is true;
- $P(A)$ is the “prior”: the probability of A being true;
- $P(B)$ is the “evidence”: the probability of B being true.

An MCMC sampler creates a chain, a stochastic process of “walkers” which explore the parameter space of the model. A set of values for the model parameters are chosen, x_i , and these start the chain. Then a random step away from these values is performed, creating a new set of parameters, x_{i+1} . The ratio of the posteriors, R , is computed according to:

$$R = \frac{P(B|A(x_{i+1}))}{P(B|A(x_i))} \times \frac{P(A(x_{i+1}))}{P(A(x_i))} \quad (2.2)$$

The value of R determines whether the new parameters are accepted, based upon the Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970). If $R \geq 1$, the new parameters are accepted with x_{i+1} becoming the new x_i , and the chain is continued by taking a new random step and repeating the process. If $R < 1$, the new parameters are accepted with a probability of R . In the case they are accepted, x_{i+1} becomes the new x_i , and the chain is continued by taking a new random step and repeating the process. In the case they are not accepted, x_i remains as x_i and the chain is continued by taking a new random step and repeating the process. At each step, the current parameters x_i are appended to the chain. The probabilistic acceptance in the case of $R < 1$ allows the chain to explore the full parameter space without getting stuck in high probability areas.

The walkers may take some time to start exploring the best regions of the parameter space, so a number of steps are discarded as “burn-in” to ensure chains used for the parameter estimation are converged. This is often done by visual inspection of the chains. Additionally, to test for non-convergence in the samples, the rank-normalised split- \hat{R} statistic can be

calculated for each parameter (Vehtari et al., 2021). The number of steps taken after burn-in are often in the thousands. Multiple chains are run in parallel and then combined to increase the number of effective samples. Parameter estimation is performed by taking the median sample of each parameter as its best-fit value, and the 16th and 84th percentiles as the one standard deviation error on the value. This is only strictly applicable in the case where the posterior (sample) distribution is Gaussian. Corner plots (see e.g., `corner` Foreman-Mackey, 2016) are used to look at the sample distribution shapes to check for a Gaussian shape and search for any correlations between the parameters.

`exoplanet` uses a variant of MCMC, Hamiltonian Monte Carlo (HMC, also known as hybrid Monte Carlo, see e.g., Betancourt, 2017), via the No-U-Turn Sampler (NUTS, Hoffman and Gelman, 2011). The base concept is the same, but the behaviour of the walkers is not a random walk as described above. Methods such as the random walk may take an exceptionally long time to converge when used with complicated models that have many parameters, as is the case with modelling exoplanetary systems. HMC can be used in the case where model parameters are continuous rather than discrete (as they are in this scenario) by changing the problem of sampling from a distribution into a problem of simulating Hamiltonian dynamics. In short, a vector field guides the walker through the parameter space. For a detailed explanation, see e.g., Betancourt (2017); Hoffman and Gelman (2011). The end result is that the walkers move across the space in larger steps, efficiently moving through the space far away from the initial point to new regions quickly, and are thus less correlated and converge more rapidly.

Chapter 3

TOI-431/HIP 26013: a super-Earth and a sub-Neptune transiting a bright, early K dwarf, with a third RV planet

Declaration and data availability

This chapter is a reproduction of the paper “TOI-431/HIP 26013: a super-Earth and a sub-Neptune transiting a bright, early K dwarf, with a third RV planet” (Osborn et al., 2021a), published in the Monthly Notices of the Royal Astronomical Society in October 2021. The majority of the analysis and writing was performed by myself. Contributions from co-authors, including observational data, analysis, and partial writing of sections, are found in Sections 3.2.1 (specifically photometry from LCOGT by Ryan Cloutier, PEST by Karen A. Collins, *Spitzer* by Ian J. M. Crossfield, and NGTS by Edward M. Bryant), 3.2.2 (specifically observations from HIRES by Fei Dai, iSHELL by Bryson Cale, FEROS by Rafael Brahm, and Minerva-Australis by Robery A. Wittenmyer), 3.2.3 (by Carl Ziegler, Elisabeth C. Matthews, Steve B. Howell, David Ciardi, and Jorge Lillo-Box), 3.2.4 (by Nuno C. Santos, Sérgio G. Sousa, Vardan Adibekyan, Elisa Delgado-Mena, Keivan G. Stassun, and Malcolm Fridlund), 3.2.5 (by Coel Hellier and Edward M. Bryant), and 3.4 (specifically the photoevaporation analysis by George W. King, composition analysis by Jon Otegi and Caroline Dorn, and the calculation of the TSM by Ian J. M. Crossfield).

The TESS data are available from the Mikulski Archive for Space Telescopes (MAST), at <https://heasarc.gsfc.nasa.gov/docs/tess/data-access.html>. The other photometry from the LCOGT, NGTS, and *Spitzer*, as well as all of the RV data, are

available for public download from the ExoFOP-TESS archive at <https://exofop.ipac.caltech.edu/tess/target.php?id=31374837>. This data is labelled “Osborn+ 2021” in their descriptions. The full HARPS RV data products are also publicly available from the ESO archive, at http://archive.eso.org/wdb/wdb/adp/phase3_main/form. The high-resolution imaging data is also available from the ExoFOP TESS archive.

Abstract

We present the bright ($V_{mag} = 9.12$), multi-planet system TOI-431, characterised with photometry and radial velocities. We estimate the stellar rotation period to be 30.5 ± 0.7 days using archival photometry and radial velocities. TOI-431 b is a super-Earth with a period of 0.49 days, a radius of $1.28 \pm 0.04 R_{\oplus}$, a mass of $3.07 \pm 0.35 M_{\oplus}$, and a density of $8.0 \pm 1.0 \text{ g cm}^{-3}$; TOI-431 d is a sub-Neptune with a period of 12.46 days, a radius of $3.29 \pm 0.09 R_{\oplus}$, a mass of $9.90^{+1.53}_{-1.49} M_{\oplus}$, and a density of $1.36 \pm 0.25 \text{ g cm}^{-3}$. We find a third planet, TOI-431 c, in the HARPS radial velocity data, but it is not seen to transit in the *TESS* light curves. It has an $M \sin i$ of $2.83^{+0.41}_{-0.34} M_{\oplus}$, and a period of 4.85 days. TOI-431 d likely has an extended atmosphere and is one of the most well-suited *TESS* discoveries for atmospheric characterisation, while the super-Earth TOI-431 b may be a stripped core. These planets straddle the radius gap, presenting an interesting case-study for atmospheric evolution, and TOI-431 b is a prime *TESS* discovery for the study of rocky planet phase curves.

3.1 Introduction

The discoveries of the *Kepler* Space Telescope (Borucki et al., 2010) provided us with the means to make statistical studies on the exoplanet population for the first time: *Kepler* has shown us that Neptune-sized planets are more common than large gas giants (Fressin et al., 2013), and that super-Earths are the most abundant planet type (Petigura et al., 2013). It became possible to look for trends that might elucidate planetary formation mechanisms; one such trend discovered is a bi-modality in the radius distribution of small planets. Often dubbed the “photoevaporation valley,” the commonly posited explanation for its existence is photoevaporation of close-in planetary atmospheres (Owen and Wu, 2017; Fulton et al., 2017; Fulton and Petigura, 2018; Van Eylen et al., 2018; Cloutier and Menou, 2020). Planets above the radius gap have retained gaseous envelopes, while planets below are theorised to have been stripped of any gas to become naked cores. Multi-planet systems have been discovered containing planets that lie both below and above the radius gap (e.g. Günther et al., 2019; Cloutier et al., 2020a), and such systems are important when considering

how evolution mechanisms may sculpt the radius gap as they allow testing of atmospheric evaporation and bulk composition models.

Further to the discovery of the radius gap, a paucity of intermediate-sized planets at short periods (≤ 3 days) dubbed the “Neptune/sub-Jovian Desert” (Szabó and Kiss, 2011; Beaugé and Nesvorný, 2013; Helled et al., 2016; Lundkvist et al., 2016; Mazeh et al., 2016; Owen and Lai, 2018), can be seen in both the mass-period and radius-period distribution of exoplanets, and Mazeh et al. (2016) and Owen and Lai (2018) derived boundaries for this triangular-shaped region, and the potential mechanisms behind their existence.

The Transiting Exoplanet Survey Satellite (*TESS*, Ricker et al., 2015) is now building upon the legacy of *Kepler*. Unlike *Kepler*, *TESS* has been optimised to look at bright stars, enabling high precision radial velocity follow up of planetary candidates to determine their masses, and additional follow-up (with JWST, for example) will allow us to study their atmospheres. Over the course of its two year primary mission, which came to an end in July 2020, over 2000 *TESS* Objects of Interest (TOIs) were released, and there have been many discoveries that contribute to fulfilling its Level-1 mission goal to measure the masses and radii of at least 50 planets with radii smaller than $4 R_{\oplus}$ (e.g. Huang et al., 2018; Gandolfi et al., 2018; Cloutier et al., 2019; Dragomir et al., 2019; Dumusque et al., 2019; Luque et al., 2019; Díaz et al., 2020; Astudillo-Defru et al., 2020; Cloutier et al., 2020b,a; Nielsen et al., 2020a; Armstrong et al., 2020).

We present here the discovery of TOI-431 b, c, and d. TOI-431 b and d are a super-Earth and sub-Neptune respectively, discovered first by *TESS* and confirmed via extensive follow up: high-precision Doppler spectroscopy from the High Accuracy Radial velocity Planet Searcher (HARPS, Pepe et al., 2002) and the High Resolution Spectrograph (HIRES, Vogt et al., 1994) which allows us to determine their masses; additional Doppler spectroscopy from iSHELL (Rayner et al., 2016), FEROS (Kaufer and Pasquini, 1998), and MINERVA-Australis (Addison et al., 2019); ground-based transit detections from NGTS (Wheatley et al., 2018) and the LCOGT 1m network (Brown et al., 2013); and a double-transit from the *Spitzer* space telescope. Both TOI-431 b and d contribute to the *TESS* Level-1 mission goal. TOI-431 c is an additional planet that we have found in the HARPS radial velocity data, and it is not seen to transit. We describe the observations made and the stellar analysis of the TOI-431 system in Section 3.2; our joint-fit model of the system in Section 3.3; and put this system into context in Section 3.4.

3.2 Observations

Here, we describe the instrumentation used and observations made to detect and characterise the TOI-431 system.

Table 3.1: Details of the TOI-431 system.

Property	Value	Source
Identifiers		
TIC ID	31374837	TICv8
HIP ID	26013	
2MASS ID	05330459-2643286	2MASS
Gaia ID	2908664557091200768	GAIA EDR3
Astrometric properties		
R.A. (J2016.0)	05:33:04.62	GAIA EDR3
Dec (J2016.0)	-26:43:25.86	GAIA EDR3
Parallax (mas)	30.65 ± 0.01	GAIA EDR3
Distance (pc)	32.61 ± 0.01	Bailer-Jones et al. (2021)
$\mu_{\text{R.A.}}$ (mas yr ⁻¹)	16.89 ± 0.01	GAIA EDR3
μ_{Dec} (mas yr ⁻¹)	150.78 ± 0.01	GAIA EDR3
Photometric properties		
TESS (mag)	8.171 ± 0.006	TICv8
B (mag)	10.10 ± 0.03	TICv8
V (mag)	9.12 ± 0.03	TICv8
G (mag)	8.7987 ± 0.0003	GAIA EDR3
J (mag)	7.31 ± 0.03	2MASS
H (mag)	6.85 ± 0.03	2MASS
K (mag)	6.72 ± 0.02	2MASS

Sources: TICv8 (Stassun et al., 2019), 2MASS (Skrutskie et al., 2006), Gaia Early Data Release 3 (Gaia Collaboration et al., 2021)

3.2.1 Photometry

TESS photometry

The TOI-431 system (TIC 31374837, HIP 26013, see Table 3.1) was observed in *TESS* Sectors 5 (Nov 15 to Dec 11 2018) and 6 (Dec 15 2018 to Jan 6 2019) on Camera 2 in the 2-minute cadence mode ($t_{\text{exp}} = 2$ min). TOI-431.01 (now TOI-431 d) was flagged on Feb 8 2019 by the MIT Quick-Look Pipeline (QLP, Huang et al., 2020) with a signal-to-noise ratio (SNR) of 58; the Sector 5 light curve reveals 2 deep transits of TOI-431 d, but further transits of this planet fell in the data gaps in S6. TOI-431 d passed all Data Validation tests (see Twicken et al., 2018) and model fitting (see Li et al., 2019); additionally, the difference image centroiding results place the transit signature source within ~ 3 arcsec of the target star. TOI-431.02 (now TOI-431 b) was flagged later, on June 6, after identification by the *TESS* Science Processing Operations Center (SPOC) pipeline (Jenkins et al., 2016) with an SNR of 24 in a combined transit search of Sectors 5-6.

We used the publicly available photometry provided by the SPOC pipeline, and used the Presearch Data Conditioning Simple Aperture Photometry (PDCSAP_FLUX), which has common trends and artefacts removed by the SPOC Presearch Data Conditioning (PDC) algorithm (Twicken et al., 2010; Smith et al., 2012; Stumpe et al., 2012, 2014). The median-normalised PDCSAP flux, without any further detrending, is shown in the top panel of Fig. 3.8.

LCOGT photometry

To confirm the transit timing and depth, and to rule out a nearby eclipsing binary (NEB) as the source of the *TESS* transit events, we obtained three seeing-limited transit observations of TOI-431 d in the zs -band. The light curves were obtained using the 1-m telescopes at the Cerro Tololo Inter-American Observatory (CTIO) and the Siding Springs Observatory (SSO) as part of the Las Cumbres Observatory Global Telescope network (LCOGT; Brown et al., 2013). Both telescopes are equipped with a 4096×4096 Sinistro camera with a fine pixel scale of $0.39'' \text{ pixel}^{-1}$.

We calibrated each sequence of images using the standard LCOGT BANZAI pipeline (McCully et al., 2018). The observations were scheduled using the *TESS* Transit Finder, a customised version of the *Tapir* software package (Jensen, 2013). The differential light curves of TOI-431, and seven neighbouring sources within $2.5'$ based on the *Gaia* DR2 (Gaia Collaboration et al., 2018), were derived from uncontaminated apertures using AstroImageJ (AIJ; Collins et al., 2017). Two partial transits were obtained on UT December 9 2019 which covered the ingress and egress events from CTIO and SSO respectively (Figure 3.9). We then obtained a second ingress observation on January 3 2020 from CTIO. Within each

light curve, we detected the partial transit event on-target and cleared the field of NEBs down to $\Delta z_s = 6.88$ mag.

PEST photometry

We also obtained a seeing-limited observation during the time of transit of TOI-431 d on UT February 13 2020 using the Perth Exoplanet Survey Telescope (PEST) near Perth, Australia. The 0.3 m telescope is equipped with a 1530×1020 SBIG ST-8XME camera with an image scale of 1.2 arcsec per pixel, resulting in a $31 \text{ arcmin} \times 21 \text{ arcmin}$ field of view. A custom pipeline based on C-Munipack¹ was used to calibrate the images and extract the differential photometry, using an aperture with a radius of 6.2 arcsec. The images have typical stellar point spread functions (PSFs) with a FWHM of ~ 5 arcsec. Because the transit depth of TOI-431 d is too shallow to detect from the ground with PEST, the target star was intentionally saturated to check the fainter nearby stars for possible NEBs that could be blended in the *TESS* aperture. The data rule out NEBs in all 17 stars within 2.5 arcmin of the target star that are bright enough (*TESS* magnitude < 17.4) to cause the *TESS* detection of TOI-431 d.

Spitzer photometry

Shortly after TOI-431 was identified and announced as a *TESS* planet candidate, we identified TOI-431 d as an especially interesting target for atmospheric characterization via transmission spectroscopy. We therefore scheduled one transit observation with the *Spitzer* Space Telescope to further refine the transit ephemeris and allow efficient scheduling of future planetary transits. We observed the system as part of *Spitzer* GO 14084 (Crossfield et al., 2018) using the $4.5 \mu\text{m}$ channel of the IRAC instrument (Fazio et al., 2004). We observed in subarray mode, which acquired 985 sets of 64 subarray frames, each with 0.4 s integration time. These transit observations spanned UT times from May 23 2019 21:13 to May 24 2019 04:42, and were preceded and followed by shorter integrations observed off-target to check for bad or hot pixels. Our transit observations used *Spitzer*/IRAC in PCRS Peak-up mode to place the star as closely as possible to the well-characterized “sweet spot” on the IRAC2 detector.

NGTS photometry

The Next Generation Transit Survey (NGTS; Wheatley et al., 2018) is an exoplanet hunting facility which consists of twelve 20 cm diameter robotic telescopes and is situated at ESO’s Paranal Observatory. Each NGTS telescope has a wide field-of-view of 8 square degrees and a plate scale of $5 \text{ arcsec pixel}^{-1}$. NGTS observations are also afforded sub-pixel level guiding

¹<http://c-munipack.sourceforge.net>

through the DONUTS auto-guiding algorithm (McCormac et al., 2013). A transit event of TOI-431 d was observed using 5 NGTS telescopes on February 20 2020. On this night, a total of 5922 images were taken across the 5 telescopes, with each telescope observing with the custom NGTS filter and an exposure time of 10 seconds. The dominant photometric noise sources in NGTS light curves of bright stars are Gaussian and uncorrelated between the individual telescope systems (Smith et al., 2020; Bryant et al., 2020). As such, we can use simultaneous observations with multiple NGTS telescopes to obtain high precision light curves.

All the NGTS data for TOI-431 were reduced using a custom aperture photometry pipeline which uses the SEP library for both source extraction and photometry (Bertin and Arnouts, 1996; Barbary, 2016). Bias, dark and flat field image corrections are found to not improve the photometric precision achieved, and so we do not apply these corrections during the image reduction. SEP and *Gaia* (Gaia Collaboration et al., 2016, 2018) are both used to identify and rank comparison stars in terms of their brightness, colour, and CCD position relative to TOI-431 (for more details on the photometry, see Bryant et al., 2020).

3.2.2 Spectroscopy

HARPS high-resolution spectroscopy

TOI-431 was observed between February 2 and October 21 2019 with the High Accuracy Radial velocity Planet Searcher (HARPS) spectrograph mounted on the ESO 3.6 m telescope at the La Silla Observatory in Chile (Pepe et al., 2002). A total of 124 spectra were obtained under the NCORES large programme (ID 1102.C-0249, PI: Armstrong). The instrument (with resolving power $R = 115,000$) was used in high-accuracy mode (HAM), with an exposure time of 900 s. Between 1 and 3 observations of the star were made per night. The standard offline HARPS data reduction pipeline was used to reduce the data, and a K5 template was used in a weighted cross-correlation function (CCF) to determine the radial velocities (RVs). Each epoch has further calculation of the bisector span (BIS), full-width at half-maximum (FWHM), and contrast of the CCF. This data is presented in Table 3.2.

In addition to this, there are 50 publicly available archival HARPS spectra dating from 2004 to 2015.

HIRES high-resolution spectroscopy

We obtained 28 high-resolution spectra of TOI-431 on the High Resolution Echelle Spectrometer of the 10m Keck I telescope (Keck/HIRES, Vogt et al., 1994). The observation spans a temporal baseline from November 11 2019 to September 27 2020. We obtained an iodine-free spectrum on November 8 2019 as the template for radial velocity extraction. All

Table 3.2: HARPS spectroscopy from February to October 2019.

Time (RJD)	RV (ms^{-1})	σ_{RV} (ms^{-1})	FWHM (ms^{-1})	Bisector (ms^{-1})	Contrast	S_{MW}
58537.53770973021	48830.979962	0.894407	6330.143967	38.148888	49.532876	0.370009
58537.655514969956	48833.848987	0.994823	6330.289387	36.923112	49.534854	0.361645
58539.53381296992	48824.538870	1.006177	6324.786911	39.072933	49.567807	0.365675
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

The full HARPS data products can be found on ExoFOP-TESS at <https://exofop.ipac.caltech.edu/tess/target.php?id=31374837>

Table 3.3: HIRES spectroscopy from November 2019 to September 2020.

Time (BJD TDB)	RV (ms^{-1})	σ_{RV} (ms^{-1})
2458796.014464	4.90676701782345	1.06348240375519
2458797.0428	6.94764041206104	1.14499938488007
2458798.095775	8.81269072598892	1.1401127576828
\vdots	\vdots	\vdots

The full HIRES data products can be found on ExoFOP-TESS at <https://exofop.ipac.caltech.edu/tess/target.php?id=31374837>

other spectra were obtained with the iodine cell in the light path for wavelength calibration and line profile modeling. Each of these spectra were exposed for 4-8 min achieving a median SNR of 200 per reduced pixel near 5500 Å. The spectra were analyzed with the forward-modelling Doppler pipeline described in Howard et al. (2010) for RV extraction. We analyzed the Ca II H & K lines and extracted the S_{HK} using the method of Isaacson and Fischer (2010). This data is presented in Table 3.3.

iSHELL spectroscopy

We obtained 108 spectra of TOI-431 during 11 nights with the iSHELL spectrometer on the *NASA Infrared Telescope Facility* (IRTF, Rayner et al., 2016), spanning 108 days from September-December 2019. The exposure times were 5 minutes, repeated 3-14 times within a night to reach a cumulative photon signal-to-noise ratio per spectral pixel varying from 131–334 at $\sim 2.4 \mu m$ (the approximate centre of the blaze for the middle order). This achieves a per-night RV precision of 3–8 ms^{-1} with a median of 5 ms^{-1} . Spectra were reduced and RVs extracted using the methods outlined in Cale et al. (2019).

FEROS spectroscopy

TOI-431 was monitored with the Fiberfed Extended Range Optical Spectrograph (FEROS, Kaufer et al., 1999), installed on the MPG2.2 m telescope at La Silla Observatory, Chile.

These observations were obtained in the context of the Warm gIaNts with tEss (WINE) collaboration, which focuses on the systematic characterization of *TESS* transiting warm giant planets (e.g., Brahm et al., 2019; Jordán et al., 2020). FEROS has a spectral resolution of $R \approx 48\,000$ and uses a comparison fibre that can be pointed to the background sky or illuminated by a Thorium-Argon lamp simultaneously with the execution of the science exposure. We obtained 10 spectra of TOI-431 between February 28 and March 12 2020. We used the simultaneous calibration technique to trace instrumental radial velocity variations, and adopted an exposure time of 300 s, which translated in spectra with a typical signal-to-noise ratio per resolution element of 170. FEROS data was processed with the *ceres* pipeline (Brahm et al., 2017), which delivers precision radial velocity and line bisector span measurements through the cross-correlation technique. The cross-correlation was executed with a binary mask reassembling the properties of a G2-type dwarf star.

MINERVA-Australis spectroscopy

MINERVA-Australis is an array of four PlaneWave CDK700 telescopes located in Queensland, Australia, fully dedicated to the precise radial-velocity follow-up of *TESS* candidates. The four telescopes can be simultaneously fiber-fed to a single KiwiSpec R4-100 high-resolution ($R=80,000$) spectrograph (Barnes et al., 2012; Addison et al., 2019, 2020). TOI-431 was observed by MINERVA-Australis in its early operations, with a single telescope, for 16 epochs between 2019 Feb 12 and 2019 April 17. Each epoch consists of two 30-minute exposures, and the resulting radial velocities are binned to a single point. Radial velocities for the observations are derived for each telescope by cross-correlation, where the template being matched is the mean spectrum of each telescope. The instrumental variations are corrected by using simultaneous Thorium-Argon arc lamp observations.

3.2.3 High resolution imaging

High angular resolution imaging is needed to search for nearby sources that can contaminate the *TESS* photometry, resulting in an underestimated planetary radius, or that can be the source of astrophysical false positives, such as background eclipsing binaries. The contrast curves from all of the sources of high resolution imaging described below are displayed in Fig. 3.1.

SOAR HRCam

We searched for stellar companions to TOI-431 with speckle imaging with the 4.1-m Southern Astrophysical Research (SOAR) telescope (Tokovinin, 2018) on UT March 17 2019, observing in the Cousins I-band, a similar visible bandpass to *TESS*. More details of the

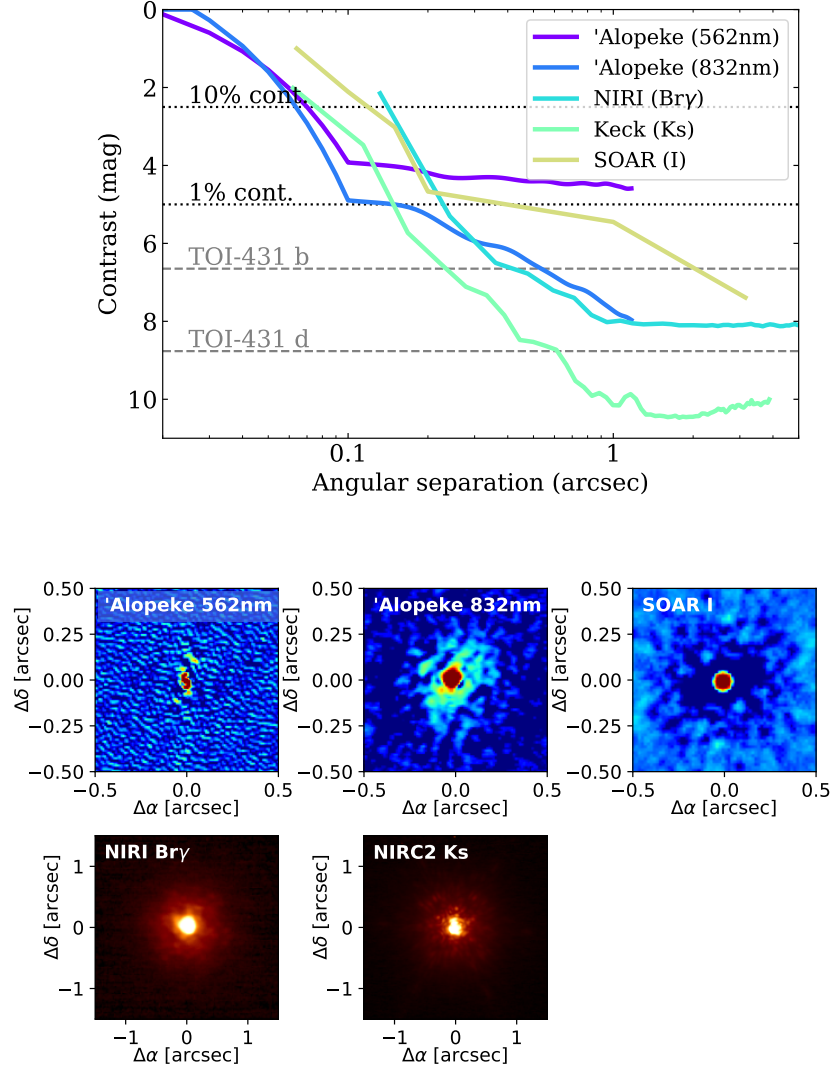


Figure 3.1: *Top*: 5σ contrast curves for all of the sources of high-resolution imaging described in Section 3.2.3. The 10 and 1 per cent contamination limits are given as the black dotted lines. The grey dashed lines labelled TOI-431 b and d represent the maximum contrast magnitude that a blended source could have in order to mimic the planetary transit depth if it were an eclipsing binary. *Bottom*: a compilation of reconstructed images from 'Alopeke and SOAR and AO images from NIRI and NIRC2, with the instrument and filter labelled. No additional companions are seen.

observation are available in Ziegler et al. (2020). The 5σ detection sensitivity and speckle auto-correlation functions from the observations are shown in Fig. 3.1. No nearby stars were detected within 3 arcsec of TOI-431 in the SOAR observations.

Gemini NIRI

We collected high resolution adaptive optics observations using the Gemini/NIRI instrument (Hodapp et al., 2003) on UT March 18 2019. We collected nine images in the Bry filter, with exposure time 0.6 s per image. We dithered the telescope by 2 arcsec between each exposure, allowing for a sky background to be constructed from the science frames themselves. We corrected individual frames for bad pixels, subtracted the sky background, and flat-corrected frames, and then co-added the stack of images with the stellar position aligned. To calculate the sensitivity of these observations, we inject fake companions and measure their S/N, and scale the brightness of these fake companions until they are recovered at 5σ . This is repeated at a number of locations in the image. We average our sensitivity over position angle, and show the sensitivity as a function of radius in Fig. 3.1. Our observations are sensitive to companions 4.6 mag fainter than the host at 0.2 arcsec, and 8.1 mag fainter than the host in the background limited regime, at separations greater than 1 arcsec.

Gemini 'Alopeke

TOI-431 was observed on UT Oct 15 2019 using the 'Alopeke speckle instrument on Gemini-North². 'Alopeke provides simultaneous speckle imaging in two bands, 562 nm and 832 nm, with output data products including a reconstructed image, and robust limits on companion detections (Howell et al., 2011). Fig. 3.1 shows our results in both 562 nm and 832 nm filters. Fig. 3.1 (bottom) shows the 832 nm reconstructed speckle image from which we find that TOI-431 is a single star with no companion brighter than within 5-8 magnitudes of TOI-431 detected within 1.2 arcsec.

The inner working angle of the 'Alopeke observations are 17 mas at 562 nm and 28 mas at 832 nm.

Keck NIRC2

As part of our standard process for validating transiting exoplanets to assess the possible contamination of bound or unbound companions on the derived planetary radii (Ciardi et al., 2015), we observed TOI-431 with infrared high-resolution Adaptive Optics (AO) imaging at Keck Observatory. The Keck Observatory observations were made with the NIRC2 instrument on Keck-II behind the natural guide star AO system. The observations were

²<https://www.gemini.edu/sciops/instruments/alopeke-zorro/>

made on UT March 25 2019 in the standard 3-point dither pattern that is used with NIRC2 to avoid the left lower quadrant of the detector, which is typically noisier than the other three quadrants. The dither pattern step size was 3 arcsec and was performed three times.

The observations were made in the Ks filter ($\lambda_o = 2.196$; $\Delta\lambda = 0.336\mu\text{m}$) with an integration time of 1 second and 20 coadds per frame for a total of 300 seconds on target. The camera was in the narrow-angle mode with a full field of view of ~ 10 arcsec and a pixel scale of 0.099442 arcsec per pixel. The Keck AO observations revealed no additional stellar companions detected to within a resolution ~ 0.05 arcsec FWHM (Fig. 3.1).

The sensitivities of the final combined AO image were determined by injecting simulated sources azimuthally around the primary target every 45° at separations of integer multiples of the central source's FWHM (Furlan et al., 2017). The brightness of each injected source was scaled until standard aperture photometry detected it with 5σ significance. The resulting brightness of the injected sources relative to the target set the contrast limits at that injection location. The final 5σ limit at each separation was determined from the average of all of the determined limits at that separation and the uncertainty on the limit was set by the rms dispersion of the azimuthal slices at a given radial distance. The sensitivity curve is shown in Fig. 3.1 (top), along with an image centred on the primary target showing no other companion stars (bottom).

Unbound Blended Source Confidence (BSC) analysis

We finally analyse all contrast light curves available for this target to estimate the probability of contamination from unbound blended sources in the *TESS* aperture that are undetectable from the available high-resolution images. This probability is called the Blended Source Confidence (BSC), and the steps for estimating it are fully described in Lillo-Box et al. (2014a). We use a Python implementation of this approach (bsc, by J. Lillo-Box) which uses the TRILEGAL³ galactic model (v1.6 Girardi et al., 2012) to retrieve a simulated source population of the region around the corresponding target⁴. This is used to compute the density of stars around the target position (radius $r = 1^\circ$), and to derive the probability of chance alignment at a given contrast magnitude and separation. We used the default parameters for the bulge, halo, thin/thick disks, and the lognormal initial mass function from Chabrier (2001).

The contrast curves of the high-spatial resolution images are used to constrain this parameter space and estimate the final probability of undetected potentially contaminating sources. We consider as potentially contaminating sources those with a maximum contrast magnitude corresponding to $\Delta m_{\text{max}} = -2.5 \log \delta$, with δ being the transit depth of

³<http://stev.oapd.inaf.it/cgi-bin/trilegal>

⁴This is done in Python by using the Bhatti et al. (2020) implementation.

the candidate planet in the *TESS* band. This offset from the target star magnitude gives the maximum magnitude that a blended star can have to mimic this transit depth. We convert the depth in the *TESS* passband to each filter (namely 562 nm and 832 nm for the Gemini/Alopeke images and Ks for the rest) by using simple conversions using the TIC catalogue magnitudes and linking the 562 nm filter to the SDSSr band, the 832 nm filter to the SDSSz band and the Ks band to the 2MASS Ks filter. The corresponding conversions imply $\Delta m_{562\text{ nm}} = 0.954\Delta m_{\text{TESS}}$, $\Delta m_{832\text{ nm}} = 0.920\Delta m_{\text{TESS}}$, and $\Delta m_{\text{Ks}} = 0.919\Delta m_{\text{TESS}}$. In Fig. 3.2 we show an example of the BSC calculation for the Keck/NIRC2 image that illustrates the method.

We applied this technique to TOI-431. The transits of the two planets in this system could be mimicked by blended eclipsing binaries with magnitude contrasts up to $\Delta m_{\text{b,max}} = 6.65$ mag and $\Delta m_{\text{d,max}} = 8.76$ mag in the *TESS* passband. This analysis is then especially relevant for the smallest planet in the system as the probability of a chance-aligned star increases rapidly with fainter magnitudes. However, the high quality of the high-spatial resolution images provide a very low probability for an undetected source capable of mimicking the transit signal. For TOI-431 b, we find 0.034 per cent (Alopeke/562 nm), 0.019 per cent (Alopeke/832 nm), 0.13 per cent (Keck/NIRC2/Ks), and 0.54 per cent (Gemini-North/NIRI/Ks). For TOI-431 d we find 0.009 per cent (Alopeke/562 nm), 0.002 per cent (Alopeke/832 nm), 0.04 per cent (Keck/NIRC2/Ks), and 0.16 per cent (Gemini-North/NIRI/Ks).

3.2.4 Stellar analysis

The parameters of the host star are required in order to derive precise values for the planetary ages, as well as the masses and radii, leading to bulk densities. This requires a good spectrum with high enough signal-to-noise and high spectral resolution. Our radial velocity spectra fulfil these requirements after co-adding the 124 individual HARPS spectra, resulting in a spectrum with a signal-to-noise of about 380 per pixel at 5950Å. We perform 2 independent spectroscopic analysis methods to derive the host star parameters, presented in Table 3.4, and further SED fitting.

Method 1: equivalent widths with ARES+MOOG:

The stellar atmospheric parameters (T_{eff} , $\log g$, microturbulence, and [Fe/H]) and respective error bars were derived using the methodology described in Santos et al. (2013); Sousa (2014). In brief, we make use of the equivalent widths (EW) of iron lines, as measured in the combined HARPS spectrum of TOI-431 using the ARES v2 code⁵ (Sousa et al., 2015),

⁵The last version of ARES code (ARES v2) can be downloaded at <http://www.astro.up.pt/~sousasag/ares>.

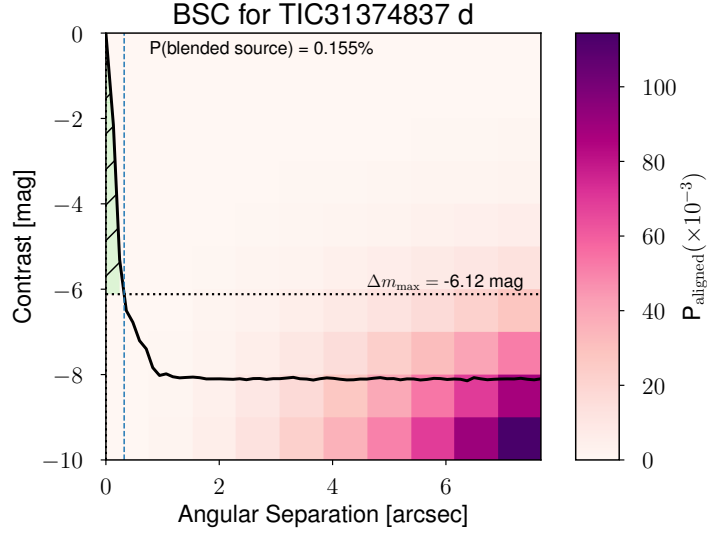


Figure 3.2: Contrast curve of TOI-431 from the Keck/NIRC2 instrument for the Ks filter (solid black line). The colour (P_{aligned}) on each angular separation and contrast bin represents the probability of a chance-aligned source with these properties at the location of the target, based on TRILEGAL model (see Sect. 3.2.3 within the main text). The maximum contrast of a blended binary capable of mimicking the planet transit depth is shown as a dotted horizontal line. The hatched green region between the contrast curve and the maximum contrast of a blended binary (Δm_{\max} line) represents the non-explored regime by the high-spatial resolution image. $P(\text{blended source})$ is the Blended Source Confidence (BSC), and this corresponds to the integration of P_{aligned} over the shaded region.

Table 3.4: Stellar parameters for TOI-431. Section references describing the method used to find the parameters are given in the Table footer.

Parameter (unit)	Value	Ref
Effective temperature T_{eff} (K)	4850 ± 75	1
Surface gravity $\log g$ (cgs)	4.60 ± 0.06	1
Microturbulence $v_{\text{tur,mic}}$ (km s ⁻¹)	0.8 ± 0.1 (fixed)	1
Macroscopic turbulence $v_{\text{tur,mac}}$ (km s ⁻¹)	0.5 ± 0.1 (fixed)	1
Bolometric flux F_{bol} (10 ⁻⁹ erg s ⁻¹ cm ⁻²)	7.98 ± 0.19	2
Stellar radius R_* (R _⊙)	0.731 ± 0.022	2
Stellar mass M_* (M _⊙)	0.78 ± 0.07	2
Rotation period P_{rot} (days)	30.5 ± 0.7	3
$v \sin i_*$ (km s ⁻¹)	2.5 ± 0.6	1
Chemical Abundances (dex)	Value	Ref
Metallicity [Fe/H]	0.2 ± 0.05	1
[Na/H]	0.22 ± 0.14	4
[Mg/H]	0.10 ± 0.07	4
[Al/H]	0.21 ± 0.10	4
[Si/H]	0.11 ± 0.13	4
[Ca/H]	0.06 ± 0.15	4
[Ti/H]	0.17 ± 0.17	4
[Cr/H]	0.12 ± 0.11	4
[Ni/H]	0.14 ± 0.08	4

1: Section 3.2.4

2: Section 3.2.4

3: From WASP-South, see Section 3.2.5

4: Section 3.2.4

and we assume ionization and excitation equilibrium. The process makes use of a grid of Kurucz model atmospheres (Kurucz, 1993) and the radiative transfer code MOOG (Snedden, 1973). This analysis results in values of effective temperature $T_{\text{eff}} = 4740 \pm 94$ K, surface gravity $\log g = 4.20 \pm 0.27$, microturbulence $v_{\text{tur,mic}} = 0.62 \pm 0.28$, and metallicity $[\text{Fe}/\text{H}] = 0.06 \pm 0.04$ dex. The value for $\log g$ can be corrected according to Mortier et al. (2014), to give 4.46 ± 0.27 (corrected for asteroseismology $\log g$ values) and 4.63 ± 0.28 (corrected for transit $\log g$ values).

Stellar abundances of the elements were derived using the classical curve-of-growth analysis method assuming local thermodynamic equilibrium (e.g. Adibekyan et al., 2012, 2015; Delgado Mena et al., 2017). For the abundance determinations we used the same tools and models as for stellar parameter determination. Unfortunately, due to the low T_{eff} of this star, we could not determine reliable abundances of carbon and oxygen. The derived abundances are presented in Table 3.4 and they are normal for a star with a metallicity close to solar.

In addition, we derived an estimated age by using the ratios of certain elements (the so-called chemical clocks) and the formulas presented in Delgado Mena et al. (2019). Since this star has a close to solar metallicity and is very cool (and thus probably outside the applicability limits of formulas using stellar parameters in addition to the chemical clock) we chose to use the 1D formulas presented in Table 5 of Delgado Mena et al. (2019). Due to the high error in Sr abundances we derived ages only from the abundance ratios $[\text{Y}/\text{Mg}]$, $[\text{Y}/\text{Zn}]$, $[\text{Y}/\text{Ti}]$, $[\text{Y}/\text{Si}]$, $[\text{Mg}/\text{Fe}]$, $[\text{Ti}/\text{Fe}]$, $[\text{Si}/\text{Fe}]$ and $[\text{Zn}/\text{Fe}]$. The abundance errors of cool stars are quite large and in turn the individual age errors of each chemical clock are also large ($\gtrsim 3$ Gyr) but the dispersion among them is smaller. We obtained a weighted average age of 5.1 ± 0.6 Gyr which is significantly older than the age obtained in Section 3.2.4. Nevertheless, we note that ages for very cool stars obtained from chemical clocks are affected by large errors and must be taken with caution.

Method 2: synthesis of the entire optical spectrum

We also derived stellar properties by analysing parts of the optical spectrum in a different way by comparing the normalized, co-added spectrum with modelled synthetic spectra obtained with the Spectroscopy Made Easy (SME) package (Valenti and Piskunov, 1996) version 5.22, with atomic parameters from the VALD database (Piskunov et al., 1995). The 1-D, plane-parallel LTE synthetic spectra are calculated using stellar parameters obtained from either photometry or a visual inspection of the spectrum as a starting point. The synthetic spectrum is automatically then compared to a grid of stellar atmospheric models. The grid we used in this case is based on the MARCS models (Gustafsson et al., 2008). An iterative χ^2 minimization procedure is followed until no improvement is achieved. We refer

to recent papers, e.g., Persson et al. (2018) and Gustafsson et al. (2008), for details about the method. In order to limit the number of free parameters we used empirical calibrations for the $v_{\text{tur,mic}}$ and $v_{\text{tur,mac}}$ turbulence velocities (Bruntt et al., 2010). The value of T_{eff} was determined from fitting the Balmer $H\alpha$ line wings. We used the derived T_{eff} to fit a large sample of [Fe I], Mg I and Ca I lines, all with well established atomic parameters in order to derive the abundance, [Fe/H], the rotation, and the surface gravity, $\log g$. We found the star to be slowly rotating, with $v \sin i_{\star} = 2.5 \pm 0.6 \text{ km s}^{-1}$. The star is cool, and the effective temperature as derived from the $H\alpha$ line wings is $T_{\text{eff}} = 4846 \pm 73 \text{ K}$. Using this value for T_{eff} , we found the [Fe/H] to be 0.20 ± 0.05 and the surface gravity $\log g$ to be 4.60 ± 0.06 (Table 3.4).

In order to check our result, we also analysed the same co-added spectrum using the public software package SpecMatch-Emp (Yee et al., 2017). This program extracts part of the spectrum and attempts to match it to a library of about 400 well characterized spectra of all types. Our input spectrum has to conform to the format of SpecMatch-Emp and we refer to Hirano et al. (2018) to describe our procedure for doing this. We derive a T_{eff} of $4776 \pm 110 \text{ K}$, an [Fe/H] of $0.15 \pm 0.09 \text{ dex}$, and a stellar radius of $R_{\star} = 0.76 \pm 0.18 R_{\odot}$. The former two values are in good agreement with the results from the SME analysis.

Because of the higher precision in the SME analysis, the final adopted value of T_{eff} for TOI-431 is $4850 \pm 75 \text{ K}$. Note that the error here is the internal errors in the synthesis of the spectra and does not include the inherent errors of the model grid itself, as well as those errors caused by using 1-D models.

The results from this method are in agreement with those found in Section 3.2.4, with T_{eff} and [Fe/H] (using SpecMatch-Emp) agreeing within error. The value for $\log g$ also agrees with the corrected $\log g$ values from the previous method. We therefore adopt the results from this method to take forward.

SED fitting

As an independent check on the derived stellar parameters, and in order to determine an estimate for stellar age, we performed an analysis of the broadband spectral energy distribution (SED). Together with the *Gaia* EDR3 parallax, we determine an empirical measurement of the stellar radius following the procedures described in Stassun and Torres (2016a); Stassun et al. (2017, 2018). We pulled the $B_T V_T$ magnitudes from *Tycho-2*, the *grizy* magnitudes from Pan-STARRS, the JHK_S magnitudes from 2MASS, the W1–W4 magnitudes from *WISE*, and the $GG_{\text{RP}}G_{\text{BP}}$ magnitudes from *Gaia*. Together, the available photometry spans the full stellar SED over the wavelength range $0.35\text{--}22 \mu\text{m}$ (see Figure 3.3). In addition, we pulled the NUV flux from *GALEX* in order to assess the level of chromospheric activity, if any.

We performed a fit using Kurucz stellar atmosphere models, with the effective temperature (T_{eff}) and metallicity ($[\text{Fe}/\text{H}]$) adopted from the spectroscopic analysis (Section 3.2.4). The extinction (A_V) was set to zero because of the star being very nearby (Table 3.1). The resulting fit is excellent (Figure 3.3) with a reduced χ^2 of 3.3 (excluding the *GALEX* NUV flux, which is consistent with a modest level of chromospheric activity; see below). Integrating the (unreddened) model SED gives the bolometric flux at Earth of $F_{\text{bol}} = 7.98 \pm 0.19 \times 10^{-9} \text{ erg s}^{-1} \text{ cm}^{-2}$. Taking the F_{bol} and T_{eff} together with the *Gaia* EDR3 parallax, with no systematic offset applied (see, e.g., Stassun and Torres, 2021), gives the stellar radius as $R = 0.731 \pm 0.022 R_{\odot}$. Finally, estimating the stellar mass from the empirical relations of Torres et al. (2010) and a 6 per cent error from the empirical relation itself gives $M = 0.77 \pm 0.05 M_{\odot}$, whereas the mass estimated empirically from the stellar radius together with the spectroscopic $\log g$ gives $M = 0.78 \pm 0.07 M_{\odot}$.

We can also estimate the stellar age by taking advantage of the observed chromospheric activity together with empirical age-activity-rotation relations. For example, taking the chromospheric activity indicator $\log R'_{HK} = -4.69 \pm 0.05$ from the archival HARPS data and applying the empirical relations of Mamajek and Hillenbrand (2008) gives a predicted age of $1.9 \pm 0.3 \text{ Gyr}$. Finally, we can further corroborate the activity-based age estimate by also using empirical relations to predict the stellar rotation period from the activity. For example, the empirical relation between R'_{HK} and rotation period from Mamajek and Hillenbrand (2008) predicts a rotation period for this star of $29.8 \pm 3.7 \text{ d}$, which is compatible with the rotation period inferred from the WASP-South observations (see Section 3.2.5). All of the stellar parameter values derived in this section can also be found in Table 3.4.

3.2.5 Stellar activity monitoring

Two instruments were used during different time periods to monitor TOI-431 in order to investigate the rotation period of the star. This is important to disentangle the effect of stellar activity when fitting for any planets present in the system.

WASP-South, located in Sutherland, South Africa, was the southern station of the WASP transit survey (Pollacco et al., 2006). The data reported here were obtained while WASP-South was operating as an array of 85mm, f/1.2 lenses backed by 2048x2048 CCDs, giving a plate scale of 32 arcsec per pixel. The observations spanned 180 days in 2012, 175 days in 2013 and 130 days in 2014. Observations on clear nights, with a typical 10-min cadence, accumulated 52 800 photometric data points.

We searched the datasets for rotational modulations, both separately and by combining the three years, using the methods described by Maxted et al. (2011). We detect a persistent modulation with an amplitude of 3 mmag and a period of $30.5 \pm 0.7 \text{ d}$ (where the error makes allowance for phase shifts caused by changing starspot patterns). The

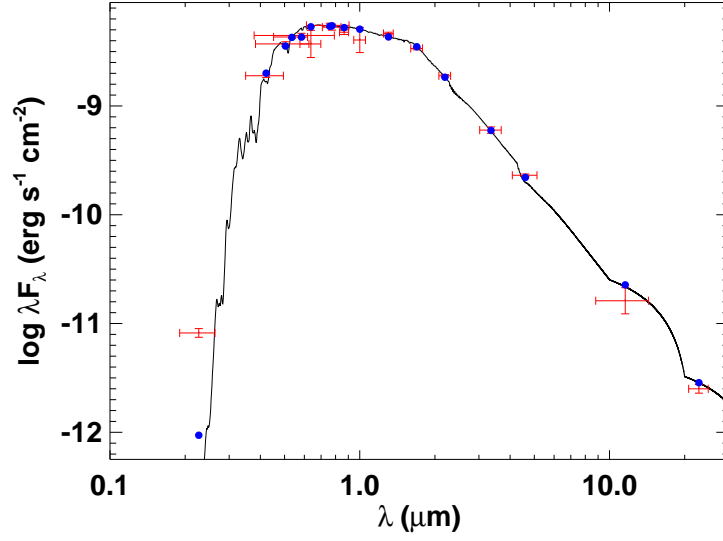


Figure 3.3: Spectral energy distribution (SED) of TOI-431. Red symbols represent the observed photometric measurements, where the horizontal bars represent the effective width of the passband. Blue symbols are the model fluxes from the best-fit Kurucz atmosphere model (black).

periodogram from the combined 2012–2014 data is shown in Fig. 3.4. The modulation is significant at the 99.9 per cent level (estimated using methods from Maxted et al. 2011). In principle, it could be caused by any star in the 112 arcsec photometric extraction aperture, but all the other stars are more than 4 magnitudes fainter.

Given the near-30-day timescale, we need to consider the possibility of contamination by moonlight. To check this, we made identical analyses of the light curves of 5 other stars of similar brightness nearby in the same field. None of these show the 30.5 d periodicity.

A single NGTS telescope was used to monitor TOI-431 between the dates of 2019 October 11 and 2020 January 20. During this time period a total of 79 011 images were taken with an exposure time of 10 seconds using the custom NGTS filter (520 - 890 nm). This data shows a significant periodicity at 15.5 days, at approximately half the period of the WASP-South modulation.

As the WASP-South period agrees with the activity signal we see in the HARPS data (see Fig. 3.5), we therefore take the 30.5 d period value forward.

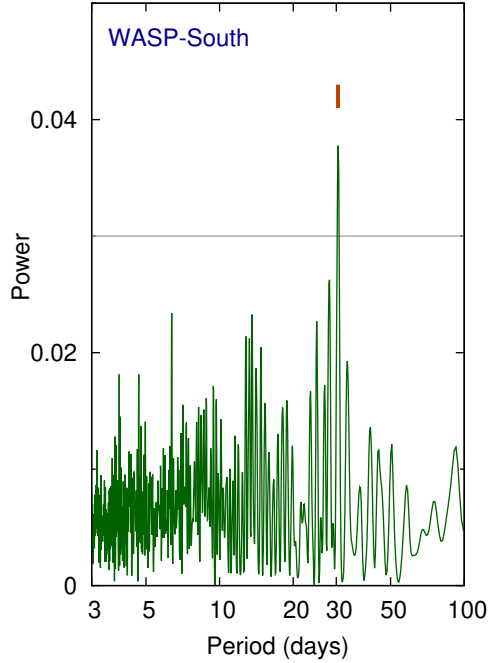


Figure 3.4: The periodogram of the WASP-South data for TOI-431 from 2012–2014. The orange tick is at 30.5 d, while the horizontal line is at the estimated 1 per cent false-alarm probability.

3.3 The joint fit

3.3.1 The third planet found in the HARPS data

We initially ran a joint fit which included only the planets flagged by the *TESS* pipelines, i.e. TOI-431 b and d. We then removed the signals of these planets from the raw HARPS radial velocities, and examined the residuals. This led to the discovery of an independent sinusoidal signal being seen as a significant peak in a periodogram of the residuals. This is shown in Fig. 3.5: from the periodogram of the raw RV data produced on DACE⁶, signals from TOI-431 b and d can be seen at 0.491 and 12.57 d respectively, with false-alarm probabilities (FAP) of < 0.1 per cent. A large signal can also be seen at 29.06 d; this is near the rotation period of the star found with WASP-South (see Section 3.2.5). Removing the fit for these two planets and the stellar activity reveals another signal at 4.85 d which does not correlate with any of the activity indicators (FWHM, BIS, S-Index and H α -Index; see Fig. 3.6 and Fig. 3.7 for periodograms of these indicators for both the current and archival HARPS data, respectively), and which is not an alias of the other planetary signals.

Phase folding the *TESS* photometry on the RV period reveals no transit (see Fig. 3.8,

⁶The DACE platform is available at <https://dace.unige.ch>

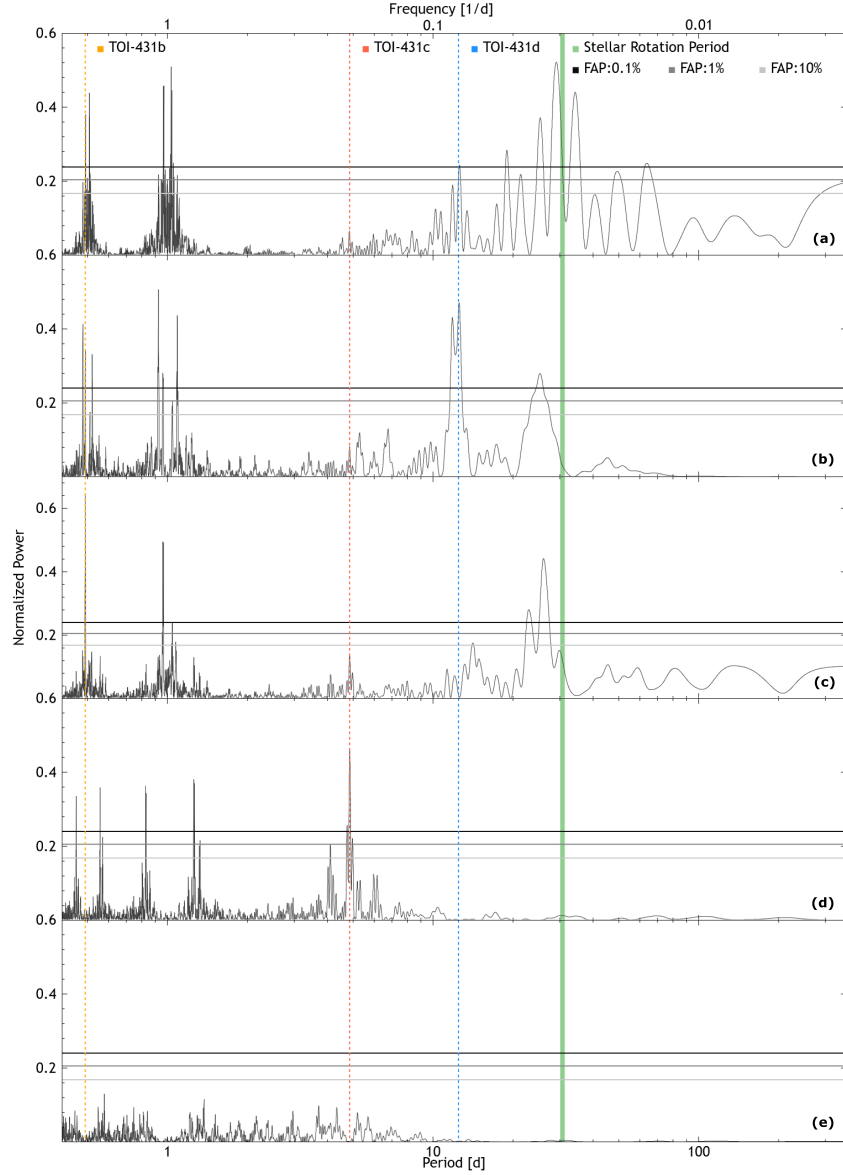


Figure 3.5: Periodograms for the HARPS data, where (going from top to bottom) the highest power peak has been sequentially removed until there is no power left. The best fit periods (see Table 3.6) of TOI-431 b (yellow), c (red), and d (blue), have been denoted by dotted lines, and the 1 standard deviation interval of the rotation period of the star has been shaded in green. The periodogram for the raw RV data is shown in panel (a); (b) has the stellar activity GP model removed; (c) has the best fit model for planet d removed also. Panel (d) has planet b removed, meaning that there should be no further power left. However, there is a peak evident at 4.85 days above the 0.1 per cent FAP that does not correlate with any stellar activity indicators, and it is not an alias of any other peaks. Taking this as an extra planet in the system (TOI-431 c) and removing the best fit model for this leaves a periodogram with no further signals, shown in panel (e).

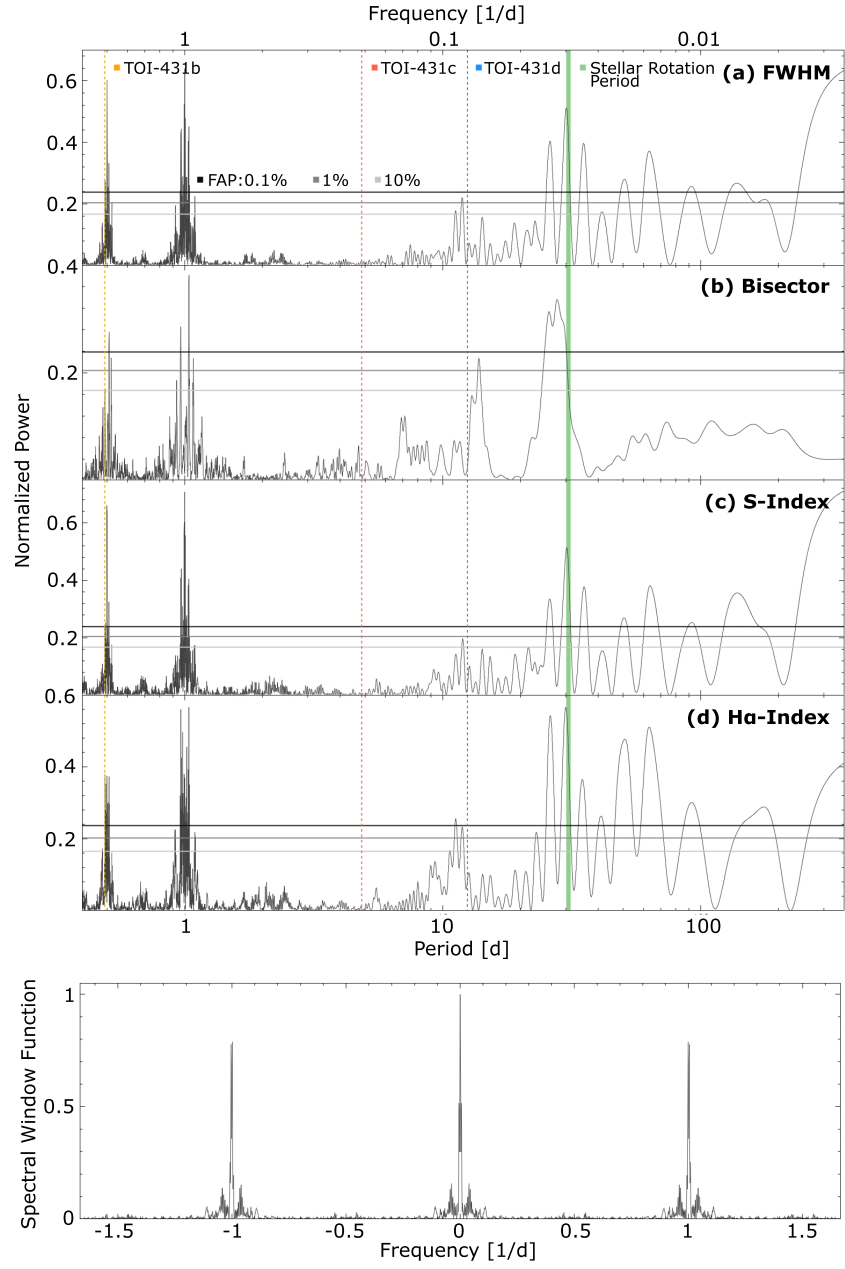


Figure 3.6: Periodograms for the activity indicators (top row) and window functions (bottom row) from the purpose-collected HARPS data from February to October 2019 (left), illustrating that there is no significant power at the 4.85 day period of TOI-431 c. The best fit periods (see Table 3.6) of TOI-431 b (yellow), c (red), and d (blue), have been denoted by dotted lines, and the 1 standard deviation interval of the rotation period of the star has been shaded in green.

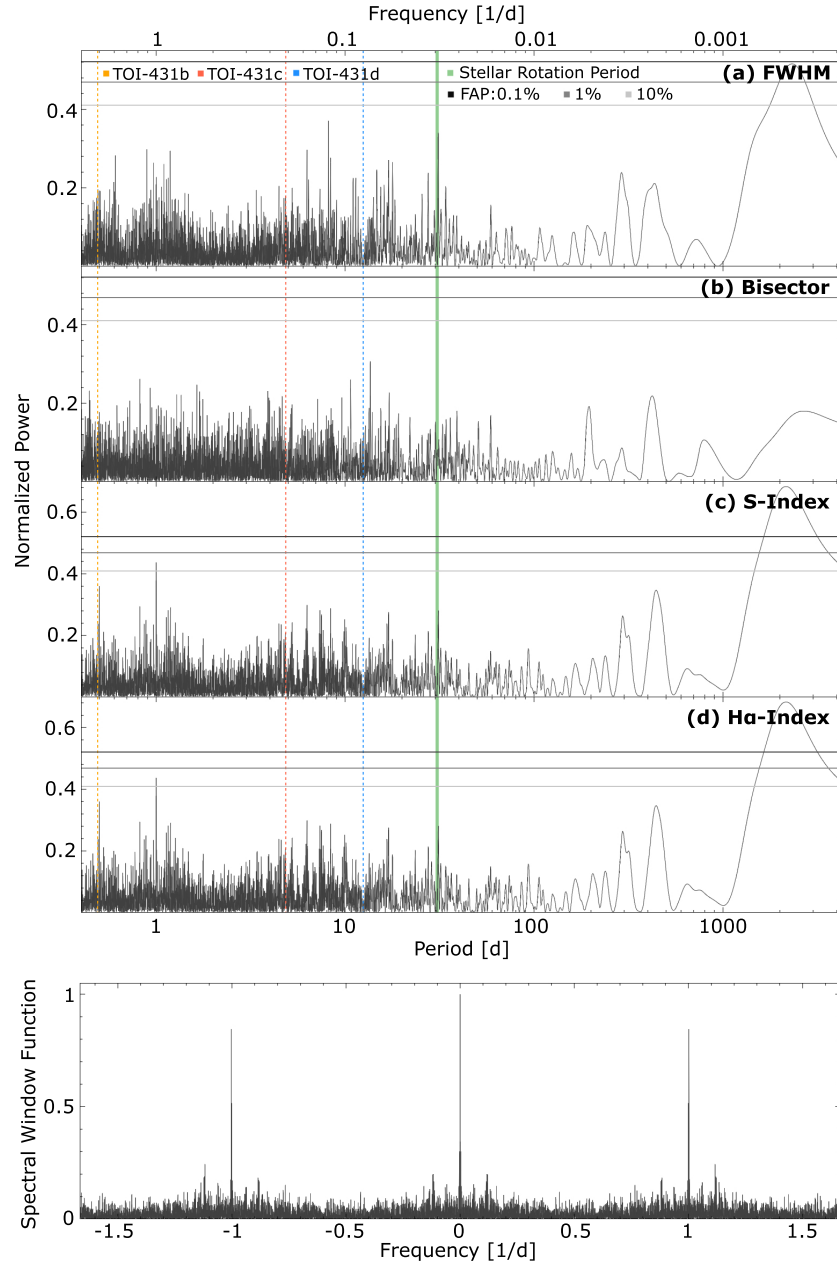


Figure 3.7: Periodograms for the activity indicators (top row) and window functions (bottom row) from the archival HARPS data from 2004 to 2015, illustrating that there is no significant power at the 4.85 day period of TOI-431 c. The best fit periods (see Table 3.6) of TOI-431 b (yellow), c (red), and d (blue), have been denoted by dotted lines, and the 1 standard deviation interval of the rotation period of the star has been shaded in green.

bottom plot, middle panel). We also attempted to use Transit Least Squares (TLS, Hippke and Heller, 2019) to recover this planet; it did not return any evidence of a transit at or near the RV period. As this planet is not evident in the *TESS* data, but is large enough to be detectable (see Section 3.3.3), we therefore make the assumption that it does not transit. As such, we conclude that this is a further, apparently non-transiting planet, and include it in the final joint fit model (described in Section 3.3.2) when fitting the RV data.

3.3.2 Construction of the joint fit model

Using the *EXOPLANET* package (Foreman-Mackey et al., 2020), we fit the photometry from *TESS*, LCOGT, NGTS, and *Spitzer* and the RVs from HARPS and HIRES simultaneously with Gaussian Processes (GPs) to remove the effects of stellar variability. *EXOPLANET* utilises the light curve modelling package *STARRY* (Luger et al., 2019), *PyMC3* (Salvatier et al., 2016), and *CELERITE* (Foreman-Mackey et al., 2017) to incorporate GPs. While we use a GP kernel included in the *EXOPLANET* package for the *TESS* data, we construct our own GP kernel using *PyMC3* for the HARPS and HIRES data. For consistency, all timestamps were converted to the same time system used by *TESS*, i.e. BJD - 2457000. All prior distributions set on the fit parameters of this model are given in Table 3.5.

Photometry

The flux is normalised to zero for all of the photometry by dividing the individual light curves by the median of their out-of-transit points and taking away one. To model the planetary transits, we used a limb-darkened transit model following the Kipping (2013) quadratic limb-darkening parameterisation, and Keplerian orbit models. This Keplerian orbit model is parameterised for each planet individually in terms of the stellar radius R_* in solar radii, the stellar mass M_* in solar masses, the orbital period P in days, the time of a reference transit t_0 , the impact parameter b , the eccentricity e , and the argument of periastron ω . While a similar Keplerian orbit model is parameterised for the third planet, b is not defined in this case as no transit is seen in the photometric data. We find the eccentricity of all planets to be consistent with 0: when eccentricity is a fit parameter in an earlier run of this model, we find the 95 per cent confidence intervals for the eccentricity of TOI-431 b, c and d to be 0 to 0.28, 0 to 0.22, and 0 to 0.31 respectively.

Therefore, we fix e and ω to 0 for all planets in the final joint fit model. These parameters are then input into light curve models created with *STARRY*, together with parameters for the planetary radii R_p , the time series of the data t , and the exposure time t_{exp} of the instrument. As we are modelling multiple planets and multiple instruments with different t_{exp} , a separate light curve model is thus created per instrument for the planets that

are expected to have a transit event during that data set. In some cases, TOI-431 b and d will have model light curves (e.g. in the *TESS* and *Spitzer* observations); in others (e.g. the LCOGT and NGTS observations), only TOI-431 d is expected to be transiting. TOI-431 c is not seen to transit, therefore we do not need to model it in this way. We use values from the *TESS* pipelines to inform our priors on the epochs, periods, transit depths and radii of the transiting planets.

Table 3.5: The prior distributions input into our joint fit model (described fully in Section 3.3), and the fit values resulting from the model. The priors are created using distributions in PyMC3, and the relevant inputs to each distribution are listed. The fit values are given as the median values of our samples, and the uncertainties are given as the 16th and 84th percentiles. Where necessary, the specific planet a parameter is describing is noted in square brackets.

Parameter	Prior Distribution	Fit Value	Parameter	Prior Distribution	Fit Value
Planets			Spitzer		
Period P [b] (days)	$\mathcal{N}(0.4900657, 0.001)$	$0.490047^{+0.000010}_{-0.000007}$	Jitter	$\mathcal{N}(337.0, 20.0)$	345 ± 8
Period P [c] (days)	$\mathcal{N}(4.849427, 0.1)$	$4.8494^{+0.0003}_{-0.0002}$	Pixel coefficient c_1	$\mathcal{N}(1236218, 10^5)$	$1448286^{+68271}_{-69627}$
Period P [d] (days)	$\mathcal{N}(12.46109, 0.01)$	12.46103 ± 0.00002	Pixel coefficient c_2	$\mathcal{N}(468921, 10^5)$	408211^{+14963}_{-14570}
Ephemeris t_0 [b] (BJD-2457000)	$\mathcal{N}(1627.533, 0.1)$	$1627.538^{+0.003}_{-0.002}$	Pixel coefficient c_3	$\mathcal{N}(-917568, 10^5)$	$-832924^{+62790}_{-62527}$
Ephemeris t_0 [c] (BJD-2457000)	$\mathcal{N}(1625.888, 0.1)$	1625.87 ± 0.10	Pixel coefficient c_4	$\mathcal{N}(465062, 10^5)$	428366^{+16837}_{-16824}
Ephemeris t_0 [d] (BJD-2457000)	$\mathcal{N}(1627.545, 0.1)$	1627.5453 ± 0.0003	Pixel coefficient c_5	$\mathcal{N}(693929, 10^5)$	688664^{+10881}_{-10749}
$\log(R_p)$ [b] (R_\odot)	$\mathcal{N}(-4.35^*, 1.0)$	-4.44 ± 0.03	Pixel coefficient c_6	$\mathcal{N}(554898, 10^5)$	542039^{+12467}_{-12391}
$\log(R_p)$ [d] (R_\odot)	$\mathcal{N}(-3.41^*, 1.0)$	-3.50 ± 0.03	Pixel coefficient c_7	$\mathcal{N}(-205010, 10^5)$	$-194425^{+61256}_{-59207}$
Star			Pixel coefficient c_8	$\mathcal{N}(564035, 10^5)$	522150^{+12762}_{-12784}
Mass (M_\odot)	$\mathcal{N}_{\mathcal{T}}(0.77, 0.7, 0.0, 3.0)$	0.81 ± 0.05	Pixel coefficient c_9	$\mathcal{N}(618285, 10^5)$	669652^{+22918}_{-22697}
Radius (R_\odot)	$\mathcal{N}_{\mathcal{T}}(0.729, 0.022, 0.0, 3.0)$	0.72 ± -0.02	Time dependent ramp coefficient f	$\mathcal{N}(0.0, 170000)$	2017^{+9457}_{-9651}
TESS			Time dependent ramp coefficient g	$\mathcal{N}(0.0, 170000)$	618^{+522}_{-518}
Mean	$\mathcal{N}(0.0, 1.0)$	0.00006 ± 0.00006	Offset constant h	$\mathcal{N}(0.0, 10^4)$	-1543^{+3755}_{-3760}
GP log ($s2$)	$\mathcal{N}(-15.257^\dagger, 0.1)$	-15.539 ± 0.008	HARPS and HIRES		
GP log ($w0$)	$\mathcal{N}(0.0, 0.1)$	0.19 ± 0.08	HARPS Offset	$\mathcal{N}(48830.87, 10.0)$	48828 ± 2
GP log ($Sw4$)	$\mathcal{N}(-15.257^\dagger, 0.1)$	-15.37 ± 0.09	log (Jitter _{HARPS})	$\mathcal{N}(-0.2661^\ddagger, 5.0)$	$-5.06^{+2.10}_{-3.37}$
LCOGT (ingress)			HIRES Offset	$\mathcal{N}(0.01, 10.0)$	-2.07 ± 2.34
Mean	$\mathcal{N}(0.0, 1.0)$	-0.00044 ± 0.00008	log (Jitter _{HIRES})	$\mathcal{N}(-0.2659^\ddagger, 5.0)$	$-0.05^{+0.36}_{-0.43}$
LCOGT (egress)			GP recurrence timescale T (stellar rotation period) (days)	$\mathcal{N}(30.5, 0.7)$	30.7 ± 0.6
Mean	$\mathcal{N}(0.0, 1.0)$	0.00002 ± 0.00006	GP amplitude η	$\mathcal{HC}(5.0)$	$5.48^{+1.12}_{-0.83}$
NGTS			GP lengthscale l_e	$\mathcal{N}_{\mathcal{T}}(30.0, 20.0, 25.0, -)$	$31.5^{+6.3}_{-4.2}$
Mean	$\mathcal{N}(0.0, 1.0)$	$-0.00015^{+0.00008}_{-0.00007}$	GP lengthscale l_p	$\mathcal{N}_{\mathcal{T}}(0.1, 10.0, 0.0, 1.0)$	$0.47^{+0.70}_{-0.09}$

Distribution descriptions:

$\mathcal{N}(\mu, \sigma)$: a normal distribution with a mean μ and a standard deviation σ ;

$\mathcal{N}_{\mathcal{B}}(\mu, \sigma, a, b)$: a bounded normal distribution with a mean μ , a standard deviation σ , an lower bound a , and an upper bound b (bounds optional);

$\mathcal{N}_{\mathcal{T}}(\mu, \sigma, a, b)$: a truncated normal distribution with a mean μ , a standard deviation σ , a lower bound a , and an upper bound b (bounds optional);

$\mathcal{HC}(\beta)$: a Half-Cauchy distribution with a single beta parameter β .

Prior values:

* equivalent to $0.5(\log(D)) + \log(R_*)$ where D is the transit depth and R_* is the value of the prior on the stellar radius (R_\odot);

† equivalent to the log of the variance of the *TESS* flux; ‡ equivalent to 2 times the log of the minimum error on the HARPS or HIRES RV data, respectively.

TESS

Both TOI-431 b and d are transiting in the *TESS* light curve, so we first create model light curves for each using *STARRY*.

As seen in Fig. 3.8, the *TESS* Sector 5 and 6 light curves show some stellar variability. This variability was thus modelled with the SHOTerm GP given in *EXOPLANET* ⁷, which represents a stochastically-driven, damped harmonic oscillator. We set this up using the hyperparameters $\log(s2)$, $\log(Sw4)$, $\log(w0)$, and Q . The prior on Q was set to $1/\sqrt{2}$. Priors on $\log(s2)$ and $\log(Sw4)$ were set as normal distributions with a mean equal to the log of the variance of the flux and a standard deviation of 0.1. The prior on $\log(w0)$ was also set as a normal distribution but with a mean of 0 and a standard deviation of 0.1 (see Table 3.5).

We then take the sum of our model light curves and subtract these from the total PDCSAP flux, and this resultant transit-free light curve is the data that the GP is trained on to remove the stellar variability. The GP model can be seen in Fig. 3.8 (top plot, top panel), and the resultant best fit model in the middle panel. Further to this, phase folds of the *TESS* data for all planets in the system can also be seen in Fig. 3.8 (bottom plot), where TOI-431 c has been folded on its period determined from the radial velocity data, and no dip indicative of a transit can be seen.

LCOGT

No further detrending to that outlined in Section 5.2.1 was included for the LCOGT data. Only TOI-431 d is transiting in this data, so we create a model light curve of TOI-431 d using *STARRY* (as outlined above) per LCOGT dataset to produce 2 model light curves overall, as there are 2 transit events - an ingress and an egress - on separate nights. For each dataset, we use a normal prior with the model light curve as the mean and a standard deviation set to the error on the LCOGT data points, and this is then compared to the observed light curve. The best fit model for both the ingress and egress data is shown in Fig. 3.9 (top 2 panels).

NGTS

No further detrending was needed for the NGTS data after the pipeline reduction outlined in Section 3.2.1, and again, only TOI-431 d is evident in this data. Thus the same simple method used for the LCOGT data above is also applied here, creating a singular model light curve of TOI-431 d for the NGTS data and comparing this to the observed light curve, with a standard deviation set to the error on the NGTS data points. The best fit model for the NGTS data is shown in Fig. 3.9 (bottom panel).

⁷<https://docs.exoplanet.codes/en/stable/user/api/#exoplanet.gp.terms.SHOTerm>

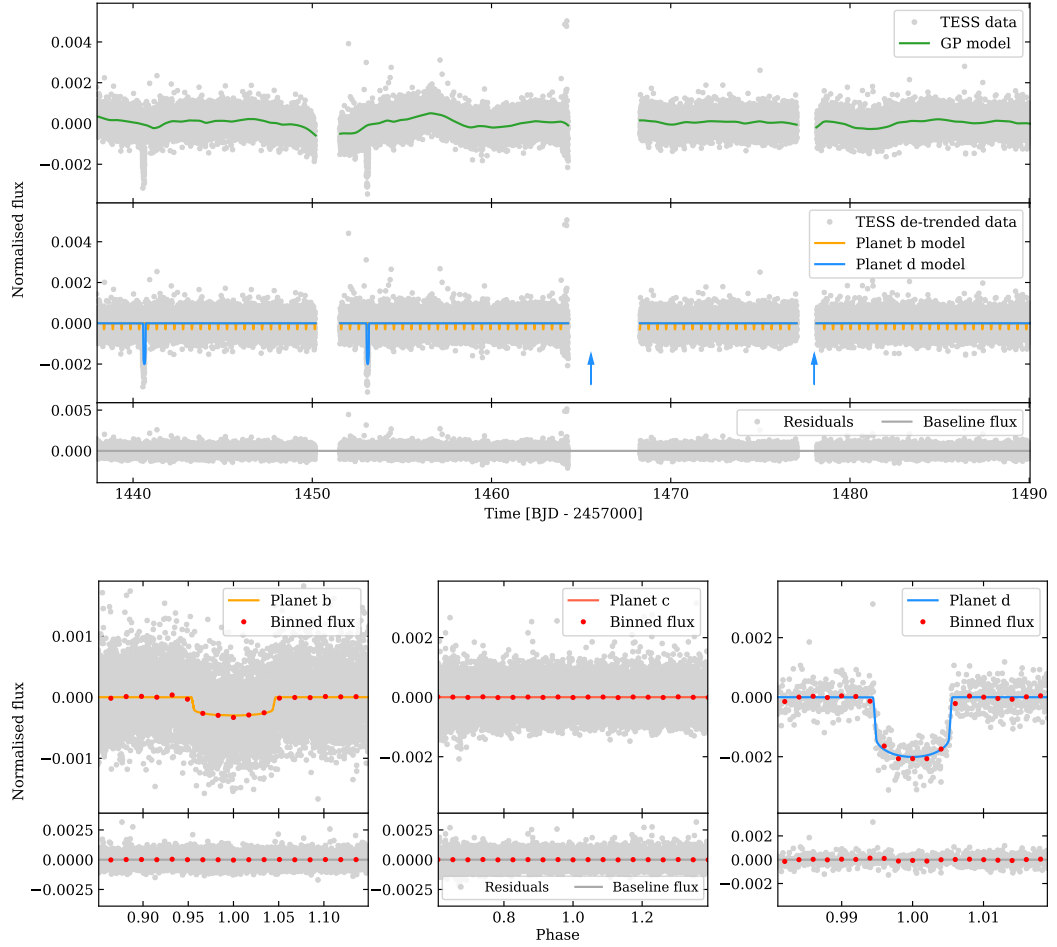


Figure 3.8: The *TESS* data for TOI-431 in Sectors 5 and 6. **Top plot:** detrending the *TESS* light curves and fitting models for TOI-431 b and c. *Top:* the full, 2-min cadence PDCSAP lightcurve, with no detrending for stellar activity, is shown in grey. Each sector has 2 segments of continuous viewing, and the gaps in the data correspond to the spacecraft down-linking the data to Earth after a *TESS* orbit of 13.7 days. Overlaid in green is the GP model that has been fit to this data (described in Section 5.4), in order to detrend the stellar activity. *Middle:* the flux detrended with the GP model, with the transit models for TOI-431 b (orange) and d (blue) overlaid. The expected transit times for the 2 further transits of TOI-431 d, both of which fall in the data down-link, are marked with blue arrows. *Bottom:* residuals when the best fit model and GP have been subtracted from the PDCSAP flux. The baseline flux (normalised to 0) is shown in dark grey. **Bottom plot:** phase folds of the *TESS* data for TOI-431 b (left), c (middle, with no transit evident), and d (right), with the flux binned as red circles, and the residuals of the folds once the best fit models have been subtracted from the flux shown in the bottom panels.

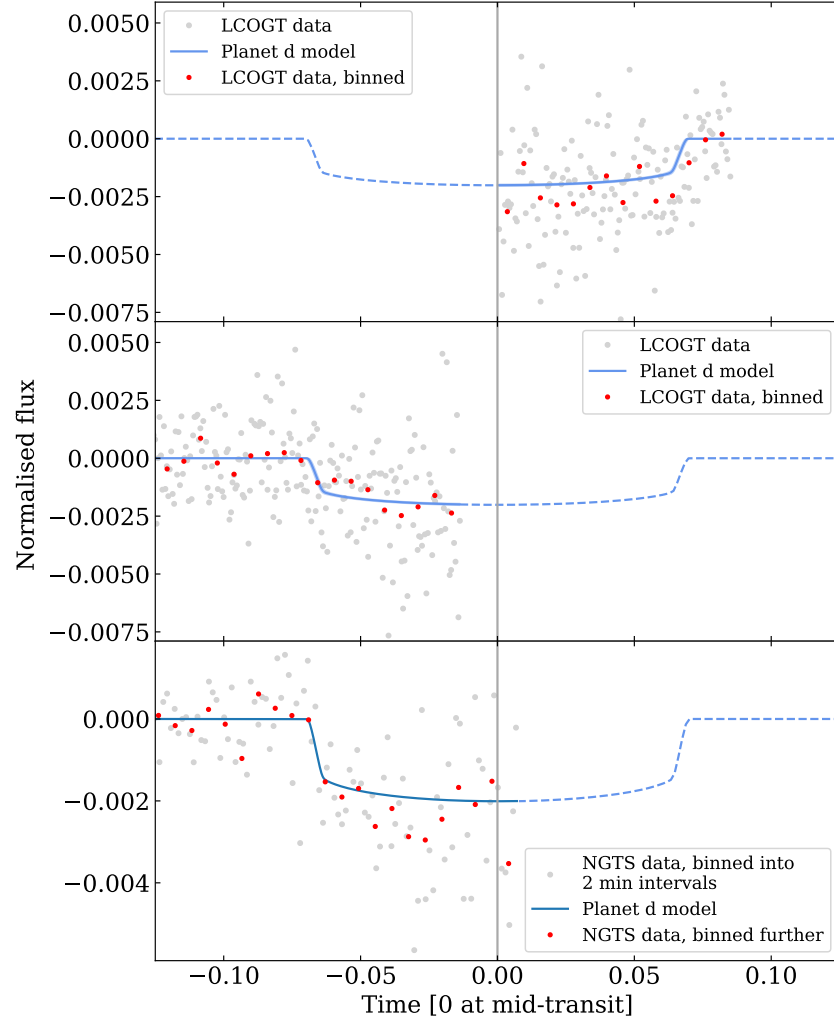


Figure 3.9: Best fit models of TOI-431 d to the LCOGT ingress (top), egress (middle) and NGTS light curves (bottom). In the LCOGT panels (top and middle), the observed flux is shown as light grey circles, the binned flux as red circles. In the NGTS panel (bottom), the flux is binned to 2 minute intervals in light grey. In all panels, the fit model is given as the blue line, solid where there are photometry points and dashed where there are not.

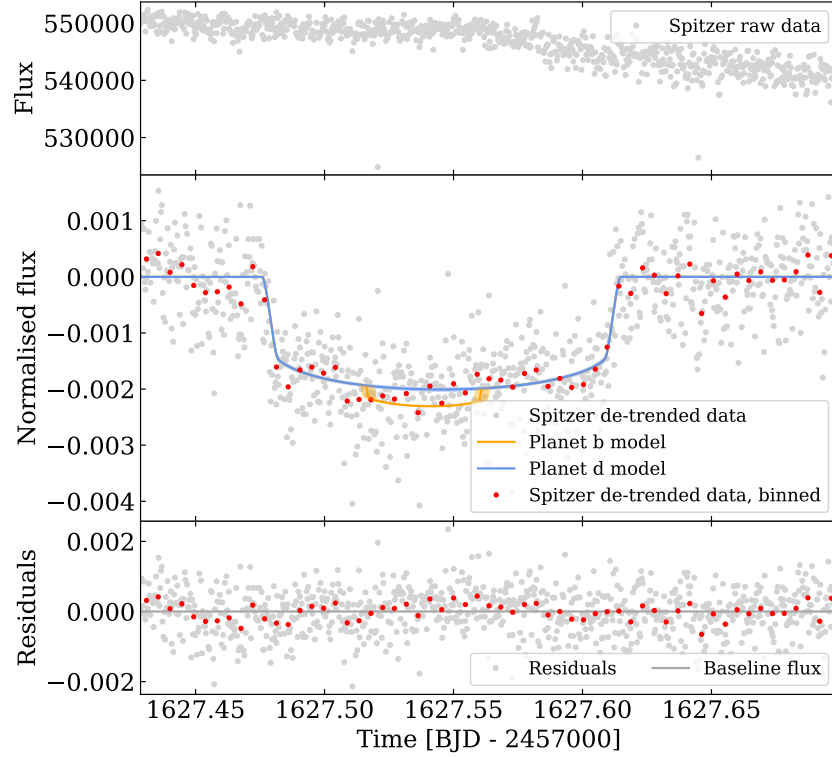


Figure 3.10: The *Spitzer* double-transit. *Top*: the raw *Spitzer* data, without any PLD applied. *Middle*: the *Spitzer* light curve detrended with PLD in grey and binned as red circles, with the best fit models of planet b (orange) and d (blue) overlaid. *Bottom*: the residuals when the best fit model has been subtracted from the detrended flux.

***Spitzer* and Pixel Level Decorrelation**

For the *Spitzer* double-transit observation, model light curves are created for both TOI-431 b and d. *Spitzer* data is given as N pixel values on a grid; in this instance, the grid is 3x3 pixels as in figure 1 of Deming et al. (2015). We follow the Pixel Level Decorrelation (PLD) method of Deming et al. (2015) (summarised below) to remove the systematic effect caused by intra-pixel sensitivity variations. Together with pointing jitter, these variations mask the eclipses of exoplanets in the photometry with intensity fluctuations that must be removed. We outline our PLD implementation as follows:

First, the intensity of pixel i at each time step t , i.e. P_i^t , is normalised such that the sum of the 9 pixels at one time step is unity, thus removing any astrophysical variations:

$$\hat{P}_i^t = \frac{P_i^t}{\sum_{i=1}^N P_i^t}. \quad (3.1)$$

PLD makes the simplification that the total flux observed can be expressed as a

linear equation:

$$\Delta S^t = \sum_i^N c_i \hat{P}_i^t + DE(t) + ft + gt^2 + h, \quad (3.2)$$

where ΔS^t is the total fluctuation from all sources. The normalised pixel intensities are multiplied by some coefficient c_i , and summed with the eclipse model $DE(t)$, a quadratic function of time $ft + gt^2$ which represents the time-dependent “ramp”, and an offset constant h . We use the eclipse model set up earlier using EXOPLANET as $DE(t)$, where D is the eclipse depth. This allows us to remove the intra-pixel effect, while solving for the eclipse amplitude and temporal baseline effects. Overall, the PLD alone has 14 free parameters that we solve for: 9 pixel coefficients, the depth of eclipse and the eclipse model, 2 time coefficients, and an offset term.

We add an additional fit parameter by introducing a *Spitzer* “jitter” term. We can estimate a prior for this fit parameter by removing our best fit model from the total raw flux from *Spitzer*, and calculating the standard deviation of the residual flux, which is approximately 337 ppm.

Our overall model for the *Spitzer* data is the PLD terms multiplied by the sum of the individual light curve models for each planet, b and d. We use a normal distribution with this model as the mean and a standard deviation set by the jitter parameter, and this is fit to the observed *Spitzer* flux. This can be seen in Fig. 3.10.

RVs

We do not include the iSHELL, FEROS, or MINERVA-Australis RVs in our joint fit, as they were not found to improve the fit due to large error bars in comparison to the HARPS and HIRES data; however, they are shown to be consistent with the result of our fit (see Fig. 3.11). We also do not include the archival HARPS data due to a large scatter in cadence and quality in comparison to the purpose-collected HARPS data.

HARPS and HIRES fitting

In this joint fit model, we fit the HARPS and HIRES data using the same method and so they are described here in tandem. We first find predicted values of radial velocity for each planet at each HARPS and HIRES timestamp using EXOPLANET. We set a wide uniform prior on K for each planet, the uniform distributions centred upon K values found when fitting the RV data with simple Keplerian models for all of the planets in DACE. We fit separate “offset” terms for HARPS and HIRES to model the systematic radial velocity, giving this a normal prior with a mean value predicted in DACE. We also fit separate “jitter” terms, setting wide normal priors on these, the means of which are set to double the log of the minimum error on the HARPS and HIRES data respectively.

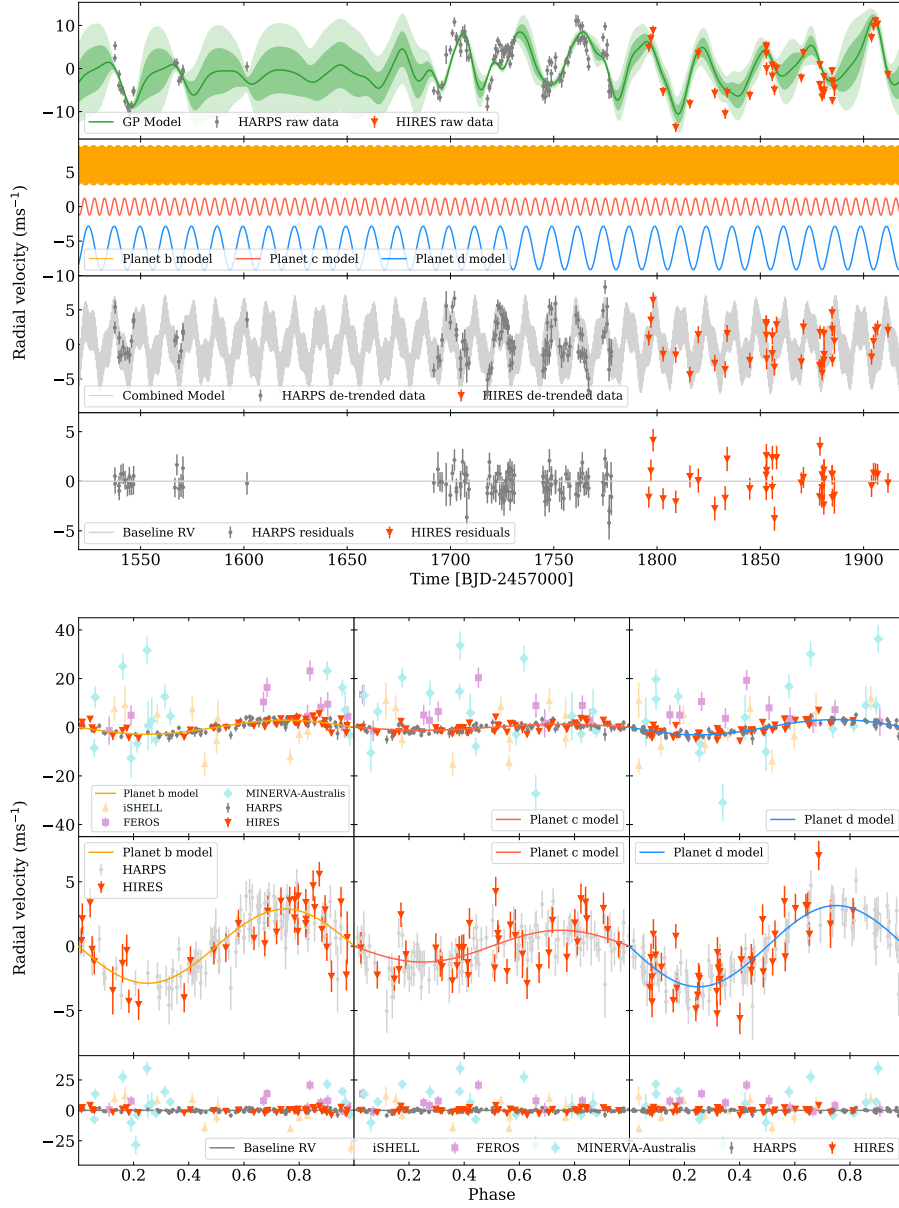


Figure 3.11: RV data plots, where the HARPS data is denoted as grey circles, HIRES as red upside down triangles, iSHELL as pale orange triangles, FEROS as pale pink squares, and MINERVA-Australis as pale turquoise diamonds. **Top plot:** the RV data, showing the GP and planet models that have been fit. *Top:* the best-fit GP used to detrend the stellar activity in the HARPS data is shown as the green line. The green shaded areas represent the 1 and 2 standard deviations of the GP fit. *Upper middle:* the separate models for each planet, b (orange, offset by $+6 \text{ m s}^{-1}$), c (red), and d (blue, offset by -6 m s^{-1}). *Lower middle:* the total model, representing the addition of the models for planets b, c, and d, is plotted in black, and over plotted is the HARPS and HIRES data. *Bottom:* the residuals after the total model, GP and offsets have been subtracted from the RV data. **Bottom plot:** the phase folds for each planet model, b (left), c (middle), and d (right), with the RV data over plotted. The top row shows all of the RV data (where the GP has been subtracted from each data set), the middle just the HARPS and HIRES data, and the bottom the residuals when the planet models have been subtracted from the RVs.

The RV data also shows significant stellar variability due to stellar rotation, and so we model this variability using another GP (see Fig. 3.11, top panel of top plot). This activity can be modelled as a Quasi-Periodic signal as starspots moving across the surface of the star evolve in time and are modulated by stellar rotation. In this case, we create our own Quasi-Periodic kernel using PyMC3, as no such kernel is available in EXOPLANET. PyMC3 provides a range of simple kernels⁸ which are easy to combine. We use their Periodic:

$$k(x, x') = \eta^2 \exp \left(-\frac{\sin^2(\pi|x - x'| \frac{1}{T})}{2l_p^2} \right), \quad (3.3)$$

and ExpQuad (squared exponential):

$$k(x, x') = \eta^2 \exp \left(-\frac{(x - x')^2}{2l_e^2} \right) \quad (3.4)$$

kernels. The hyperparameters are η (the amplitude of the GP), T (the recurrence timescale, equivalent to the P_{rot} of the star), l_p (the smoothing parameter), and l_e (the timescale for growth and decay of active regions) (see e.g. Rasmussen and Williams, 2006; Haywood et al., 2014; Grunblatt et al., 2015).

We multiply these kernels together to create our final Quasi-Periodic kernel:

$$k(x, x') = \eta^2 \exp \left(-\frac{\sin^2(\pi|x - x'| \frac{1}{T})}{2l_p^2} - \frac{(x - x')^2}{2l_e^2} \right). \quad (3.5)$$

We use the same GP to fit the HARPS and HIRES data together using the same hyperparameters. We use a normal distribution with a mean equal to the rotation period of the star found by WASP-South (see Section 3.2.5 and Table 3.4) to set a wide prior on T .

To bring everything together, we add the predicted radial velocities together with the offsets, and subtract these from their respective observed radial velocity values. This is then used as the prior on the GP, which is also given a noise term that is equal to an addition of the jitters with the squared error on the RV data.

3.3.3 Fit results

We first use `exoplanet` to maximise the log probability of the PyMC3 model. We then use the fit parameter values this obtains as the starting point of the PyMC3 sampler, which draws samples from the posterior using a variant of Hamiltonian Monte Carlo, the No-U-Turn Sampler (NUTS). By examining the chains from earlier test runs of the model, we allow for 1000 burn-in samples which are discarded, and 5000 steps with 15 chains. We present our best fit parameters for the TOI-431 system from our joint fit in Table 3.6.

⁸<https://docs.pymc.io/api/gp/cov.html>

TOI-431 b is a super-Earth with a mass of $3.07^{+0.35}_{-0.34} M_{\oplus}$ and a radius of $1.28 \pm 0.04 R_{\oplus}$, and from this we can infer a bulk density of $7.96^{+1.05}_{-0.99} \text{ g cm}^{-3}$. This puts TOI-431 b below the radius gap, and it is likely a stripped core with no gaseous envelope. A period of 0.49 days puts TOI-431 b in the rare Ultra-Short Period (USP) planet category (defined simply as planets with $P < 1$ day); examples of systems which have USP planets include Kepler-78 (Winn et al., 2018), WASP-47 (Becker et al., 2015), and 55 Cancri (Dawson and Fabrycky, 2010).

TOI-431 c has a minimum mass of $2.83^{+0.41}_{-0.34} M_{\oplus}$, but the lack of transits does not allow us to fit a radius. We can use the mass-radius relation via FORECASTER (Chen and Kipping, 2017) to estimate a radius of $1.44^{+0.60}_{-0.34} R_{\oplus}$, which would place this planet as another super-Earth.

TOI-431 d is a sub-Neptune with a mass of $9.90^{+1.53}_{-1.49} M_{\oplus}$ and a radius of $3.29^{+0.09}_{-0.08} R_{\oplus}$, implying a bulk density of $1.36^{+0.25}_{-0.24} \text{ g cm}^{-3}$. This lower density implies that TOI-431 d probably has a gaseous envelope. We further analyse these planets in the following section.

Table 3.6: The parameters for the planets TOI-431 b, c, and d, calculated from our joint fit model described fully in Section 3.3. The values are given as the median values of our samples, and the uncertainties are given as the 16th and 84th percentiles. The bulk densities are then calculated using the masses and radii, assuming a spherical planet of uniform density. A calculation of the radius of TOI-431 c can be found in Section 3.3.3, and discussion of the inclinations of the planets can be found in Section 3.4. The equilibrium temperature is calculated assuming an albedo of zero. Further joint fit model parameters to those presented here can be found in Table 3.5.

Parameter	TOI-431 b	TOI-431 c	TOI-431 d
Period P (days)	$0.490047^{+0.000010}_{-0.000007}$	$4.8494^{+0.0003}_{-0.0002}$	12.46103 ± 0.00002
Semi-major axis a (AU)	$0.0113^{+0.0002}_{-0.0003}$	0.052 ± 0.001	0.098 ± 0.002
Ephemeris t_0 (BJD-2457000)	$1627.538^{+0.003}_{-0.002}$	1625.9 ± 0.1	1627.5453 ± 0.0003
Radius R_p (R_\oplus)	1.28 ± 0.04	-	3.29 ± 0.09
Impact parameter b	$0.34^{+0.07}_{-0.06}$	-	$0.15^{+0.12}_{-0.10}$
Inclination i (degrees)	$84.3^{+1.1}_{-1.3}$	$< 86.35^{+0.04}_{-0.09}$	89.7 ± 0.2
Eccentricity e	0 (fixed)	0 (fixed)	0 (fixed)
The argument of periastron ω	0 (fixed)	0 (fixed)	0 (fixed)
Radial velocity semi-amplitude K ($m s^{-1}$)	2.88 ± 0.30	$1.23^{+0.17}_{-0.14}$	3.16 ± 0.46
Mass M_p (M_\oplus)	3.07 ± 0.35	$2.83^{+0.41}_{-0.34}$ ($M \sin i$)	$9.90^{+1.53}_{-1.49}$
Bulk density ρ ($g cm^{-3}$)	8.0 ± 1.0	-	1.36 ± 0.25
Equilibrium temperature T_{eq} (K)	1862 ± 42	867 ± 20	633 ± 14

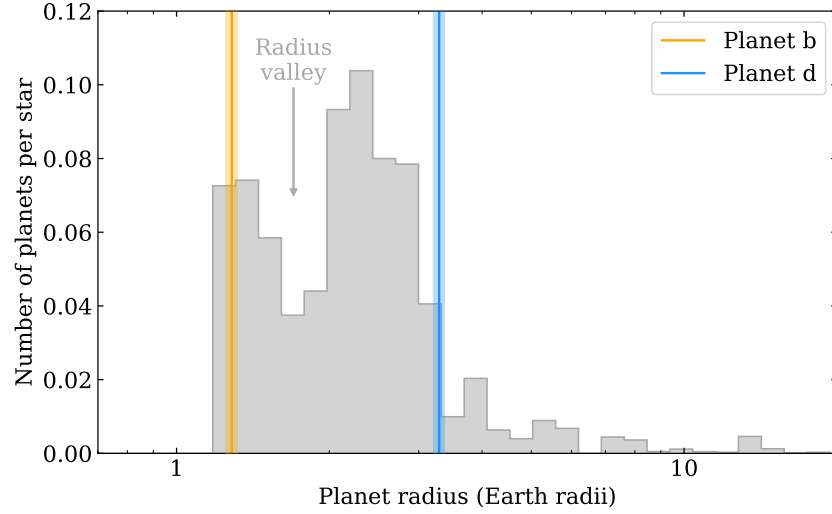


Figure 3.12: A histogram of planet radius for planets with orbital periods less than 100 days, as given in Fulton and Petigura (2018). The radius valley can be seen at $1.7 R_{\oplus}$: below the gap are rocky super-Earths, above the gap are gaseous sub-Neptunes. TOI-431 b (orange, with 1σ confidence intervals shaded) is the former, while TOI-431 d (blue) is the latter.

3.4 Discussion

The architecture of this system is unusual in that the middle planet, TOI-431 c, is non-transiting, while the inner and outer planets are both seen to transit. Examples of this can be seen in Kepler-20 (Buchhave et al., 2016), a 6-planet system where the fifth planet out from the star does not transit, but the sixth does, and HD 3167 (Vanderburg et al., 2016; Gandolfi et al., 2017; Christiansen et al., 2017), a 3-planet system where the middle planet does not transit as is the case with TOI-431. Using the impact parameter b from Table 3.6, we calculate inclinations for TOI-431 b and d of $(84.5^{+1.1}_{-1.3})^{\circ}$ and $89.7 \pm 0.2^{\circ}$, respectively (Table 3.6). We can calculate a limit on the inclination for TOI-431 c assuming $b = 1$, which results in an inclination that must be $< (86.35^{+0.04}_{-0.09})^{\circ}$ in order for TOI-431 c to be non-transiting.

The TOI-431 system is a good target system for studying planetary evolution. TOI-431 b and d reside either side of the radius-period valley described in Fulton et al. (2017); Fulton and Petigura (2018); Van Eylen et al. (2018) (see Fig. 3.12), providing a useful test-bed for the theorised mechanisms behind it. X-ray and EUV-driven photoevaporation is one of the two main proposed mechanisms (Owen and Wu, 2017), and we investigated its effect both now and in the past in the TOI-431 system. As no direct X-ray observations of the system exist, we had to make use of empirical formulae for relating the ratio of the X-ray and bolometric luminosities to age (Jackson et al., 2012) and Rossby number (related

to P_{rot} , Wright et al. (2011, 2018)). We extrapolate to the EUV using the relations of King et al. (2018). Under the assumption of energy-limited escape (Watson et al., 1981; Erkaev et al., 2007), we estimate a current mass loss rate for TOI-431 d between 5×10^8 and 5×10^9 g s^{-1} . The same assumptions yield a current rate of 10^{10} to 10^{11} g s^{-1} for TOI-431 b, but since that planet is unlikely to retain much, if any, atmosphere, the likely true rate is much lower.

Integrating the Jackson et al. (2012) relations across the lifetime of the star, and again assuming energy-limited escape, lifetime-to-date mass loss estimates of 44 per cent and 1.0 per cent for TOI-431 b and d respectively are found. Adding 2 per cent extra mass and doubling the radius to account for a primordial envelope around TOI-431 b raises the lifetime loss to 94 per cent. Again, the true value will be lower as XUV photoevaporation will not affect the rocky core, but rather the estimates calculated here demonstrate TOI-431 b would easily have lost a typical envelope with a mass fraction of a few per cent. The value for TOI-431 d is consistent with the density of the planet, which suggests it retains a substantial envelope.

In order to characterize the composition of TOI-431 b and TOI-431 d, we model the interior considering a pure-iron core, a silicate mantle, a pure-water layer, and a H-He atmosphere. The models follow the basic structure model of Dorn et al. (2017), with the equation of state (EOS) of the iron core taken from Hakim et al. (2018), the EOS of the silicate-mantle from Connolly (2009), and SCVH (Saumon et al., 1995) for the H-He envelope assuming protosolar composition. For water we use the QEOS of Vazan et al. (2013) for low pressures and the one of Seager et al. (2007) for pressures above 44.3 GPa.

Fig. 3.13 shows M-R curves tracing compositions of pure-iron, Earth-like, pure-water and a planet with 95 per cent water and 5 per cent H-He atmosphere subjected to a stellar radiation of $F/F_{\oplus} = 50$ (comparable to the case of the TOI-431 planets), and exoplanets with accurate and reliable mass and radius determinations. It should be noted that the position of the water line in the diagram is very sensitive to used EOS (e.g. Haldemann et al., 2020). Fig. 3.13 shows two water lines using QEOS and EOS from Sotin et al. (2007). As shown in Fig. 3.13, TOI-431 b is one of the many super-Earths following the Earth-like composition line. This suggests that it is mostly made of refractory materials. TOI-431 d, instead, sits above the two pure-water curves and below the 5 per cent curve, implying that the H-He mass fraction is unlikely to exceed a few per cent. Its density is lower than most of the observed sub-Neptunes. There are three planets in the catalogue presented in Otegi et al. (2020a) with masses below $10 M_{\oplus}$ and radii above $3 R_{\oplus}$ (Kepler-11 d,e and Kepler-36 c), and all of their masses have been determined with TTVs. As shown in Otegi et al. (2020b), reducing the uncertainties in this M-R regime would lead to significant improvements on the determination of the volatile envelope mass. As TOI-431 is in the ESPRESSO GTO target

Table 3.7: Inferred interior structure properties of TOI-431 b and d.

Interior Structure:	TOI-431 b	TOI-431 d
$M_{\text{core}}/M_{\text{total}}$	$0.51^{+0.15}_{-0.14}$	$0.29^{+0.16}_{-0.13}$
$M_{\text{mantle}}/M_{\text{total}}$	$0.37^{+0.27}_{-0.18}$	$0.34^{+0.23}_{-0.12}$
$M_{\text{water}}/M_{\text{total}}$	$0.15^{+0.12}_{-0.09}$	$0.33^{+0.21}_{-0.15}$
$M_{\text{H-He}}/M_{\text{total}}$	-	$0.036^{+0.012}_{-0.009}$

list, more observations will help to further constrain the internal structure of TOI-431 d.

We then quantify the degeneracy between the different interior parameters and produce posterior probability distributions using a generalised Bayesian inference analysis with a Nested Sampling scheme (e.g. Buchner, 2014). The interior parameters that are inferred include the masses of the pure-iron core, silicate mantle, water layer and H-He atmospheres. For the analysis, we use the stellar Fe/Si and Mg/Si ratios as a proxy for the planet. Table 3.7 lists the inferred mass fractions of the core, mantle, water-layer and H-He atmosphere from the interior models.

It should be noted, however, that our estimates have rather large uncertainties. Indeed, in this regime of the M-R relation there is a large degeneracy, and therefore the mass ratio between the planetary layers is not well-constrained. Nevertheless, we find that TOI-431 b has a negligible H-He envelope of $1.2 \times 10^{-9} M_{\oplus}$.

The larger companion TOI-431 d is expected to have a significant volatile layer of H-He and/or water of about 3.6 or 33 per cent of its total mass, respectively. The nature of the volatile layer is degenerate.

Considering the future observation prospects of this system, for TOI-431 d we calculate a transmission spectroscopy metric (TSM; Kempton et al., 2018) of 215 ± 58 , after propagating the uncertainties on all system parameters. The relatively large uncertainty is dominated by the uncertainty on the planet’s mass; nonetheless, this TSM value indicates that TOI-431 d is likely among the best transmission spectroscopy targets known among small, cool exoplanets ($< 4R_{\oplus}$, < 1000 K; see Table 11 of Guo et al., 2020).

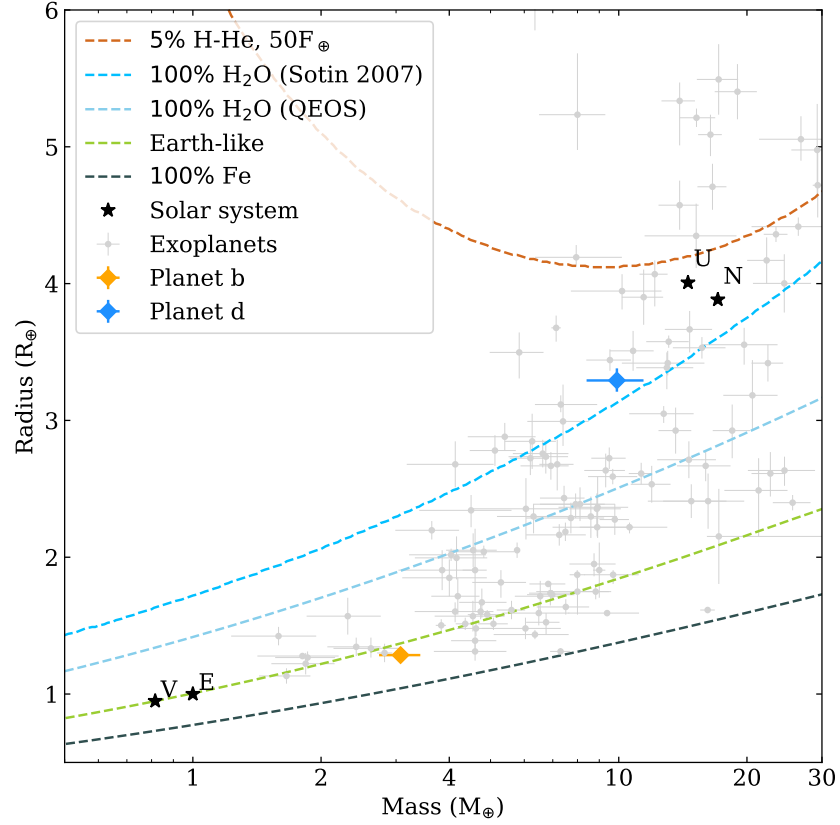


Figure 3.13: Mass-radius diagram of known exoplanets with mass determinations better than 4σ from the NASA exoplanet archive (<https://exoplanetarchive.ipac.caltech.edu>, as of 22 September 2020) shown in grey. TOI-431 b (orange) and d (blue) are denoted as diamonds, and the Solar System planets Venus (V), Earth (E), Uranus (U), and Neptune (N) are marked as black stars. Also shown are the composition lines of iron (dark grey), Earth-like (green), and pure-water planets (pale blue and mid blue, using QEOS and EOS from Sotin et al. (2007) respectively), plus an additional line representing a planet with a 95 per cent water and a 5 per cent H-He envelope with $F/F_{\oplus}=50$, comparable to the case of the TOI-431 planets (brown).

3.5 Conclusion

We have presented here the discovery of three new planets from the *TESS* mission in the TOI-431 system. Our analysis is based upon 2-min cadence *TESS* observations from 2 sectors, ground-based follow-up from LCOGT and NGTS, and space-based follow-up from *Spitzer*. The photometric data was modelled jointly with RV data from the HARPS spectrograph, and further RVs from iSHELL, FEROS, and MINERVA-Australis are included in our analysis. We find evidence to suggest that the host star is rotating with a period of 30.5 days, and account for this in our joint-fit model. Nearby contaminating stellar companions are ruled out by multiple sources of high resolution imaging.

TOI-431 b is a super-Earth characterised by both photometry and RVs, with an ultra-short period of 0.49 days. It likely has a negligible envelope due to substantial atmosphere evolution via photoevaporation, and an Earth-like composition.

TOI-431 c is found in the HARPS RV data and is not seen to transit. It has a period of 4.84 days and a minimum mass similar to the mass of TOI-431 b; extrapolating this minimum mass to a radius via the MR relation places it as a likely second super-Earth.

TOI-431 d is a sub-Neptune with a period of 12.46 days, characterised by both photometry and RVs. It has likely retained a substantial H-He envelope of about 4 per cent of its total mass. Additionally, TOI-431 b and d contribute to the *TESS* Level-1 mission requirement.

This system is a candidate for further study of planetary evolution, with TOI-431 b and d either side of the radius valley. The system is bright, making it amenable to follow-up observations. TOI-431 b, in particular, would potentially be an interesting target for phase-curve observations with JWST.

Chapter 4

Nomads: uncovering the origin of remnant planets in the hot Neptunian desert

Declaration and data availability

The Nomads large HARPS programme (ID 1108.C-0697) is headed by PI David Armstrong, and I am co-I. This chapter has not been published, though there is a summary of the work performed in “TOI-908: a planet at the edge of the Neptune desert transiting a G-type star” (Hawthorn et al., 2023b) and “TOI-2498 b: a hot bloated super-Neptune within the Neptune desert” (Frame et al., 2023), and work in preparation performed by co-authors. Aside from those works, to which I contributed the observation data and assistance towards characterisation of the planetary systems, and the submission of the initial Nomads proposal by David Armstrong, the work presented in this Chapter was wholly performed and written by myself. This includes the sample compilation, observation scheduling, first-look analysis of targets, and preliminary analysis of the sample.

The HARPS RV data products are available from the ESO archive, at http://archive.eso.org/wdb/wdb/adp/phase3_main/form, but any data taken within the last year are proprietary and not publicly available. I do not make this data available, as it is to be formally published elsewhere. TOI data is available from MIT at <https://tess.mit.edu/toi-releases/>. *TESS* data products are available from the Mikulski Archive for Space Telescopes (MAST), at <https://heasarc.gsfc.nasa.gov/docs/tess/data-access.html>.

4.1 Abstract

The Neptunian desert, a region in the planetary distribution at short orbital periods where Neptunian planets are scarce, is one of a few exoplanetary population features that provides valuable evidence towards elucidating the formation, migration, and evolution mechanisms of planetary systems. Here, I present the first steps towards analysing planets within the Neptunian desert in a homogeneous and statistically significant way. To do this, I compile a sample of *TESS* Objects of Interest within the desert, quantifying the biases and using ranking criteria. The sample comprises 47 targets that have already been followed up, and 26 targets that have not. I subsequently observe these in the Nomads large HARPS programme over two years in order to confirm their planetary nature and determine precise mass values. Preliminary analysis is performed on the observed targets, looking at their CCFs, stellar activity, and performing simple fits on their RV data. From this, the planetary nature (or not) of the target is determined. 3 targets have been published as planetary, a further 4 are in-prep, and another 1 has a confirmed planetary signal, where all of these have precise mass determinations. 3 further targets are possibly planetary, but need dedicated stellar activity detrending to pull the signal out and determine the mass precision. 13 targets are null results, and 2 are confirmed false positives (spectroscopic double-lined binaries). Finally, I examine the published planets within the full sample in the context of the Neptunian desert and planetary composition.

4.2 Introduction

The population of exoplanets bears fingerprints of the formation, migration and evolution mechanisms that sculpt planetary systems (Section 1.4). Through observations of many systems (while quantifying the biases of such observations), we can constrain the plausible scenarios. One such region of interest is the Neptunian desert, a marked lack of Neptunian planets with short orbital periods (Section 1.4.1). Since its discovery (Szabó and Kiss, 2011), the desert has been the subject of much study (e.g., Beugé and Nesvorný, 2013; Kurokawa and Nakamoto, 2014; Lundkvist et al., 2016; Matsakos and Königl, 2016; Mazeh et al., 2016; Owen and Lai, 2018; Ionov et al., 2018; Vissapragada et al., 2022; Szabó et al., 2023), and at present, the boundaries of the desert are thought to be set by a combination of photoevaporation (lower boundary) and tidal disruption (upper boundary) (Owen and Lai, 2018).

The Transiting Exoplanet Survey Satellite, *TESS*, started an all-sky survey to find transiting exoplanets in 2018. Prior to *TESS*, few “nomad” exoplanets had been found within the desert, and none far from its boundaries in what we call the “deep desert”. *TESS* soon

discovered several nomad planets, which, at the time of programme proposal (April 2021), included TOI-849 b (Armstrong et al., 2020), the core of what could have been a gas giant in another life, and LTT-9779 b (Jenkins et al., 2020), an ultra-hot planet that has somehow managed to retain a large H-He envelope. These discoveries pose the new question: what pathways can leave planets within the desert?

The nature of TESS SG4 (precise radial velocity) follow-up requires teams to bid for telescope time for planets they are interested in; there is no unified or overarching strategy other than to be collaborative, for example, informing others of the targets on which observations are intended to be made, in order to minimise overlap. Thus, there are many TOIs that lie within the desert that have varying quality of (or no) radial velocity follow-up. Precise radii and masses are required to calculate the bulk density of a planet, which, when combined with the stellar spectra that allow us to determine a star’s chemical composition, can be used to model the interior structure of the planet (e.g., Dorn et al., 2015, 2017). Not having these quantities makes an overarching study of the desert difficult, given the purported mechanisms that sculpt it – without knowing whether or not there is a H-He envelope, how do you investigate photoevaporation histories?

Thus, the “Nomads” programme (ID 1108.C-0697, PI: Armstrong) was born. Its overarching goal is to precisely characterise approximately 30 nomad planets in order to substantially increase the number of planets with precisely measured radii and masses (with errors on the latter determined to better than 20 per cent) within the Neptunian desert, particularly in the deep desert. Follow-up of bright TESS Objects of Interest (TOIs) within the Mazeh et al. (2016) desert boundaries is performed within a large HARPS programme in order to enable investigation of the desert planets in a homogeneous and statistically significant way.

4.3 The sample

Investigating a planetary population in a homogeneous and statistically significant way requires a carefully curated sample that acknowledges and tries to reduce or account for its biases. We choose to base our sample on the list of released TOIs (see Fig. 4.1). As explained in Section 2.1.1, the TOI list itself unfortunately has inherent biases due to the way the Candidate Target List is built and which stars are prioritised for 2 m cadence observations from which TOIs are alerted. In order to compile an unbiased sample set, one would have to do their own transit search on the Full Frame Images (see e.g., Bryant et al., 2023). However, this was not possible for us in the timeframe of the awarded Nomads observation programme, and so would be a worthwhile future project.

The TOI list is collaboratively worked on by members of each Sub Group in a shared

Disposition	Name	Description
SB1	Spectroscopic Binary 1	Single-lined spectra showing an in-phase RV variation too large to be caused by a planet
SB2	Spectroscopic Binary 2	Double-lined spectra in-phase with the photometric orbit
SEB	Spectroscopic data imply Eclipsing Binary	Spectroscopic data imply an EB, i.e. all that do not fall into SEB1 or SEB2)
SEB1	Spectroscopic Eclipsing Binary 1 <i>with orbital solution</i>	Single-lined spectra with an orbital solution that shows period and epoch that match the transit ephemeris
SEB2	Spectroscopic Eclipsing Binary 2 <i>with orbital solution</i>	Double-lined spectra with an orbital solution that shows period and epoch that match the transit ephemeris

Table 4.1: Select spectroscopic disposition tags. Reproduced from the SG2/SG4 Observation Coordinator spreadsheet, which is not publicly accessible.

spreadsheet (called the “SG2/SG4 Observation Coordinator”). As well as presenting the stellar and planetary parameters from the TOI release (Jenkins et al., 2016; Guerrero et al., 2021), it contains a large amount of commentary which is very useful for vetting targets in our Nomads sample.

Each TOI is allocated an “SG1 Disposition”, which is updated as SG1 follow up is performed (though there may be some lag, so the disposition needs to be cross-checked). Dispositions are listed with explanations in Table 4.2.

Notes on each TOI are presented, which generally provide information on: whether the target is already known prior to *TESS* (and other identifiers); whether the *TESS* data is suspicious (e.g., if the transit has an odd shape, whether there are centroid shifts); whether there are any close companions; the status and results of SG1 follow up (e.g., whether the event is confirmed on target, and whether the field is cleared of Nearby Eclipsing Binaries (NEBs)); status and results of any other SG follow-up; and other comments that might be useful to anyone following up the target. There are also more dispositions for spectroscopic follow up that are used in these notes, and an explanation of relevant ones are presented in Table 4.1.

There are also specific columns for SG2 and SG4 follow up. The SG2 (reconnaissance, low-resolution spectroscopy) column will note how many exposures a facility has made and relevant conclusions from the data, e.g., the calculated $v \sin i$ of the star. The SG4 (precise radial velocity follow up) columns will note which PIs/teams/facilities are interested in follow up, whether they have data and how much, and any relevant conclusions from the data, e.g., whether there is drift, stellar activity, a flatline, and/or a conclusive and publishable mass determination made.

There is also a calculation of the predicted mass of each TOI (following Chen and Kipping (2017)), and a corresponding estimate of the radial velocity semi-amplitude. These values must be taken with a grain of salt, but are useful indicators of what to expect from your RV data.

I use all of this information to create our Nomads sample as outlined below.

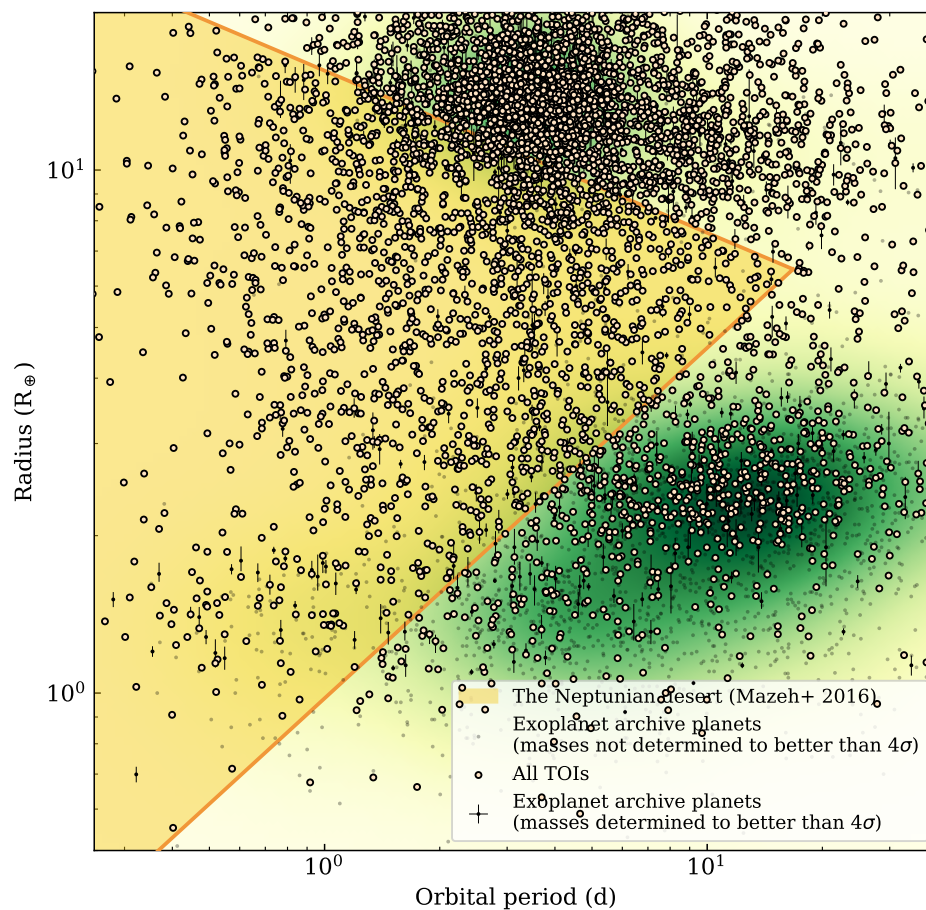


Figure 4.1: All TOIs within and surrounding the Neptunian desert as of 20 July 2022, with periods and radii from their TOI release data. The TOI catalogue is publicly available from MIT via <https://tess.mit.edu/toi-releases/>.

Disposition	Name	Description
P	Confirmed Planet	A confirmed planet with a mass measurement.
VP	Validated Planet	A statistically validated planet that does not have a mass measurement.
KP	Known Planet	A candidate that was known to be a published confirmed/validated planet prior to its <i>TESS</i> detection.
BD	Brown Dwarf	A confirmed brown dwarf with a mass measurement.
VPC+(?)(-)	Verified Planet Candidate, <i>achromatic</i>	Event has been verified by SG1 to: occur within target star follow up aperture; have no strong filter dependent depth chromaticity (i.e. it is achromatic); have no other obvious <i>Gaia</i> DR2 or TIC stars contaminating the aperture that are bright enough to cause the detection. -+ = event appears achromatic, but a known star contaminates the follow up aperture.
VPC(?)	Verified Planet Candidate	Event has been verified by SG1 to: occur within target star follow up aperture; have no other obvious <i>Gaia</i> DR2 or TIC stars contaminating the aperture that are bright enough to cause the detection.
VPC-(?)	Verified Planet Candidate, <i>aperture contaminated</i>	Event has been verified by SG1 to: occur within target star follow up aperture. However, there are other obvious <i>Gaia</i> DR2 or TIC stars contaminating the aperture that are bright enough to cause the detection.
CPC(?)	Cleared Planet Candidate	Follow up has ruled out NEBs in all nearby (within ~ 2.5 arcmin) <i>Gaia</i> DR2 or TIC stars contaminating the aperture that are bright enough to cause the detection.
CPC-	Cleared Planet Candidate, <i>outside target PSF</i>	Follow up has ruled out NEBs in all nearby (within ~ 2.5 arcmin) <i>Gaia</i> DR2 or TIC stars <i>outside the target star PSF</i> that are bright enough to cause the detection. Neighbouring stars within the target star PSF could still be the source of the detection.
PPC	Promising Planet Candidate	Follow up has ruled out <i>most</i> NEBs in all nearby (within ~ 2.5 arcmin) <i>Gaia</i> DR2 or TIC stars that are bright enough to cause the detection, and there are no obvious NEBs in the uncleared stars.
PC	Planet Candidate	A planet candidate with no or so far inconclusive follow up observations.
LEPC	Lost Ephemeris Planet Candidate	Photometric follow up has not detected the event on target or in a nearby neighbour when it should have, within a several sigma ephemeris coverage.
STPC	Single Transit Planet Candidate	A planet candidate from a single (“mono”) transit detection in the <i>TESS</i> data.
PNEB(?)	Possible Nearby Eclipsing Binary	Follow up has detected a nearby EB event in one epoch that could be the source of the detection. A second epoch is required to confirm or deny the NEB.
NEB	Nearby Eclipsing Binary	Follow up has determined that the detection was caused by a nearby EB that is contaminating the <i>TESS</i> aperture.
NPC(?)	Nearby Planet Candidate	Event was on a nearby star, but is still consistent with a planet around that nearby star. The nearby star is assigned a new TOI number, and the NPC TOI is retired.
APC	Amibiguous Planet Candidate	Initial follow up is suggestive of a false positive, or that confirmation of a planet would be difficult.
BEB(?)	Blended Eclipsing Binary	Detection likely caused by a blended eclipsing binary in the <i>TESS</i> and follow up apertures. They are generally determined as such via: a strongly chromatic transit depth in follow up photometry; a lack of detectable RV variation when the estimated planet mass should have been detected; and/or bisector span variations that are correlated with the RVs.
EB(?)	Eclipsing Binary	Event apparently occurring on target, but too deep relative to the stellar radius to be caused by a planet. Also used for systems that follow up observations, vetting diagnostic reports and/or other published data determines have: a clear odd-even depth difference; synchronised out-of-eclipse phase variations; a secondary event too deep for a planet; unequal times between odd/even events; and/or a host star so big that the transit depth cannot be caused by a planet.
FA	False Alarm	TOI signal has been found to be due to instrumental noise or a systematic.

Table 4.2: All SG1 Disposition tags. Where ? is added, it implies a tentative detection or inconsistencies between multiple observations. Reproduced from the SG2/SG4 Observation Coordinator spreadsheet, which is not publicly accessible.

4.3.1 Stage 1: cuts based on desert boundaries and observability

It is first necessary to make basic cuts to the full TOI list based upon the boundaries of the Neptune desert and objects that HARPS can effectively observe.

These cuts are as follows:

- **Remove all TOIs with stellar $V_{\text{mag}} > 13$.** Bright stars require less exposure time to reach an acceptable signal-to-noise and RV precision. Removing faint stars therefore maximises the time the program can devote to as many targets as possible. Also, many faint stars will never allow us reach the precision required to detect a Neptunian-sized planet in the RV data.
- **Remove all TOIs with stellar $\text{Dec} \geq +20$.** As HARPS is located in the southern hemisphere, targets in the northern sky are impossible to observe (see Section 2.2.2).
- **Remove all TOIs with stellar $R_{\star} \geq 1.5R_{\odot}$.** This removes larger A-type stars where the semi-amplitude of a Neptunian-sized planet would become too small to detect with HARPS.
- **Remove all TOIs with stellar $T_{\text{eff}} \geq 6400 \text{ K}$.** Again, this is to remove larger stars, as radius scales with effective temperature. Additionally, hot stars rotate faster, which broadens the spectral lines and makes it difficult to obtain precise radial velocities.
- **Remove all TOIs with SG1 dispositions explicitly marked as “FA” or “EB”.** As explained above, these are known False Alarms and Eclipsing Binaries, and are therefore not planets.
- **Remove all TOIs with planetary P and R_p outside of the Mazeh et al. (2016) boundaries.** While these boundaries are empirical and perhaps outdated (and thus perhaps can be labelled as generous in what they consider as a desert planet), there is not yet any newer, widely accepted and used definition of the boundaries of the desert in the literature.
- **Remove all TOIs with planetary $P \geq 10 \text{ days}$.** This cuts off the very tip of the Mazeh et al. (2016) desert triangle. Planets with periods longer than 10 days are generally considered “warm”, rather than “hot”, and also take more time to follow up.

This results in a preliminary sample of 272 planets (shown in Fig. 4.2), which is far too many for a single large programme to follow up (as stated earlier, the goal for Nomads is to follow up ~ 30 targets). However, a lot of these are still poor candidates for follow up for many reasons, and manual vetting is now required to remove these and refine the sample further.

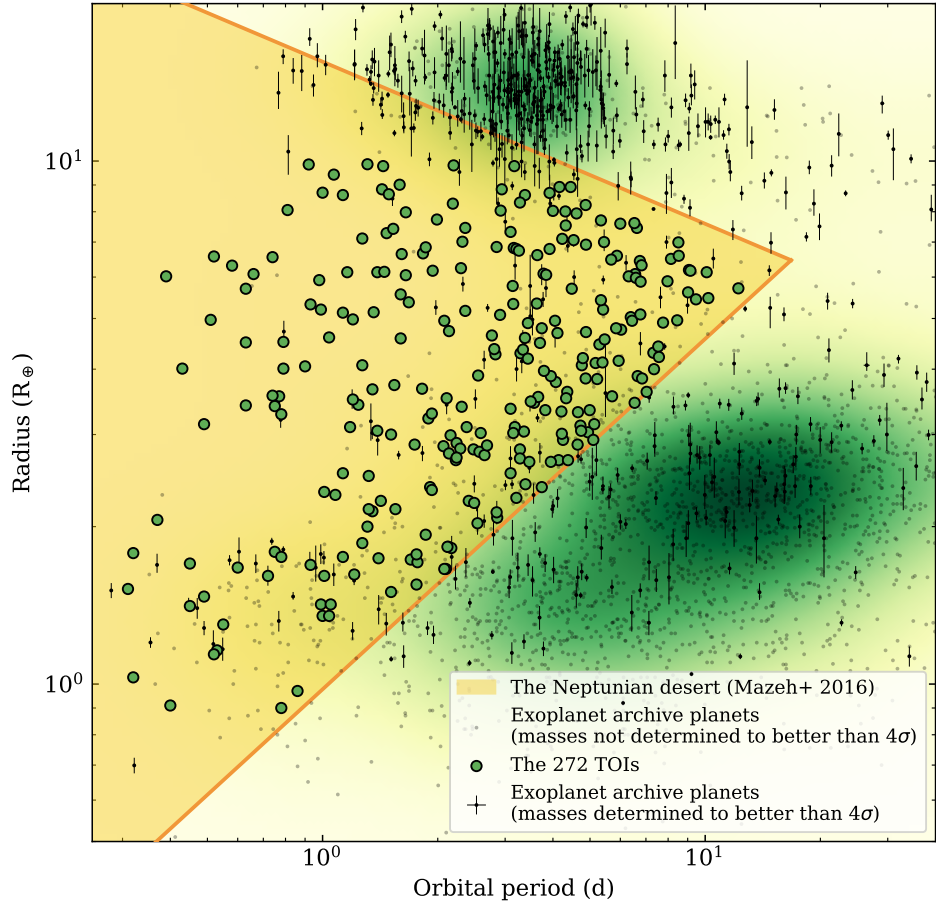


Figure 4.2: The 272 remaining TOIs after basic cuts are performed based on observability and the Neptunian desert boundaries as described in Section 4.3.1, with periods and radii taken from their TOI release data. These 272 targets are listed in Table 4.3.

4.3.2 Stage 2: manual vetting and ranking

Manual vetting

Manual vetting is performed based on information available on the SG2/SG4 Observation Coordinator spreadsheet¹, the ExoFOP-TESS pages of each target², the “TESS wiki” on Confluence³, and the NASA Exoplanet Archive⁴. For each target, I consider the following questions to determine whether or not it is eligible for the programme. My conclusions for each target, summarised in brief, are presented in the final column labelled “Notes (brief)” of Table 4.3.

- **Is the target already published, in preparation to be published, or otherwise currently being followed up by another PI/team?** As *TESS* has been operating for several years already, a lot of high interest targets have already been followed up. I perform a search of the NASA Exoplanet Archive (where confirmed planets are recorded), the *TESS* wiki (where abstracts for in preparation *TESS* papers are published), and the SG2/SG4 Observation Coordinator (where SG4 PIs/teams record their observations/intent to observe). In some cases, I contacted individual PIs to determine whether they’ve got enough RVs to publish a 20 per cent mass error, as I can offer to observe the target on HARPS to make up the shortfall if not. If the answer to the overarching question is yes, these targets are marked as a “0” in the Rank column of Table 4.3, and the publication or PI is noted. These are taken as inclusions to the sample, but I do not have to perform any follow up in the Nomads programme.
- **Do the notes label the candidate as some kind of (likely or confirmed) false positive scenario?** At the opposite end of the scale to published systems, there are plenty of candidates that are definitely (or very likely) not planets. Any candidates that are labelled as any flavour of eclipsing binary by photometric or spectroscopic follow up (see Tables 4.2 and 4.1) are dismissed from the sample. These are marked as “B” in the Rank column of Table 4.3. As well as the notes in the SG2/SG4 spreadsheet, data and associated notes are uploaded to ExoFOP-TESS, so I cross-check there.
- **Is there past SG2/4 follow up that indicates a lack of planet detection?** For some candidates, spectroscopic follow up will note that the target is “flatlining”, implying there is no planetary signal present. Several of these are from the past NCORES programme ((ID 1102.C-0249, PI: Armstrong), the predecessor to the

¹The SG2/SG4 Observation Coordinator spreadsheet is not publicly available.

²ExoFOP-TESS is available at <https://exofop.ipac.caltech.edu/tess/>.

³The “TESS wiki” or tessbook is available at <https://tessbook.mit.edu/display/TESS/Home>.

⁴The NASA Exoplanet Archive is available at <https://exoplanetarchive.ipac.caltech.edu/index.html>.

Nomads programme with many of the same consortium members, so it is easy for me to go back and examine that data. In the case where the data has been taken by another PI, I contact them to confirm the flatline. These are marked as “B” in the Rank column of Table 4.3, but are not automatically dismissed. I still rank these targets, as they would still contribute to the final sample as a null result (and we can obtain an upper mass limit of the planet that may be present) if they made the ranking cuts described below. Again, as well as the notes in the SG2/SG4 spreadsheet, data and associated notes are uploaded to ExoFOP-TESS, so I cross-check there.

- **Does the target have a reported $v \sin i$?** If a star spins rapidly, the lines in its spectra are broadened, making it hard to obtain precise radial velocities from the cross-correlation function. A high rotation rate is also usually indicative of more stellar activity, which also complicates PRV follow up. I dismiss any targets with reported $v \sin i > 6 \text{ km s}^{-1}$ from the sample. These are marked as “B” in the Rank column of Table 4.3, and their reported $v \sin i$ (and the facility this was measured with) is recorded in the Notes column.
- **Is the target a known young star?** Similarly to the above, young stars spin rapidly and also exhibit high levels of stellar activity, making PRV very difficult, and are so dismissed from the sample. These are marked as “B” in the Rank column of Table 4.3, and their known young star status is recorded in the Notes column.
- **Is the target lacking a stellar radius?** Unfortunately, as remarked upon many times now in this thesis, obtaining a planetary radius is predicated on us knowing the stellar radius. It is rare, given the work of *Gaia*, but some stars lack a radius measurement. A radius *can* be obtained from stellar spectra, but it takes many stacked together to obtain a good measurement and thus could be a very large waste of HARPS time. So, I dismiss any targets that do not have a stellar radius measurement from the sample, and these are marked as “B” in the Rank column of Table 4.3, with “no stellar radius” remarked in the Notes column.

75 targets are dismissed from the sample as bad (B in the “Rank” column), and 65 targets could theoretically make the final sample, but have already been followed up and thus do not need to be observed in the Nomads programme. Additionally, 6 targets are known flatlines, so again could theoretically make the final sample but I do not need to observe them. This leaves 126 targets, cutting the number by more than half, but still too many to follow up within the time constraints of a large HARPS programme. Thus, I turn to ranking the targets to prioritise which are followed up.

Ranking

I rank in two ways: one based on a “merit function” that prioritises targets based on visual magnitude and position within the desert, and the other based on how much follow up a target already has. Both are based upon minimising wastage of observing time on HARPS.

The merit function, M , is a multiplication of a two metric values: one for stellar magnitude, M_{mag} , and one for the planet’s “distance” from the desert boundaries in period-radius space, M_{dist} . These are calculated as follows.

First, the metric value based on stellar magnitude:

$$M_{\text{mag}} = |V_{\text{mag}} - 13| + 1. \quad (4.1)$$

This metric value scales linearly with brightness. As we already cut all targets with $V_{\text{mag}} > 13$, the smallest this value can be is 1. For our brightest target, with $V_{\text{mag}} = 7.95$, $M_{\text{mag}} = 6.05$.

Second, to create a metric value that quantifies the distance of a target from the desert boundaries, we first use the equations that define the boundaries in period-radius ($P - R_P$) space as given in Mazeh et al. (2016), where ldb stands for lower desert boundary, and udb stands for upper desert boundary:

$$\begin{aligned} R_{\text{ldb}, \text{Mazeh}} &= 0.67P - 0.01; \\ R_{\text{udb}, \text{Mazeh}} &= -0.33P + 1.17. \end{aligned} \quad (4.2)$$

We modify these slightly:

$$\begin{aligned} R_{\text{ldb}} &= 10^{0.67 \log_{10}(P) - 0.01}; \\ R_{\text{udb}} &= \min\{10, 10^{-0.33 \log_{10}(P) + 1.17}\}. \end{aligned} \quad (4.3)$$

First, this is to transform the boundary out of log space. Then, a minimum clause is added to the upper boundary equation. At very short periods, the Mazeh et al. (2016) boundary extends to very large radii, which, if unchanged, would heavily bias the distance metric to ultra-short period planets. The minima modifies the boundary to not extend to radii greater than 10 Earth radii, effectively flattening it.

Then, the minimum value produced by either equation is taken as the value for the distance metric where:

$$M_{\text{dist}} = \min\{|R_P - R_{\text{ldb}}|, |R_P - R_{\text{udb}}|\} \quad (4.4)$$

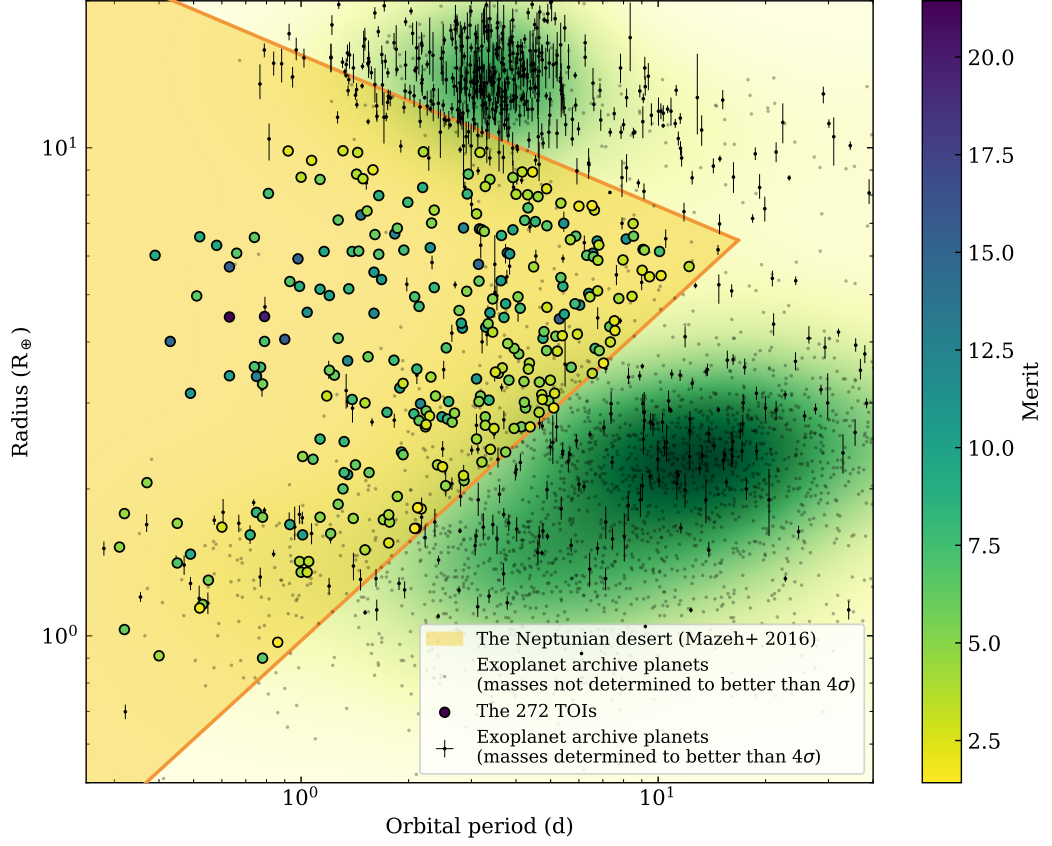


Figure 4.3: The 272 TOIs, this time showing a colour corresponding to their “Merit” in Table 4.3, explained in Section 4.3.2.

The final merit function is calculated as a simple multiplication of the two merit values as follows:

$$M = M_{\text{mag}} \times M_{\text{dist}} \quad (4.5)$$

As mentioned in Section 2.2.2, a brighter target requires a shorter exposure time, and thus uses less observing time to obtain a conclusive result. We prioritise candidates deeper within the desert due to their high importance when it comes to studying the formation and evolution mechanisms sculpting the desert - they present a significant science return, and fit the goals of the Nomads programme.

The ranking based on follow up is perhaps a little subjective. Targets from manual vetting that have been published, are in-prep, or are being followed up already have a rank of 0, and false positives a rank of B. Each remaining target is rated from 1-4 based on the photometric (SG1), spectroscopic (SG2), and high-resolution (SG3) follow up it already

has. A ranking of 1 means it is a good target to observe; it is confirmed on target and has SG1, SG2 and SG3 follow up. A ranking of 2 is similar; however, it may not be fully confirmed on target or the field entirely cleared of NEBs, but the outlook is positive and it is unlikely to be a false positive. A ranking of 3 implies some amount of follow up by any sub group, but there is no confirmation on target or NEB clearing, so it could be a false positive. A ranking of 4 means there is no follow up on the target at all, and this poses the most risk for time wasted on following up false positives.

After ranking each target via both systems, with the merit function recorded in the Merit column of Table 4.3 and the ranking based on follow up recorded in the “Rank” column, we make a cut on both to end up with a sample that contains close to 30 targets that will be observed by our programme on HARPS. This is a cut based on retaining all follow up rankings from 0-3 (removing ranks of 4 and B, excepting B rankings that come from known flatlines) and then only retaining targets with a merit function result of ≥ 5 . This is shown in Fig. 4.4.

It is important to note here that this sample was updated multiple times throughout the programme running - the TOI list is ever-changing with new releases and status updates on targets. Additionally, the ranking system was in flux before being finalised. The sample presented in this Chapter is the final version compiled on 20 July 2022. Thus, there are notes in Table 4.2 indicating that some targets which did not make the final cut were actually observed earlier in the programme (i.e. “observed as part of older version of sample”) before the ranking system was finalised or when their ranking may have been more favourable. I do not discuss these targets further in this thesis, but the data will still be analysed and published as an outcome of the programme.

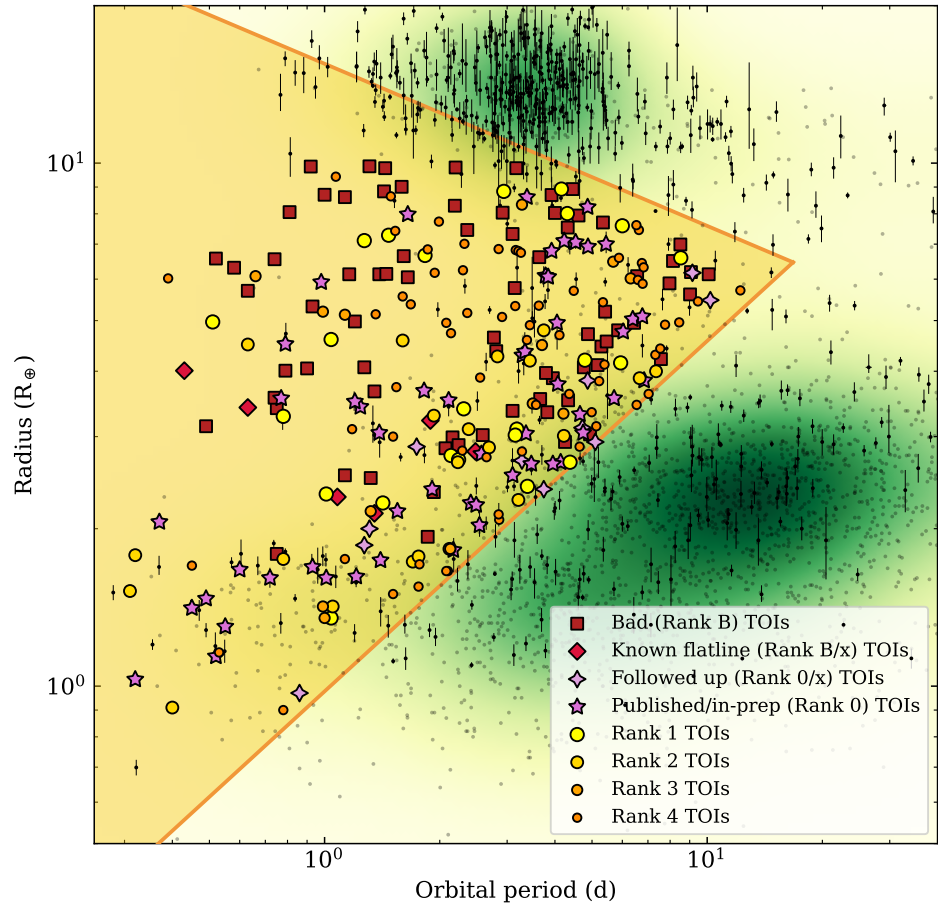


Figure 4.4: The 272 TOIs, this time showing a colour and symbol corresponding to their “Rank” in Table 4.3, explained in Section 4.3.2.

4.3.3 Stage 3: the final sample

The final sample has two forms.

First, the “full” sample. This is the cut explained above: retaining all follow up rankings from 0-3 (removing ranks of 4 and B, excepting B rankings that come from known flatlines) and then only retaining targets with a merit function result of ≥ 5 . It is the homogeneous sample where an overall analysis of planets within it should inform our theories on formation and evolution processes that sculpt the desert. Thus it deliberately includes those that are already published/followed up and those that are known flatlines. This is displayed in Fig. 4.5. Targets that are already published/followed up and known flatlines are also presented in Table 4.4, where the former are noted with the publication values of their periods, radii, and masses. There are 6 known flatlines, 29 published, and 12 in-prep/being followed up, for a total of 47 targets that contribute to the full sample but which the Nomads programme does not need to observe.

From the full sample, I pull an “observation” sub-sample. These are the targets that have not been followed up and do not have mass determinations with errors to better than 20 per cent. Therefore, these are the targets that the Nomads programme will observe to complete the full sample. This sub-sample is shown in Fig. 4.6, and also presented in Table 4.5. There are a total of 26 targets to be observed.

The full sample is an addition of the targets in Tables 4.4 and 4.5.

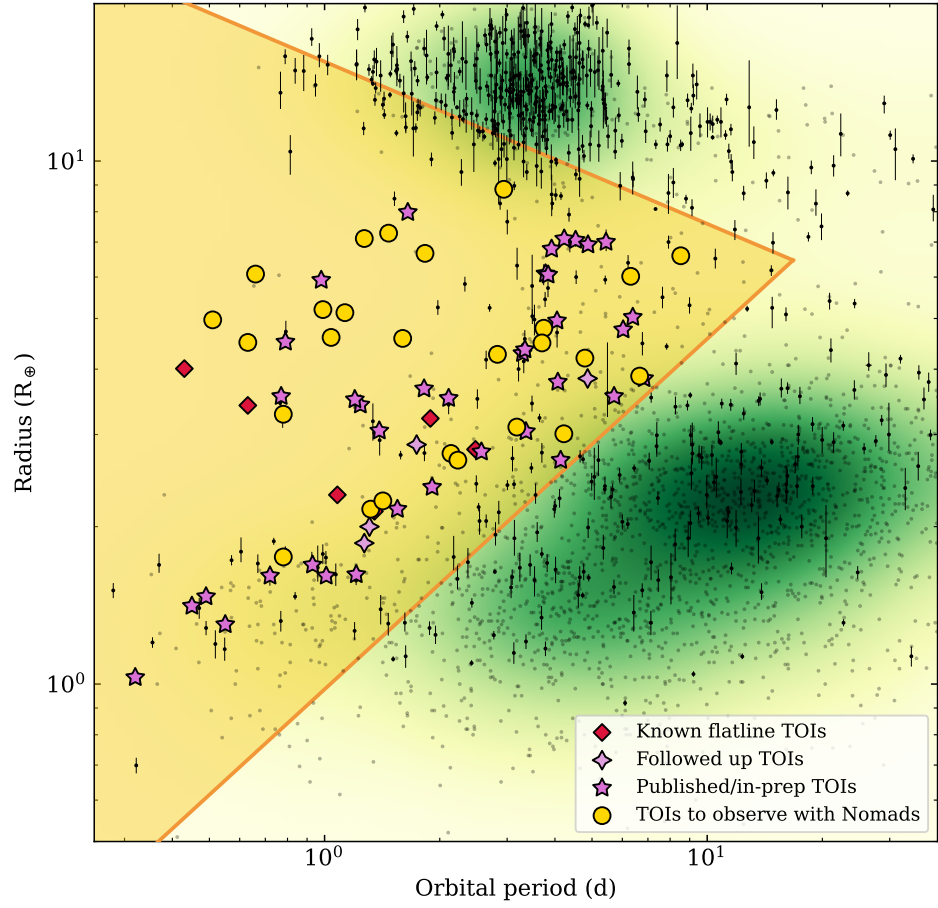


Figure 4.5: The final full sample, explained in Section 4.3.3 and displayed in Tables 4.4 and 4.5. The colours and symbols correspond to their “Rank” in Table 4.3, explained in Section 4.3.2.

Table 4.3: The 272 targets from the TOI list (downloaded on 20 July 2022) after simple cuts on brightness and observability, as described in Section 4.3.1. The Merit and Rank columns are described in Section 4.3.2. I colour code the Rank column: targets with a bad (B) rank are dark red; targets with a known null result (B/x) are bright red; targets that are published or in prep with ≤ 20 per cent mass errors (0) are dark purple; targets that are currently being followed up (0/x) are light purple; and targets with a rank of 1-4 (based on amount of follow up) are shades of yellow, from light yellow to dark yellow respectively. The TOI column is also colour coded: published/in-prep/followed up targets in the final sample are purple, known nulls in the final sample are red, and targets that make up the final observation subsample are yellow (see Section 4.5). All of these are then taken forward into Tables 4.4 and 4.5. The dividing line marks the cut-off at a Merit of 5.

Target		Nomads		SG1 status and TOI release values							Notes (brief)
TOI	TIC	Merit	Rank	SG1 Disp.	P (days)	R _P (R _⊙)	M _P (M _⊙)	K (m s ⁻¹)	V _{mag}	N _{pl}	
422.01	117979455	21.44	2	CPC-	0.63	4.50	18.49	12.87	9.51	1	⇒ final sample; close companion causing ambiguity
193.01	183985250	19.72	0	P	0.79	4.51	18.60	13.49	9.78	1	Published (Jenkins et al., 2020)
2226.01	403135192	15.80	B	BEB?	0.90	4.05	15.50	9.99	10.19	1	Likely BEB
4391.01	66620917	15.67	B	PC	5.29	4.46	18.21	6.68	7.66	1	CHIRON $v \sin i = 40$; no stellar radius
1047.01	370745311	15.39	B	NPC	0.63	5.70	27.69	20.69	11.10	1	NPC
129.01	201248411	15.25	0	P	0.98	5.92	29.49	24.06	11.00	1	Published (Nielsen et al., 2020b)
2431.01	258804746	14.82	0	VPC	0.22	4.14	16.06	22.75	10.90	1	In prep (Malavolta, TESS wiki)
4605.01	19155785	14.82	B	PNEB	3.14	5.77	28.22	12.21	10.82	1	PNEB
1967.01	320079492	14.50	B/2	VPC?	0.43	4.01	15.22	12.85	10.75	1	Observed (Jenkins, priv. comm.), null result
1036.01	146172354	13.47	0	VPC	3.78	6.09	30.94	12.45	10.97	1	In prep (Nielsen, priv. comm.)
570.01	126733133	13.39	1	VPC	1.47	7.28	41.99	20.14	10.40	1	⇒ final sample
2671.01	190081497	13.26	4	PC	3.16	6.82	37.53	14.64	10.83	1	Centroid offset possibly onto another star; possibly EB
865.01	44797824	12.39	B	APC	0.75	3.40	11.51	8.40	10.56	1	SB2
2665.01	186936449	12.38	4	PC	2.85	6.85	37.85	15.55	11.02	1	-
855.01	269558487	12.22	1	VPC	1.83	6.66	36.08	15.00	11.18	1	⇒ final sample; potential false positive from ShARCS AO (Dressing in prep)
4175.01	952046774	12.19	B	PC	2.16	2.99	9.25	4.94	8.82	1	No stellar information
728.01	96097215	12.05	2	VPC?	1.60	4.58	19.05	10.21	11.15	1	⇒ final sample
2201.01	219306934	11.76	B	VPC?	1.35	3.66	13.01	6.63	10.60	1	CHIRON $v \sin i = 7.5$
5291.01	250983039	11.52	4	PC	2.30	6.24	32.31	15.75	11.58	1	-
355.01	183593642	11.49	1	VPC	1.04	4.60	19.26	11.21	11.50	1	⇒ final sample; possible odd-even
576.01	408310006	11.40	0	P	5.44	7.00	39.28	13.07	9.35	1	Published (Hellier et al., 2019b)

Continued on the next page...

Target		Nomads		SG1 status and TOI release values							Notes (brief)
TOI	TIC	Merit	Rank	SG1 Disp.	P (days)	R _P (R _⊙)	M _P (M _⊙)	K (m s ⁻¹)	V _{mag}	N _{pl}	
2539.01	201508515	8.92	3	PC	0.99	5.20	23.69	15.04	12.30	1	⇒ final sample
189.01	278866211	8.83	B	APC	2.19	8.29	52.34	22.68	10.74	1	APC/BEB
564.01	1003831	8.78	0	P	1.65	7.98	48.99	26.91	11.10	1	Published (Davis et al., 2020)
2243.01	1990842033	8.71	B	APC	5.41	5.20	23.66	7.31	11.25	1	APC/binary; two stars in one pixel; stellar radius unknown
4461.01	149282072	8.65	3	PC	3.70	4.49	18.48	8.61	11.25	1	⇒ final sample
2673.01	1018843	8.50	0	CPC-	1.91	2.38	6.25	3.24	9.45	1	In prep (Serrano, TESS wiki)
2922.01	45492609	8.41	4	PC	1.36	5.14	23.23	12.73	12.30	1	-
1972.01	267414551	8.40	B	PNEB	4.76	4.07	15.64	5.29	10.34	1	PNEB
3450.01	126125073	8.39	B	PC	3.64	6.61	35.58	15.09	11.93	1	Two very close (less than 1 arcsec) stars
1943.01	382980571	8.35	0/3	VPC	1.74	2.87	8.64	4.50	10.60	1	Under observation (Quinn/Gandolfi, 20/22 on SG4 coordinator respectively)
4541.01	293432942	8.33	B	APC	1.13	2.53	6.98	4.28	10.63	1	SB; TRES $v \sin i = 10$
5556.01	55315929	8.31	0	KP	0.72	1.61	3.23	2.20	9.45	1	Published (Frustagli et al., 2020)
1060.01	101230735	8.25	B	APC	2.07	2.85	8.49	4.23	10.34	1	HARPS variation of $\sim \text{km s}^{-1}$ in a few nights; very close companion at less than 1 arcsec
2996.01	35855047	8.24	4	PC	2.36	5.17	23.47	10.70	12.14	1	-
735.01	150271680	8.22	B	SEB2	2.59	3.02	9.42	4.64	10.22	1	SEB2
181.01	76923707	8.21	0	VPC+	4.53	7.06	39.81	18.47	11.19	1	Published (Mistry et al., 2023)
261.01	63898957	8.18	0	PC	3.36	3.04	9.53	3.89	9.56	2	In prep (Gandolfi, priv. comm.)
3505.01	390988385	8.17	B	PC	2.92	8.04	49.62	19.67	11.24	1	TRES $v \sin i = 12$
132.01	89020549	8.15	0	P	2.11	3.52	12.18	6.47	11.20	1	Published (Díaz et al., 2020)
2590.01	234832821	8.12	B	PC	0.75	1.79	3.85	2.44	9.90	1	TRES $v \sin i = 11$; CHIRON $v \sin i = 10$; possible blend
949.01	189013222	8.08	B	NPC	1.32	2.50	6.82	4.27	10.52	1	Event on a different star
2688.01	98728690	7.87	4	PC	2.14	4.73	20.16	10.20	12.08	1	-
854.01	160222069	7.85	B	APC	5.85	4.79	20.60	6.94	10.98	1	NEB
2958.01	463203560	7.82	4	PC	1.99	7.73	46.48	20.92	11.61	1	-
426.01	189013224	7.71	3	PPC	1.32	2.16	5.32	3.06	10.12	1	⇒ final sample; ambiguity between two stars
5522.01	366409685	7.66	B	PC	1.39	6.13	31.36	15.67	12.43	1	TRES $v \sin i = 10$; possible secondary
824.01	193641523	7.60	0	P	1.39	3.05	9.56	6.94	11.31	1	Published (Burt et al., 2020)

Continued on the next page...

Target		Nomads		SG1 status and TOI release values							Notes (brief)
TOI	TIC	Merit	Rank	SG1 Disp.	P (days)	R _P (R _⊕)	M _P (M _⊕)	K (m s ⁻¹)	V _{mag}	N _{pl}	
880.02	34077285	7.55	0	PC	2.57	2.78	8.16	4.38	10.10	3	In prep (Nielsen, priv. comm.)
969.01	280437559	7.55	0	VPC+	1.82	3.67	13.08	9.03	11.65	1	Published (Lillo-Box et al., 2023)
849.01	33595516	7.54	0	P	0.77	3.55	12.39	9.19	11.98	1	Published (Armstrong et al., 2020)
462.01	420049884	7.44	4	VPC	4.11	4.69	19.89	7.30	11.66	1	Stellar companion or hierarchical triple
4230.01	128127712	7.40	4	PC	1.94	6.17	31.66	16.22	12.47	1	-
4537.01	251039147	7.37	2	VPC?	6.66	3.88	14.42	4.60	8.74	1	⇒ final sample
235.01	280095254	7.32	B	VPC+	10.09	6.13	31.30	8.58	9.86	1	Very close companion (less than 1 arcsec) causing significant line asymmetry
293.01	355637190	7.31	B	APC	0.81	8.06	49.86	31.96	11.51	1	APC/BEB
940.01	248434716	7.12	B	APC	0.74	3.56	12.42	9.78	12.11	1	APC/BEB
4179.01	297148017	7.10	4	PC	2.55	3.89	14.43	6.75	11.68	1	Possible depth-aperture correlation
5494.01	14440334	7.08	4	PC	2.08	4.95	21.80	11.43	12.38	1	-
2744.01	279989567	7.07	B	PNEB	0.79	4.01	15.20	10.52	12.31	1	PNEB; no stellar radius
4800.01	280307604	6.95	B	PC	4.32	7.53	44.39	14.93	11.33	1	TRES $v \sin i = 12$; possibly on different star; ShARCS AO suggest false positive (Dressing in prep)
5632.01	420779000	6.93	0	KP	4.23	7.10	40.17	17.75	11.76	1	Published (Hartman et al., 2011)
561.02	377064495	6.91	0	P	0.45	1.41	2.59	2.24	10.25	3	Published (Weiss et al., 2021)
2653.01	149394317	6.81	B	PC	0.93	5.32	24.65	15.27	12.74	1	No stellar radius
2340.01	29959761	6.78	B	NPC	1.65	6.05	30.60	18.66	12.63	1	NPC; event not on this target
2427.01	142937186	6.77	0/1	VP	1.31	2.00	4.65	3.66	10.30	1	Under observation (Gandolfi/Quinn, 101/33 on SG4 coordinator respectively); validated in Giacalone et al. (2022)
745.01	444842193	6.73	B/3	APC	1.08	2.30	5.90	3.51	11.03	1	Known null from HARPS NCOES
2335.01	160003961	6.69	B	APC	1.45	6.14	31.44	16.57	12.62	1	APC/SB2
2365.01	344085117	6.68	3	VPC?	0.66	6.08	30.92	21.98	12.64	1	⇒ final sample; v-shaped; confusion with neighbour
502.01	172193428	6.64	1	VPC	2.94	8.83	58.29	24.30	10.94	1	⇒ final sample
2641.01	162802770	6.56	0	VPC+	4.88	6.92	38.50	12.57	11.69	1	Published (Psaridi et al., 2023)
5005.01	282485660	6.48	3	VPC?	6.31	6.02	30.33	10.16	11.87	1	⇒ final sample
4029.01	393633044	6.48	4	VPC	5.88	6.59	35.45	12.40	11.55	1	ShARCS AO validation suggests false positive (Dressing in prep)
801.01	177258735	6.44	4	PC	0.78	0.90	0.68	0.41	8.04	1	Centroid variation with sector

Continued on the next page...

Target		Nomads		SG1 status and TOI release values							Notes (brief)
TOI	TIC	Merit	Rank	SG1 Disp.	P (days)	R _P (R _⊕)	M _P (M _⊕)	K (m s ⁻¹)	V _{mag}	N _{pl}	
2355.01	124515764	6.40	B	PC	1.27	4.07	15.62	8.03	12.37	1	CHIRON $v \sin i = 8$
3105.01	124454726	6.39	4	PC	5.34	5.42	25.39	8.56	12.13	1	-
273.01	279740441	6.36	B	NPC	0.74	6.55	35.09	32.37	12.57	1	NEB
5088.01	234388232	6.32	4	PC	7.32	4.30	17.13	5.64	10.02	2	Possible centroid offset
3071.01	452006073	6.29	1	VPC-?	1.27	7.11	40.30	20.74	12.38	1	⇒ final sample
2224.01	388198242	6.28	1	VPC	0.51	4.97	21.93	18.88	12.83	1	⇒ final sample; v-shaped, possible odd-even
2725.01	100780304	6.25	4	PC	0.18	4.10	15.83	16.77	12.70	1	-
1839.01	381714186	6.21	1	APC	1.42	2.24	5.66	3.36	10.90	1	⇒ final sample; possible APC flagged after observations taken
5559.02	456945304	6.18	1	PC	2.14	2.76	8.08	4.21	11.11	1	⇒ final sample; K2-370 b; validated in Christiansen et al. (2022)
5018.01	294691204	6.15	4	PC	6.59	5.98	30.04	8.78	11.92	1	-
4599.02	307809773	6.15	0	P	5.71	3.55	12.35	7.38	9.63	2	Published (Luque et al., 2022)
731.01	34068865	6.04	0	P	0.32	1.03	1.07	1.69	10.15	1	Published (Lam et al., 2021)
835.01	405700729	6.04	1	VPC+	4.79	4.20	16.46	6.24	11.49	1	⇒ final sample
1117.01	295541511	6.01	2	CPC	2.23	2.68	7.69	3.72	11.02	1	⇒ final sample
3379.01	130191319	6.01	B	PC	1.61	6.64	35.89	18.06	12.62	1	TRES $v \sin i = 8$
1853.01	73540072	6.00	0	VPC+	1.24	3.42	11.63	7.89	12.18	1	Published (Naponiello et al., 2023)
644.01	63303499	5.99	B	PNEB	1.93	2.35	6.13	2.85	10.73	1	NEB
4245.01	52928939	5.98	4	PC	3.15	5.33	24.67	10.83	12.58	1	-
544.01	50618703	5.96	0	P	1.55	2.16	5.32	3.88	10.78	1	In prep (Gandolfi, priv. comm.)
3268.01	410358456	5.91	4	PC	1.86	6.84	37.76	20.89	12.58	1	-
2486.01	369376388	5.86	4	PPC	1.54	3.73	13.43	9.81	12.29	1	-
3350.01	4959676	5.84	B	PNEB	2.24	2.89	8.75	5.09	11.37	1	PNEB, possible secondary
2595.01	151284882	5.84	B	APC	3.82	3.34	11.14	4.22	10.99	1	APC; TRES $v \sin i = 8$
499.01	123702439	5.81	1	VPC-	8.53	6.59	35.39	9.75	10.59	1	⇒ final sample; close neighbour
2196.01	372172128	5.79	0	VPC+	1.20	3.50	12.10	7.19	12.30	1	Published (Persson et al., 2022)
5484.01	443961200	5.76	B	PC	1.13	8.61	55.75	29.50	11.59	1	TRES $v \sin i = 8$
179.01	207141131	5.74	0	VPC+	4.14	2.68	7.65	3.40	8.99	1	Published (Vines et al., 2023; Desidera et al., 2023)
332.01	139285832	5.70	1	VPC+	0.78	3.28	10.81	8.18	12.35	1	⇒ final sample
500.01	134200185	5.68	0	P	0.55	1.30	2.23	2.21	10.54	1	Published (Serrano et al., 2022)

Continued on the next page...

Target		Nomads		SG1 status and TOI release values							Notes (brief)
TOI	TIC	Merit	Rank	SG1 Disp.	P (days)	R _P (R _⊕)	M _P (M _⊕)	K (m s ⁻¹)	V _{mag}	N _{pl}	
5534.01	176868951	4.79	0	KP	0.37	2.06	4.91	4.82	12.13	1	K2-131 b; published (Dai et al., 2017)
4724.01	119638424	4.77	4	PC	3.57	3.45	11.79	4.86	11.79	1	ShARCS AO suggest false positive (Dressing in prep)
539.01	238004786	4.72	2	CPC-	0.31	1.52	2.94	3.22	11.73	1	Marginal signal; under observation by Veloce; validated in Giacalone et al. (2022)
755.01	73228647	4.72	0	P	2.54	2.03	4.79	2.20	10.09	1	Published (Osborn et al., 2021b)
740.01	310009611	4.71	3	PC	2.13	1.83	4.00	1.86	10.09	1	Possible odd-even
2731.01	34729329	4.69	3	CPC	1.00	1.35	2.39	1.52	10.59	1	-
125.01	52368076	4.68	0	P	4.65	3.31	10.99	4.48	11.02	3	Published (Nielsen et al., 2020a)
3225.01	401952328	4.68	0/4	PC	9.15	6.17	31.66	9.67	11.61	1	Observed by CORALIE (7 on SG4 coordinator)
5448.01	291016451	4.67	4	PC	3.21	6.79	37.30	16.96	12.89	1	-
4640.01	337216118	4.66	2	PC	2.69	2.86	8.58	3.70	11.63	1	Validated by ShARCS AO (Dressing in prep)
333.01	224245334	4.65	B	VPC	3.79	3.97	15.00	5.54	12.21	1	CHIRON $v \sin i = 7$
2343.01	166834768	4.63	4	PC	3.26	6.74	36.82	15.09	12.91	1	Small depth-aperture correlation
2199.01	369455629	4.62	0	KP	3.38	8.61	55.81	23.94	11.98	1	WASP-182; published (Nielsen et al., 2019)
4401.01	178170828	4.59	4	PC	0.45	1.70	3.55	3.11	11.85	1	-
134.01	234994474	4.54	0	P	1.40	1.74	3.68	3.00	11.01	1	Published (Astudillo-Defru et al., 2020)
1973.01	47617161	4.52	B	VPC+	3.92	8.69	56.73	22.60	11.39	1	CHIRON RVs show large scatter; large astrometric excess noise
2683.01	200320748	4.52	4	PC	1.53	7.42	43.35	21.76	12.74	1	Slightly v-shaped
160.01	253917293	4.51	B	NPC	4.61	7.94	48.60	17.31	11.74	1	Event on a different star
451.02	257605131	4.44	B	P	1.86	1.93	4.40	2.35	10.94	3	Published (Newton et al., 2021); known young system so PRV difficult
2522.01	148883384	4.43	4	PC	2.10	1.66	3.40	1.73	9.80	1	-
2000.02	371188886	4.36	0	PC	3.10	2.53	6.95	3.09	10.98	2	Published (Sha et al., 2023)
4644.01	387423096	4.33	2	PC	0.32	1.78	3.84	4.75	12.14	1	Validated by ShARCS AO (Dressing in prep)
4527.01	380887434	4.26	2	PC	0.40	0.91	0.69	0.97	10.92	1	Validated by ShARCS AO (Dressing in prep); under observation (Nowak)
3409.01	405485201	4.23	4	PC	7.76	4.91	21.46	6.39	11.94	1	-
494.01	19519368	4.22	2	PC	1.70	1.73	3.66	2.22	10.85	1	Close neighbours; under observation (Veloce and Nowak)
5166.01	59582240	4.18	4	PC	1.77	1.71	3.57	1.93	10.72	1	-
Continued on the next page...											

Target		Nomads		SG1 status and TOI release values							Notes (brief)
TOI	TIC	Merit	Rank	SG1 Disp.	P (days)	R _P (R _⊙)	M _P (M _⊙)	K (m s ⁻¹)	V _{mag}	N _{pl}	
4388.01	313554509	4.17	4	PC	4.48	5.70	27.68	11.61	12.97	1	-
1027.01	20318757	4.13	0/1	VPC+	3.28	2.70	7.78	4.37	11.31	2	⇒ observed as part of older version of sample; continued from HARPS NCORES observations; under observation (Gandolfi, priv. comm.) ; validated by Louie (in prep)
3201.01	329921262	4.13	B	PC	2.36	7.45	43.65	20.95	12.84	1	No stellar radius
2216.01	39018923	4.13	3	VPC?	5.68	6.48	34.42	10.74	12.56	1	-
724.01	339961200	4.11	2	CPC?	3.21	2.27	5.80	2.66	10.39	1	Potentially v-shaped
669.01	124573851	4.10	0	VPC?	3.95	2.66	7.57	3.07	10.61	1	Observed by TESS-Keck Survey (Murphy in prep)
168.01	369457671	4.08	1	VPC+	2.31	3.39	11.46	6.77	12.48	1	⇒ observed as part of older version of sample
2374.01	439366538	4.05	1	VPC+	4.31	8.01	49.31	22.99	12.10	1	⇒ observed as part of older version of sample
2467.01	71013298	4.01	2	PC	1.76	1.77	3.80	2.34	11.03	1	Centroid offset but not centered on another star
2827.01	440775339	4.00	B	VPC	5.34	7.70	46.20	15.83	11.79	1	TRES $v \sin i = 10$
2489.01	279327604	3.94	B	APC	1.43	8.83	58.23	35.21	12.19	1	BEB
4869.01	71221700	3.92	B	PC	3.10	7.32	42.40	16.90	12.93	1	TRES $v \sin i = 8$; no stellar radius
4762.01	119069946	3.89	4	PC	4.36	3.62	12.80	4.82	12.05	1	ShARCS AO suggest false positive (Dressing in prep)
4930.01	307221757	3.87	4	PC	6.78	5.88	29.16	9.60	12.70	1	-
942.01	146520535	3.86	B	VP	4.33	3.52	12.20	5.49	11.98	1	CHIRON $v \sin i = 16$; published with mass constraint only (Zhou et al., 2021); young system
970.01	174599051	3.84	4	PNEB?	4.99	3.34	11.13	4.54	11.38	1	Some centroid offset not onto another star; flagged as PNEB
1113.01	290348382	3.84	B	NPC	6.44	4.95	21.74	11.39	12.49	1	Event on a different star
4178.01	289988797	3.79	B	SEB2	2.20	9.81	69.62	33.83	10.82	1	SEB2; TRES $v \sin i = 12$
4311.02	122617317	3.77	3	PC	0.99	1.42	2.59	1.82	11.39	2	-
5078.01	303523128	3.77	4	PC	1.51	1.50	2.85	1.96	10.87	1	-
2494.02	282576340	3.73	0	CPC	2.41	2.24	5.64	2.82	11.48	2	In prep (Quinn, TESS wiki)
2523.01	443213925	3.72	4	PC	4.64	2.87	8.62	3.21	10.72	1	Potential centroid offset
2763.01	78441371	3.69	B	PC	4.00	8.03	49.54	19.95	12.42	1	No stellar radius
2986.01	148497855	3.64	3	VPC	3.28	8.33	52.78	21.45	12.63	1	Under observation (Yee)
2768.01	443556801	3.64	4	VPC?	1.51	3.00	9.30	7.05	12.66	1	ShARCS AO suggest false positive (Dressing in prep)
5174.01	49428710	3.58	4	PC	12.21	5.71	27.75	7.70	11.58	1	-

Continued on the next page...

Target		Nomads		SG1 status and TOI release values							Notes (brief)
TOI	TIC	Merit	Rank	SG1 Disp.	P (days)	R _P (R _⊕)	M _P (M _⊕)	K (m s ⁻¹)	V _{mag}	N _{pl}	
252.01	237924601	3.55	B	APC	1.00	8.70	56.77	40.47	12.46	1	Likely BEB
2345.01	91555165	3.55	2	CPC?	1.05	1.42	2.61	2.02	11.48	2	-
2356.01	30947715	3.52	B	PC	3.66	3.54	12.29	4.59	12.40	1	TRES $v \sin i = 8$
5116.01	149496868	3.50	2	VPC?	3.44	4.19	16.39	8.13	12.82	1	Validated in Christiansen et al. (2022)
2195.01	24695044	3.43	1	VPC	4.16	8.92	59.24	25.06	11.40	1	-
3106.01	124453339	3.42	4	PC	1.49	8.64	56.18	29.61	12.55	1	-
5168.01	363548415	3.37	0	KP	6.77	5.09	22.87	8.35	12.69	1	K2-27; published (Van Eylen et al., 2016)
941.01	408137826	3.33	B	VPC+?	8.51	6.99	39.11	11.65	11.46	1	TRES $v \sin i = 8$
4301.01	307079330	3.32	4	PC	6.54	3.45	11.80	4.01	10.71	1	-
4679.01	283605976	3.28	4	PC	6.82	6.32	32.97	11.02	12.71	1	-
5172.01	53287554	3.28	0/2	PC	4.72	3.12	9.96	4.72	11.59	1	Under observation (Gandolfi, 4 on SG4 coordinator)
2315.01	302333151	3.24	B	APC	1.44	9.78	69.34	34.37	11.34	1	Strong chromaticity, likely blend
2809.01	45655786	3.17	B	VPC+	3.17	9.78	69.30	26.69	11.41	1	CHIRON $v \sin i = 12$
3500.01	443666343	3.13	2	VPC?	7.35	4.00	15.14	5.19	11.55	1	-
4971.01	453942180	3.11	B	PC	7.97	5.89	29.29	9.37	12.79	1	No stellar radius
4806.01	96317115	3.10	4	PC	8.45	4.96	21.85	6.56	12.35	1	-
833.01	362249359	3.07	1	VP	1.04	1.35	2.40	2.15	11.72	1	-
4290.01	278196907	3.03	4	PC	1.18	3.10	9.83	6.34	12.99	1	V-shaped
177.01	262530407	3.00	4	P	2.85	2.13	5.18	3.63	11.40	1	-
2238.01	231077395	2.99	1	CPC	3.39	2.41	6.40	2.48	11.50	1	Poor transit shape
3082.01	428699140	2.96	2	VPC+	1.93	3.29	10.87	7.33	12.93	1	-
4747.01	235078021	2.92	4	PC	5.22	3.83	14.08	5.15	12.44	1	-
3498.01	310294189	2.89	B	PC	4.88	4.71	20.04	9.40	13.00	1	Likely blended binary
2850.01	453105377	2.85	4	VPC	6.63	7.45	43.59	18.46	12.07	1	V-shaped
4337.01	399642071	2.82	4	PC	2.29	2.82	8.35	5.58	12.67	1	V-shaped; possible instrument systematics
2302.01	144000801	2.81	B	APC	1.31	9.86	70.27	44.36	11.54	1	Likely BEB
493.01	19025965	2.79	1	VPC+?	5.94	4.15	16.15	7.61	12.55	1	⇒ observed as part of older version of sample
5068.01	366631954	2.76	0	KP	4.74	3.05	9.54	3.96	11.83	1	K2-182 b; published (Akana Murphy et al., 2021)
1075.01	351601843	2.73	0	P	0.60	1.67	3.42	3.76	12.62	1	Published (Essack et al., 2023)
4538.02	262715204	2.71	0	KP	3.47	2.66	7.58	3.52	12.08	2	K2-285 b; published (Palle et al., 2019)
2394.01	144297844	2.67	3	PC	3.50	3.47	11.89	5.25	12.79	1	Validated by ShARCS AO (Dressing in prep)
Continued on the next page...											

Target		Nomads		SG1 status and TOI release values							Notes (brief)
TOI	TIC	Merit	Rank	SG1 Disp.	P (days)	R _P (R _⊕)	M _P (M _⊕)	K (m s ⁻¹)	V _{mag}	N _{pl}	
921.01	278775625	2.66	4	PC	5.13	3.14	10.04	3.43	11.81	1	-
4378.01	394721720	2.63	4	PC	1.07	9.42	65.06	38.64	12.33	1	V-shaped
5541.01	231721006	2.62	B	VPC	9.02	5.61	26.93	8.06	12.88	1	Source of signal not on this star
872.01	220459826	2.61	1	VPC	2.24	2.71	7.80	4.46	12.71	2	⇒ observed as part of older version of sample
3428.01	271486001	2.58	4	PC	5.39	4.12	15.96	5.82	12.77	1	-
2408.01	67630845	2.50	4	VPC	9.46	5.44	25.60	8.90	12.78	1	-
1201.01	29960110	2.47	0	P	2.49	2.22	5.57	4.28	12.26	1	Published (Kossakowski et al., 2021)
148.01	393940766	2.45	0	BD	4.87	8.24	51.82	18.02	12.40	1	Brown Dwarf; published (Grieves et al., 2021)
209.01	52204645	2.45	1	VPC	4.38	2.68	7.67	3.45	11.67	1	⇒ observed as part of older version of sample
5553.01	218354669	2.44	4	PC	1.76	1.55	3.03	1.93	11.83	1	-
275.01	373844472	2.39	B	APC	0.92	9.85	70.20	46.15	11.92	1	APC/BEB
4746.01	160180298	2.33	4	PC	7.54	4.42	17.96	6.33	12.57	1	-
2428.01	274193763	2.20	B	APC	7.57	4.22	16.63	5.41	12.46	1	No stellar radius; large <i>Gaia</i> RUWE
2311.01	25139786	2.19	B	APC	1.59	9.01	60.31	36.58	12.90	1	Definite blend
3328.01	276647494	2.15	0/4	PC	10.19	5.47	25.82	7.46	12.83	1	Observed by CORALIE (9 on SG4 coordinator)
1078.01	370133522	1.99	0	VP	0.52	1.14	1.56	2.35	12.68	1	Published (Shporer et al., 2020)
167.01	149990841	1.97	B	VPC	4.45	8.92	59.25	22.43	12.24	1	Known null from HARPS Nielsen
2397.01	306337838	1.89	1	PC	6.00	7.59	45.04	16.17	12.82	1	-
4533.01	418761354	1.88	0	KP	2.17	1.82	3.97	2.45	12.40	1	K2-216 b; published (Persson et al., 2018)
4685.01	431693152	1.77	4	PC	7.03	3.62	12.82	4.37	12.26	1	-
620.01	296739893	1.76	0/1	VP	5.10	2.93	8.91	4.97	12.26	1	Validated and upper mass estimate from PRV (Reefe et al., 2022); could follow up to confirm mass
4918.01	60859922	1.66	4	PC	6.53	7.61	45.24	15.37	12.77	1	-
4363.01	325682877	1.61	2	VPC?	2.12	1.83	4.02	2.75	12.68	1	-
696.01	77156829	1.52	0/1	VP	0.86	0.97	0.88	1.26	12.60	3	Published with a mass upper limit (Silverstein et al., 2022); would need to obtain masses with EPRV
4566.01	269728501	1.42	4	PC	2.08	1.66	3.39	2.15	12.66	1	-

Table 4.4: The TOIs that make my final cuts, but are already published/in-prep/being followed up, or are known nulls, and thus do not need to be observed by Nomads. The mass column is colour coded: those with published ≤ 20 per cent mass errors are dark purple; those that are in-prep and expected to have ≤ 20 per cent mass errors (private communication) are mid purple; those that are being followed up but with an as-yet-unknown mass error are light purple; the single target (TOI-561.02) that doesn't quite make a ≤ 20 per cent mass error is yellow; and known nulls are red and do not have mass determinations. It is worth noting that TOI-2641.01 does not have a precise radius determination, and this is because the transit is grazing.

Target			TOI release values					Publication values			
TOI	TIC	Other identifier	Merit	V_{mag}	N_{pl}	P (days)	R_{P} (R_{\oplus})	P (days)	R_{P} (R_{\oplus})	M_{P} (M_{\oplus})	Publication
193.01	183985250	LTT-9779 b	19.72	9.78	1	0.79	4.51	0.7920520 ± 0.0000093	4.72 ± 0.23	$29.32^{+0.78}_{-0.81}$	Jenkins et al. (2020)
129.01	201248411	HIP 65A b	15.25	11.00	1	0.98	5.92	0.9809734 ± 0.0000031	$22.8^{+6.8}_{-5.5}$	1021 ± 25	Nielsen et al. (2020b)
2431.01	258804746	-	14.82	10.90	1	0.22	4.14	≈ 0.224	≈ 1.5	≈ 5	Malavolta (TESS wiki)
1036.01	146172354	-	13.47	10.97	1	3.78	6.09	-	-	$\leq 20\%$ error	Nielsen (priv. comm.)
576.01	408310006	WASP-166 b	11.40	9.35	1	5.44	7.00	5.443540 ± 0.000004	7.06 ± 0.34	32.1 ± 1.6	Hellier et al. (2019a)
465.01	270380593	WASP-156 b	10.82	11.56	1	3.84	6.06	3.836169 ± 0.000003	5.72 ± 0.22	$40.7^{+3.2}_{-2.9}$	Demangeon et al. (2018)
118.01	266980320	HD 219666 b	10.50	9.81	1	6.03	4.76	$6.03607^{+0.00064}_{-0.00063}$	4.71 ± 0.17	16.6 ± 1.3	Esposito et al. (2019)
880.01	34077285	-	10.29	10.10	3	6.39	5.03	-	-	$\leq 20\%$ error	Nielsen (priv. comm.)
141.01	403224672	HD 213885 b	9.87	7.95	1	1.01	1.61	$1.008035^{+0.000021}_{-0.000020}$	$1.745^{+0.051}_{-0.052}$	$8.83^{+0.66}_{-0.65}$	Espinoza et al. (2020)
192.01	183537452	WASP-29 b	9.70	11.33	1	3.92	6.79	3.922727 ± 0.000004	$8.88^{+0.63}_{-0.39}$	77.6 ± 6.4	Hellier et al. (2010)
829.01	276128561	-	9.49	10.96	1	3.29	4.29	-	-	$\leq 20\%$ error	Nielsen (priv. comm.)
4524.01	333657795	HD 20329 b	9.23	8.76	1	0.93	1.69	0.9261 ± 0.00005	1.72 ± 0.07	7.42 ± 1.09	Murgas et al. (2022)
431.02	31374837	-	9.12	9.12	2	0.49	1.47	$0.490047^{+0.000010}_{-0.000007}$	1.28 ± 0.04	3.07 ± 0.35	Osborn et al. (2021a)
564.01	1003831	-	8.78	11.10	1	1.65	7.98	1.651144 ± 0.000018	$1.02^{+0.71}_{-0.29}$	$1.463^{+0.10}_{-0.096}$	Davis et al. (2020)
2673.01	1018843	-	8.50	9.45	1	1.91	2.38	≈ 1.9	≈ 2.44	≈ 14.64	Serrano (TESS wiki)
1943.01	382980571	-	8.35	10.60	1	1.74	2.87	-	-	?	Quinn/Gandolfi (SG4 coordinator)
5556.01	55315929	HD 80653 b	8.31	9.45	1	0.72	1.61	0.719573 ± 0.000021	1.613 ± 0.071	5.60 ± 0.43	Frustagli et al. (2020)
181.01	76923707	-	8.21	11.19	1	4.53	7.06	4.5320 ± 0.000002	$6.9559^{+0.087}_{-0.101}$	$46.1687^{+2.71}_{-7.83}$	Mistry et al. (2023)
261.01	63898957	-	8.18	9.56	2	3.36	3.04	-	-	$\leq 20\%$ error	Gandolfi (priv. comm.)
132.01	89020549	-	8.15	11.20	1	2.11	3.52	$2.1097019^{+0.000012}_{-0.000011}$	$3.42^{+0.13}_{-0.14}$	$22.40^{+1.90}_{-1.92}$	Díaz et al. (2020)
824.01	193641523	-	7.60	11.31	1	1.39	3.05	$1.392978^{+0.000018}_{-0.000017}$	$2.926^{+0.202}_{-0.191}$	$18.467^{+1.843}_{-1.875}$	Burt et al. (2020)

continued on the next page...

continued on the next page...

Target			TOI release values					Publication values			
TOI	TIC	Other identifier	Merit	V_{mag}	N_{pl}	P (days)	$R_{\text{P}} (R_{\oplus})$	P (days)	$R_{\text{P}} (R_{\oplus})$	$M_{\text{P}} (M_{\oplus})$	Publication
510.01	238086647	-	9.66	9.02	1	1.35	2.14	-	-	-	Known null from HARPS NCORES
641.01	49079670	-	9.39	10.55	1	1.89	3.22	-	-	-	Known null from HARPS NCORES / KESPRINT / PSF
745.01	444842193	-	6.73	11.03	1	1.08	2.30	-	-	-	Known null from HARPS NCORES

Target		Nomads		SG1 status and TOI release values							Nomads observations	
TOI	TIC	Merit	Rank	SG1 Disp.	P (days)	R _P (R _⊕)	M _P (M _⊕)	K (m s ⁻¹)	V _{mag}	N _{pl}	Num. of obs.	Observation notes (brief)
422.01	117979455	21.44	2	CPC-	0.63	4.50	18.49	12.87	9.51	1	6	Flatline
570.01	126733133	13.39	1	VPC	1.47	7.28	41.99	20.14	10.40	1	8	Flatline, likely false positive due to stellar activity
855.01	269558487	12.22	1	VPC	1.83	6.66	36.08	15.00	11.18	1	39	Drift; flatline, likely false positive due to stellar activity
728.01	96097215	12.05	2	VPC?	1.60	4.58	19.05	10.21	11.15	1	5	Flatline, likely false positive due to stellar activity
355.01	183593642	11.49	1	VPC	1.04	4.60	19.26	11.21	11.50	1	13	Drift; flatline, likely false positive due to stellar activity
1975.01	467281353	10.75	2	PC	2.83	4.27	16.93	7.81	10.75	1	16 (-1)	Flatline, likely false positive due to stellar activity
2498.01	263179590	9.60	2	VPC-?	3.74	4.79	20.58	8.04	11.20	1	16 (-1)	≤ 20% error on mass, published
2358.01	124095888	9.42	3	PC	1.13	5.13	23.12	11.99	12.14	1	17	Possibility of planet, requires activity detrending
2539.01	201508515	8.92	3	PC	0.99	5.20	23.69	15.04	12.30	1	7	False positive; double-lined spectroscopic binary
4461.01	149282072	8.65	3	PC	3.70	4.49	18.48	8.61	11.25	1	26 (-3)	Flatline, requires activity detrending
426.01	189013224	7.71	3	PPC	1.32	2.16	5.32	3.06	10.12	1	78 (-1)	Possibility of planet, requires activity detrending
4537.01	251039147	7.37	2	VPC?	6.66	3.88	14.42	4.60	8.74	1	7	Flatline
2365.01	344085117	6.68	3	VPC?	0.66	6.08	30.92	21.98	12.64	1	10	Flatline
502.01	172193428	6.64	1	VPC	2.94	8.83	58.29	24.30	10.94	1	10	Drift; flatline, likely false positive due to stellar activity
5005.01	282485660	6.48	3	VPC?	6.31	6.02	30.33	10.16	11.87	1	30	≤ 20% error on mass
3071.01	452006073	6.29	1	VPC-?	1.27	7.11	40.30	20.74	12.38	1	14	≤ 20% error on mass, in prep
2224.01	388198242	6.28	1	VPC	0.51	4.97	21.93	18.88	12.83	1	37	False positive; double-lined spectroscopic binary
1839.01	381714186	6.21	1	APC	1.42	2.24	5.66	3.36	10.90	1	71	Possibility of planet, requires activity detrending
5559.02*	456945304	6.18	1	PC	2.14	2.76	8.08	4.21	11.109	1	60 (-2)	≤ 20% error on mass, in prep
835.01	405700729	6.04	1	VPC+	4.79	4.20	16.46	6.24	11.49	1	12	Likely false positive due to stellar activity
1117.01	295541511	6.01	2	CPC	2.23	2.68	7.69	3.72	11.02	1	79 (-2)	≤ 20% error on mass, in prep
499.01	123702439	5.81	1	VPC-	8.53	6.59	35.39	9.75	10.59	1	12	Likely false positive due to stellar activity
332.01	139285832	5.70	1	VPC+	0.78	3.28	10.81	8.18	12.35	1	16	≤ 20% error on mass, published
908.01	350153977	5.30	1	VPC	3.18	3.10	9.81	4.38	11.32	1	45 (-3)	≤ 20% error on mass, published
2411.01	10837041	5.23	2	VP	0.78	1.75	3.71	3.46	11.27	1	54 (-2)	≤ 20% error on mass, in prep
2227.01	405425498	5.03	2	CPC	4.22	3.01	9.34	3.62	10.52	1	7	Flatline

Table 4.5: The final sub-sample of TOIs to follow up with Nomads, as described in Section 4.3.3, alongside brief notes on the observations taken, which are described in Sections 4.4 and 4.5. In the number of observations column, the negative number in brackets indicates how many points are unusable due to e.g., bad weather. The outcomes of each target’s observations are colour-coded: binaries are purple; null results are red; possible planets are turquoise; confirmed planets that are not yet in prep are blue green; in prep are yellow green; published are green.

*5559.02 is also known as K2-370 b.

4.4 The observations

The HARPS large programme ran from October 2021 - September 2023 inclusive (P108-P111), and was awarded a total of 20.75 nights in P108, 16 in P109, 21.5 in P110, and 16 in P111. A total of 20.44, 15.31, 20.78, and 16.44 nights were successfully observed in P108, P109, P110, and P111 respectively.

These 2 years afforded a baseline long enough that: (a) targets could be visited in ≥ 2 periods of visibility, allowing us to revisit a target if we did not obtain enough RVs to obtain a mass to the 20 per cent error goal before it set; (b) stellar variability of the host star could be tracked over multiple cycles if necessary; and (c) any long term trends like drift could be accounted for.

The overarching goal was to gather enough points for each target to either determine it was a null result, or obtain a precise mass determination with an error of ≤ 20 per cent. To this end, data was analysed as described in Section 4.5 daily where possible, and at the end of each run at minimum to try and achieve this goal in as few observations as possible.

To cover as much of the RV phase curve as possible, I aimed to have targets with periods $\lesssim 1.5$ days observed twice a night with as long a gap between exposures as possible, and targets with periods > 1.5 days were aimed to be observed daily. A minimum of one run's (~ 6 day's) worth of observations would be taken, and data analysed before deciding the continuing strategy. If a target was seen to flatline (i.e. where by a visual inspection, the variation in the RV points does not exceed the errors), or there were correlations with stellar activity that indicated it was a false positive, it was dropped. If a drift was seen, a gap in observations was allowed to quantify the drift over a longer baseline.

The visibility of targets over the 2 years and which months they were observed in is recorded in Table 4.6. Tracking visibility was imperative to plan ahead and ensure a target was finished before it set where necessary.

Target details			Year 1 (2021-22)												Year 2 (2022-23)												
			P108						P109						P110						P111						
TOI	RA	Dec	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	
422.01	04:47:12.76	-17:15:11.07												R	V												
570.01	09:07:49.24	-51:52:01.55		V																							
855.01	01:02:10.30	-02:08:40.31		V	V							R		V	V												
728.01	09:47:34.02	-07:19:50.65																	V								
355.01	01:11:33.12	-36:42:32.76		V																							
1975.01	11:12:18.27	-60:55:57.30										S	S						V								
2498.01	06:21:39.89	11:15:06.24													R	V	V										
2358.01	13:05:08.61	-31:59:07.14																			V	V					
2539.01	06:12:09.22	-30:47:16.42														V											
4461.01	18:36:35.30	16:27:08.73																					V	V			
426.01	05:16:24.52	-15:30:36.47													V	V	V	V	V	S							
4537.01	22:42:21.97	-06:52:16.06												V	V												
2365.01	07:12:04.17	-50:14:51.68															V										
502.01	06:34:07.68	-28:24:15.15	V																								
5005.01	15:52:25.98	-48:08:41.98																		V	V	V		V	V		
3071.01	11:33:06.98	-56:30:12.27																									
2224.01	04:04:58.33	-78:22:29.53	V		V	V										V											
1839.01	13:07:19.96	05:51:08.32									V	V												V			
5559.02	04:31:16.96	19:49:54.50														V	V	V	S								
835.01	12:48:13.87	-56:53:39.25																									
1117.01	18:14:24.53	-66:25:11.31										V	V	V	V	S							V	V	V	V	V
499.01	07:42:33.32	-43:33:13.82														V	V										
332.01	23:12:14.10	-44:52:34.77		V	S	N				R																	
908.01	03:32:38.22	-81:15:02.66	V	V	V	S																					
2411.01	01:23:41.51	-08:42:04.98												V		V	V							R	V		V
2227.01	22:30:04.28	-67:51:00.97										V	V														

Table 4.6: Target visibilities and observations made. Letters in the header denote the month, where P108/P110 are October-March, and P109/P111 are April-September. Visibility is indicated by the colour and letter in the cell: dark blue (N) indicates the target is not observable; orange (R or S) indicates that the target is becoming observable, i.e. it's rising or setting in the sky; and yellow (V) indicates that the target is easily observable at the required airmass for the entire length of exposure required. If a cell has a white circle, it indicates that the target was observed at least once in that month.

4.5 First-look analysis of target data

For each target, I go through a routine of data checks and fits using DACE. I use the simple linear regression for this first-look analysis, rather than running a full MCMC.

The first check is of the shape of the CCF for each observation, shown in Fig. 4.7. They should show a clean, Gaussian dip (see e.g., the CCF for TOI-2365). If not, this can be for a few reasons. The data might not have been reduced properly, or the observation gone awry (e.g., stopped mid-exposure, see TOI-908). A double-lined structure is indicative of a binary star system, and so in this case the TOI is a false positive (see e.g., TOI-2539).

Next, I check the stellar activity indicators (FWHM, bisector span and contrast of the CCF, and S-index and H-alpha index) in two ways. The correlation between each indicator and the RVs is examined, and the Pearson’s R statistic is calculated as an indicator of correlation strength. If the correlation is strong, it indicates that the RVs are varying due to stellar activity, rather than planetary presence. Then, I inspect the periodogram of each indicator – ideally, these are quiescent and do not display any significant peaks (quantified by the FAP). However, in a lot of cases there will be peaks showing the periods of stellar activity cycles. It is important to check whether any of these peaks are at the same period as the planet candidate. If there is a peak at the planetary period, then it is likely the candidate is a false positive caused by activity.

The RVs themselves are then appraised. In the simplest of cases, I fit a Keplerian with the eccentricity and argument of periastron fixed to 0 (as is reasonable to expect of planets on short orbital periods) and the epoch and period fixed to the values expected from the TOI release, which are presented in Table 4.6. This leaves the RV semi-amplitude as the only variable to fit. There are a few questions to consider. Are the RV points in phase with the expected sinusoid? Is the value for K comparable to the value we expect from a planet of this size? Is there an actual amplitude to the RV points, or is it on the same scale as the errors? A “flatline” may not preclude the presence of a planet, but if one *is* present then it is much less massive than expected (depending upon the number of observations taken and errors on the RVs) and HARPS precision can only obtain an upper limit on the mass. Sometimes, it may just be a false positive.

If there are enough points, then the periodogram of the RVs should start to show peaks at the periods of any signals present. These can be cross-checked against the peaks present in the activity indicators - a strong peak in the indicator periodograms will likely also show up in the RV periodogram. Ideally, a strong peak will appear at the planetary candidate period. However, this can sometimes be complicated by other signals - for example, a strong activity signal may need to be removed (by fitting out the activity with a simple sinusoid in this stage of the analysis; later, in a joint fit, this can be removed by e.g., a Gaussian process)

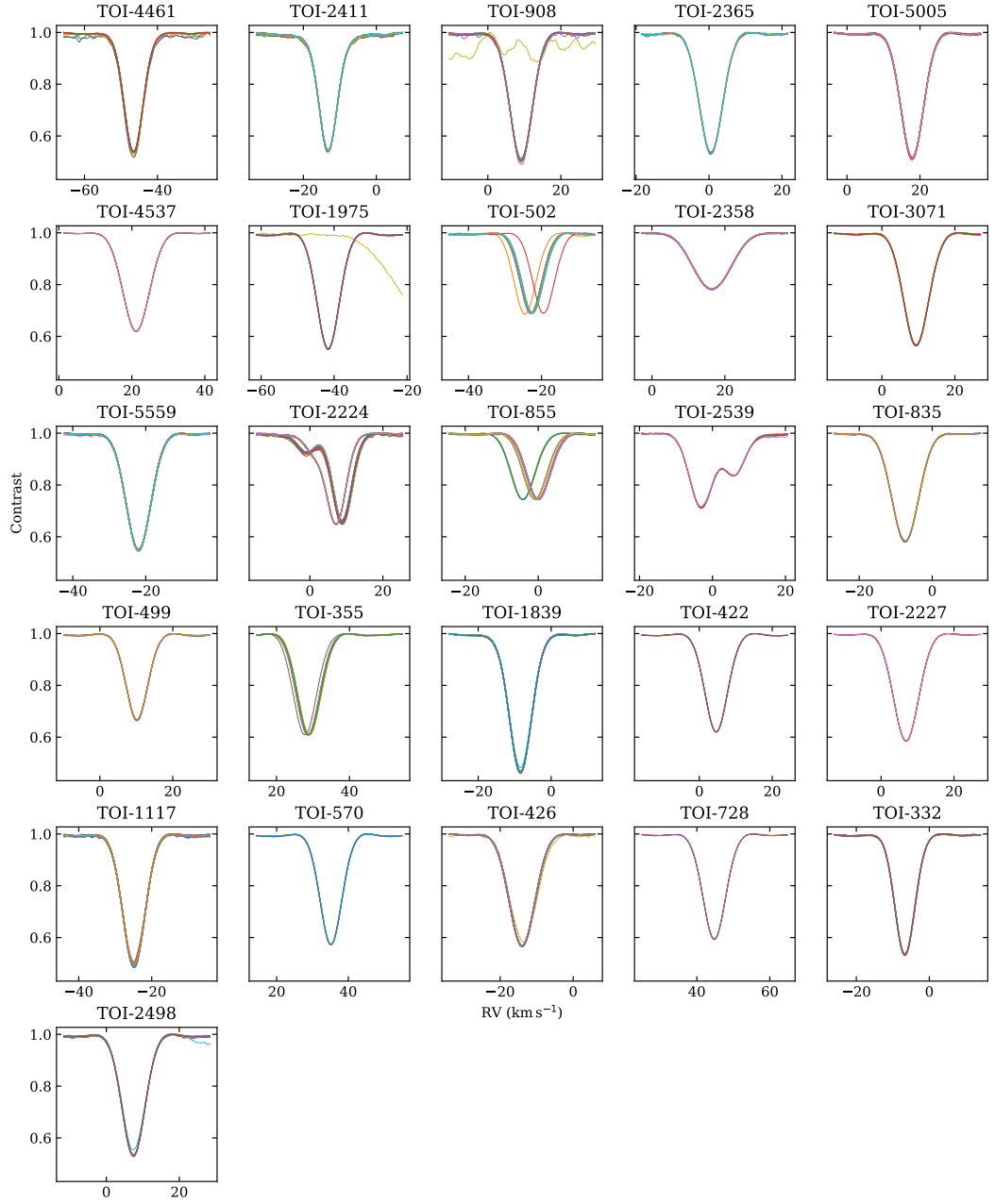


Figure 4.7: The CCFs for each observation of each target observed by Nomads.

before a peak at the planetary period will show.

If there is a visible drift, this must be removed before the RVs can be studied. In all Nomads cases, the baseline for the measurements (an absolute maximum of 2 years) is too short to see a full cycle of the drift, and so the portion that is seen can be fit with a simple quadratic of the form:

$$RV = at^2 + bt + c, \quad (4.6)$$

where t is the time (minus a reference time, t_0 , such that $t = t_{\text{HARPS,RJD}} - t_0$), and a , b , and c are the coefficients of the quadratic that are fit.

I continually performed this analysis as observation data came in. Catching false positives early means that a target can be dropped without wasting too much observation time. A target can also be stopped when enough points have been collected for a precise mass determination. It also allows me to adjust the baseline of the target - if there is a drift showing, we may wish to stop and revisit a target in a few months time to better quantify the drift signal.

If the analysis is positive, then the planet progresses forward to a proper analysis and joint fit of all the available data, photometry, spectroscopy, and high-resolution imaging.

Below, I show the data and summarise the results of these first analysis steps for each target in the Nomads sample, or provide a summary of a system when it has been published. There are several targets currently in preparation to be published - in these cases I do not show the data, but provide a quick summary, and direct the reader to their full publications when available. Brief notes on each target are available in Table 4.5, and the TOI release and predicted values from this table are reproduced in the subtitle of each system for ease of reading.

4.5.1 Systems already published - TOIs 332, 908, 2498

TOI-332 b

Release values: $P = 0.78$ days; $R_P = 3.28 R_\oplus$; $M_P = 10.81 M_\oplus$; $K = 8.18 \text{ m s}^{-1}$

TOI-332 b is the subject of Chapter 5. I will summarise in brief.

The CCFs show no abnormalities (Fig. 4.7). In the HARPS data, the star appears to be quiet, showing no correlation with the activity indicators (Fig. 5.8 and no significant peaks in the indicator periodograms (Fig. 5.9), including around the planet periodicity. The RV data is clean, with no drift (I introduced a gap in the data of ~ 150 days to confirm this), and a clear signal at the planetary period with a large, unmistakable semi-amplitude of $43 \pm 1 \text{ m s}^{-1}$ (Fig. 5.3 and Table 5.5). With just 16 observations, I obtain a very precise mass determination of $57.2 \pm 1.6 M_\oplus$. It is a clear confirmation of the planetary nature of TOI-332.01, now known as TOI-332 b.

TOI-908 b

Release values: $P = 3.18$ days; $R_P = 3.10 R_\oplus$; $M_P = 9.81 M_\oplus$; $K = 4.38 \text{ m s}^{-1}$

TOI-908 b is explored in Hawthorn et al. (2023b). Again, I will summarise in brief, and direct the reader to the publication.

The CCFs show no abnormalities (Fig. 4.7), except for one observation in particular (shown in yellow-green), taken on the night of 16 January 2022, and this was consequently removed. There were two further problematic observations: on the night of 25 December 2021, the observer noted that the exposure was stopped 15 min early due to bad weather, and so this data point has very large errors and is removed before analysis; there were similar weather issues on 23 December 2021, again removing an observation due to large errors. The RV data shows a strong signal at the periodicity of the planet, but also another signal at a longer ~ 20 day period, which is also reflected in the periodograms of the activity indicators (Fig. 4.8, left). There is determined to be activity with a period of 21.932 ± 6.167 days, which is detrended from the RV data using a GP. There is no drift. The planetary signal has a semi-amplitude of $7.244 \pm 1.768 \text{ m s}^{-1}$ (Fig. 4.8, right), and with a total of 42 observations, a mass of $16.137^{+4.112}_{-4.039} M_\oplus$ is determined. TOI-908.01, now TOI-908 b, is thus a confirmed planet.

TOI-2498 b

Release values: $P = 3.74$ days; $R_P = 4.79 R_\oplus$; $M_P = 20.58 M_\oplus$; $K = 8.04 \text{ m s}^{-1}$

TOI-2498 b is explored in Frame et al. (2023). Again, I will summarise in brief, and direct the reader to the publication.

The CCFs show no abnormalities (Fig. 4.7), aside from one observation on the night of 11 October 2022 (shown in turquoise) due to cloud coverage, and so this is removed before analysis. There is no correlation between the RVs and activity indicators, and no significant peaks in the indicator periodograms (Fig. 4.9, left). There is no drift in the RV data, and a clear signal at the planetary period with a semi-amplitude of $13.25 \pm 1.36 \text{ m s}^{-1}$ (Fig. 4.9, right). A total of 15 observations results in a precise mass determination of $34.62^{+4.10}_{-4.09} M_\oplus$, and TOI-2498.01, now TOI-2498 b, is confirmed to be a planet.

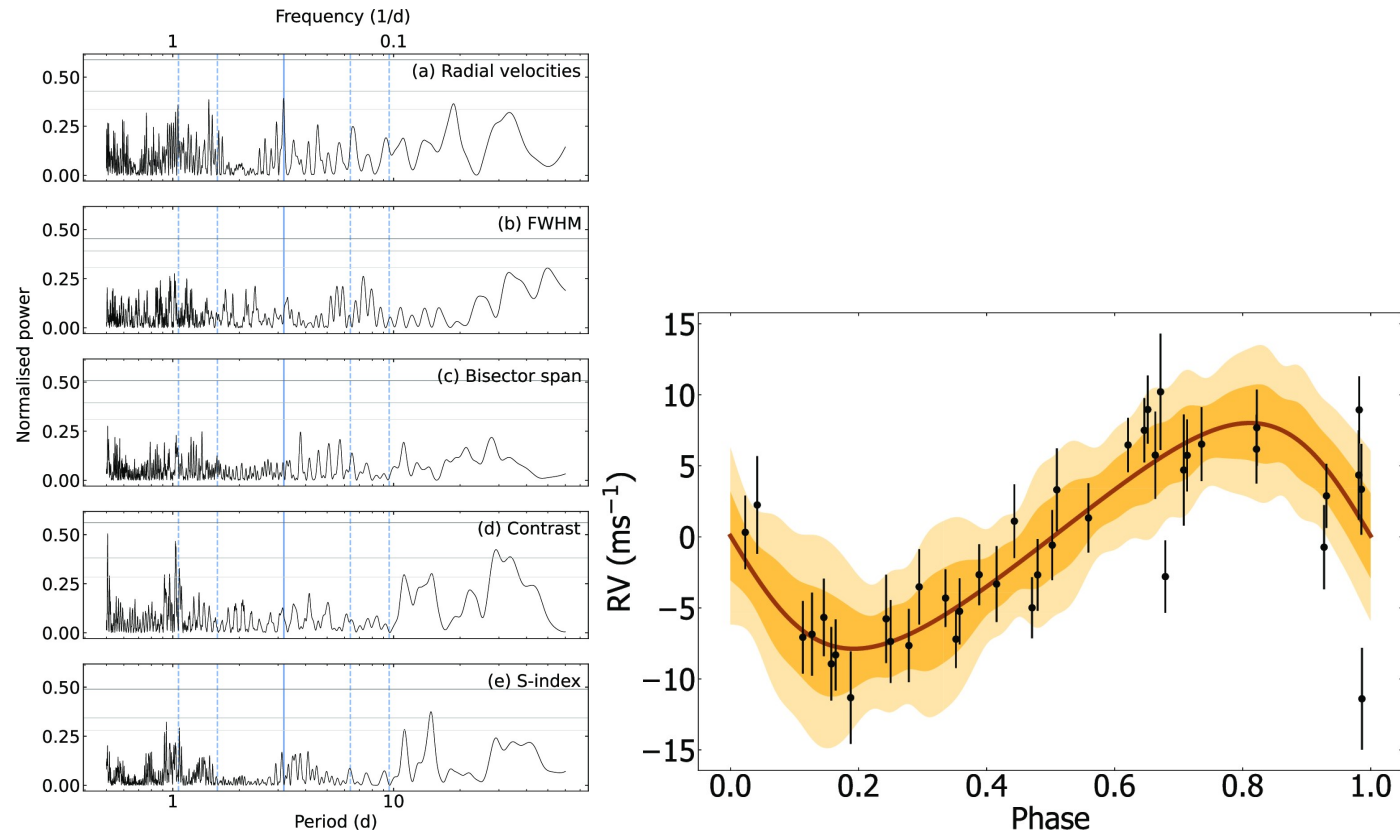


Figure 4.8: TOI-908. Left: periodograms of the RVs and select stellar activity indicators. The expected planetary period is shown as a blue line, with its aliases as dotted blue lines. Right: the phase fold of the fit to the RV data, where the stellar activity has been detrended with a GP. Reproduced from Hawthorn et al. (2023b).

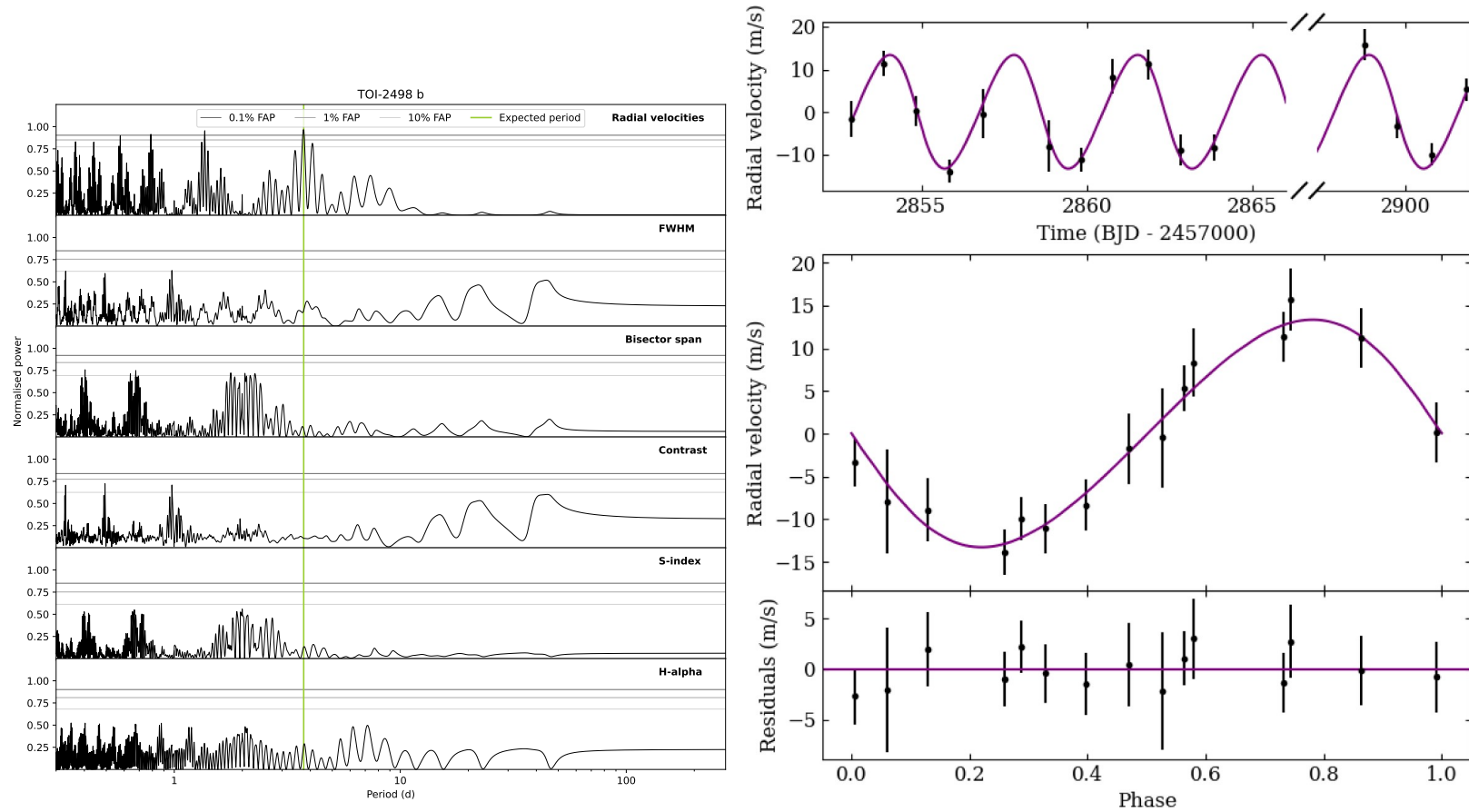


Figure 4.9: TOI-2498. Left: periodograms of the RVs and select stellar activity indicators. The expected planetary period is shown as a green line. Right: the RV data where the fit is shown in purple, with the timeseries in the top panel, phase fold in the middle panel, and residuals after the fit in the bottom panel. Left is created by me, right is reproduced from Frame et al. (2023).

4.5.2 Systems that are in-prep - TOIs 3071, 5559, 1117, 2411

There are four further systems in preparation to be published, all expected to have a precise mass determination with an error ≤ 20 per cent when their analysis is complete. Some of these are collaborations with additional RV data from other instruments. I do not show their data here, in anticipation of the publications, but summarise the important information for each.

TOI-3071.01

Release values: $P = 1.27$ days; $R_P = 7.11 R_\oplus$; $M_P = 40.30 M_\oplus$; $K = 20.74 \text{ m s}^{-1}$

This target is in preparation by Alejandro Hacker, on behalf of the Nomads consortium. The CCFs show no abnormalities (Fig. 4.7). There is a potential correlation with FWHM, but no significant peaks in any of the activity indicator periodograms. There is a very clear signal in the RV periodogram at the expected periodicity of the planet, and no other significant peaks at other periods except for a harmonic of the period. The Keplerian fit to the RVs is clean, showing a semi-amplitude of $\sim 34 \text{ m s}^{-1}$. The joint fit analysis of photometric and radial velocity data obtains a precise planetary mass of $0.22 \pm 0.01 M_J$ and radius of $0.72 \pm 0.04 R_J$.

TOI-5559.02 (a.k.a. K2-370 b)

Release values: $P = 2.14$ days; $R_P = 2.76 R_\oplus$; $M_P = 8.08 M_\oplus$; $K = 4.21 \text{ m s}^{-1}$

TOI-5559.02 was originally observed as part of the *K2* mission, and was subsequently re-observed by *TESS* and alerted as a TOI. It has been validated as a planet candidate in Christiansen et al. (2022). This target was also observed as part of the HARPS-N GTO program looking at *K2* targets, for a total of 80 observations, and Nomads obtained a further 60. As a result, this will become a joint publication between Nomads and the HARPS-N GTO team, with HARPS-N taking the lead. The CCFs show no abnormalities (Fig. 4.7). Our data shows a strong correlation with FWHM and contrast, though there are no peaks in the RV periodograms at the expected planetary periodogram. Instead, there are significant peaks around 14 days in FWHM, bisector span, contrast, and H-alpha, indicating a stellar activity cycle of this period. This peak also shows up as significant in the RV periodogram. A similar activity period is also seen in the HARPS-N data. Removing this activity signature with a simple Keplerian, the planet periodicity is fit well and results in a semi-amplitude of $5.443 \pm 0.836 \text{ m s}^{-1}$, agreeing with the HARPS-N data. A more complex treatment of the stellar activity will be applied in the further analysis by the HARPS-N GTO team.

TOI-1117.01

Release values: $P = 2.23$ days; $R_P = 2.68 R_\oplus$; $M_P = 7.69 M_\oplus$; $K = 3.72 \text{ m s}^{-1}$

This target is in preparation by Isobel Lockley, on behalf of the Nomads consortium. The CCFs show no abnormalities (Fig. 4.7). There are two observations that have large errors compared to the rest, from the nights of 24 September 2022 and 6 May 2023. Seeing was noted to be poor on those nights, and thus the points are discarded before analysis. There are no strong correlations with any of the activity indicators, nor are there any significant peaks at the expected planet periodicity in the activity indicator periodograms. There are some longer period significant peaks in the contrast at ~ 40 and 200 days. In the RV periodogram, the planet periodicity shows up above the 0.1 per cent FAP. Removing this signal reveals another above the 1 per cent FAP level at ~ 4.6 days, which does not correlate with any activity signals, and may be a second planet. Further removing this signal reveals yet another above the 0.1 per cent FAP level at ~ 8.7 days, which again does not correlate with any activity signals, and may be a third planet. After this, there are no further significant signals. The first planet, our target TOI-1117.01, has a semi-amplitude of $\sim 4.463 \pm 0.390 \text{ m s}^{-1}$ from my simple fit, which is precise enough to yield a corresponding precise mass measurement in a joint fit. The further two planet candidates have semi-amplitudes of $3.566 \pm 0.383 \text{ m s}^{-1}$ and $3.420 \pm 0.416 \text{ m s}^{-1}$ respectively, and will likely also have precise mass measurements in a joint fit. It is worth noting that the planets are in near 2:1 period commensurabilities, and thus an exploration of the stability of the system may be worthwhile.

TOI-2411.01

Release values: $P = 0.78$ days; $R_P = 1.75 R_\oplus$; $M_P = 3.71 M_\oplus$; $K = 3.46 \text{ m s}^{-1}$

This target has also been observed by Quinn using PFS, and Gaidos using MAROON-X; as such, it will become a joint publication between the three teams. Here I solely describe the Nomads HARPS data. The CCFs show no abnormalities (Fig. 4.7). There are two observations with larger errors compared to the rest, on the nights of 4 and 6 September 2022. This is likely due to poor weather, and they are discarded before analysis. There are no correlations with the activity indicators. There are also no significant peaks at the expected periodicity of the planet in the activity indicator periodograms, though there are significant peaks at longer periods: in FWHM, between ~ 21 and 23 days, and beyond 100 days; in contrast at ~ 24.5 days. In the RV periodogram, there are no significant peaks, though a cluster of peaks between ~ 21 and 25 days (corresponding to the activity peaks) almost

reach the 10 per cent FAP limit, and a peak shows at the expected periodicity of the planet. Fitting the planet signal increases the significance of the activity peak above 10 per cent, and it fits to ~ 27 days. The planet signal semi-amplitude fits to $9.615 \pm 0.711 \text{ m s}^{-1}$, which is larger than the expected amplitude of $\sim 3.5 \text{ m s}^{-1}$, suggesting a more massive planet than anticipated for its size. This is precise enough to result in a precise mass measurement in a joint fit, especially when the activity is detrended appropriately.

4.5.3 Another system with a less than 20 per cent error on mass - TOI-5005

TOI-5005.01

Release values: $P = 6.31$ days; $R_P = 6.02 R_\oplus$; $M_P = 30.33 M_\oplus$; $K = 10.16 \text{ m s}^{-1}$

The CCFs show no abnormalities (Fig. 4.7). There are no strong correlations with any of the activity indicators (Fig. 4.11). Nor are there any significant peaks in the activity indicator periodograms at the expected periodicity of the planet, though there are a few peaks beyond 10 days that might hint at longer period activity cycles, with one peak in particular around ~ 22 days. In the RV periodogram, the period of the planet shows up as a significant peak (Fig. 4.10). There is a little scatter in the fit, but fixing epoch and period to the values from the TOI release, I obtain a semi-amplitude of $11.781 \pm 0.902 \text{ m s}^{-1}$, which is on the scale of the expected semi-amplitude ($\sim 10 \text{ m s}^{-1}$) and not consistent with 0. Therefore, I conclude that this is confirmation of TOI-5005.01, which will likely result in a precise mass determination in a full joint fit analysis. The scatter would likely be reduced through a treatment of the longer period activity cycle (e.g., with a Gaussian process, see Section 3.3.2 for an example of this) – when the planet signal is removed, a peak at the earlier noted 22 days becomes significant in a periodogram of the RV residuals.

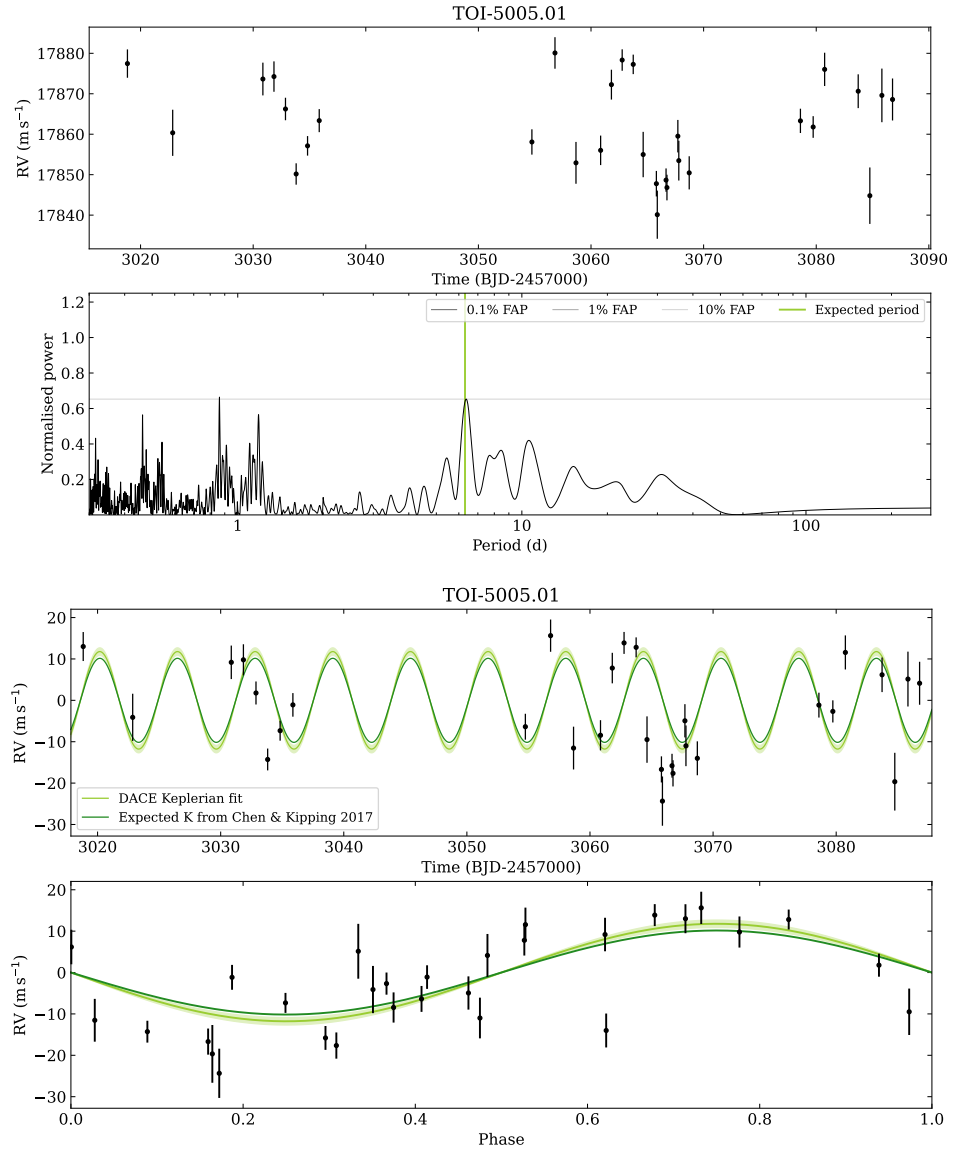


Figure 4.10: TOI-5005.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

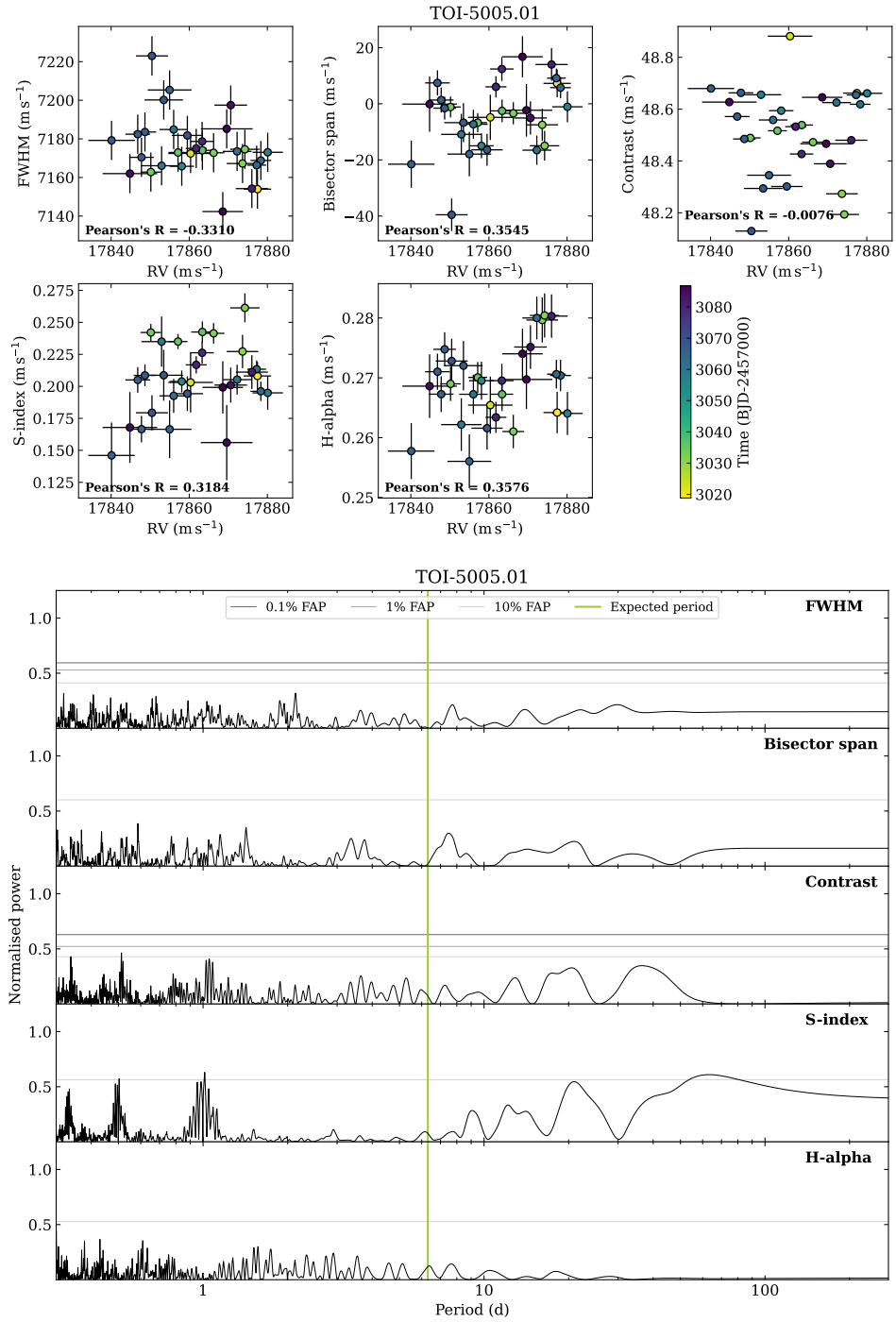


Figure 4.11: TOI-5005.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

4.5.4 Systems that may show a planet - TOIs 2358, 426, 1839

TOI-2358

Release values: $P = 1.13$ days; $R_P = 5.13 R_\oplus$; $M_P = 23.12 M_\oplus$; $K = 11.99 \text{ m s}^{-1}$

The CCFs show no abnormalities (Fig. 4.7). There are no strong correlations with the activity indicators. There is a forest of peaks around 1 day in each activity indicator periodogram due to sampling of the RV data, but no significant peaks at the expected periodicity of the planet. There are, however, significant peaks near and beyond 10 days, which are likely due to an activity cycle. In the RV periodogram, again there is a forest of peaks around 1 day which complicates seeing a peak at the expected planet period. Unfortunately, the errors on the RV data average 10.18 m s^{-1} , which is on the scale of the expected semi-amplitude of the signal ($\sim 12 \text{ m s}^{-1}$) and this will make it difficult to get a precise semi-amplitude and thus mass measurement. When fitting the planet signal with a simple Keplerian, the semi-amplitude is $4.937 \pm 3.905 \text{ m s}^{-1}$, which is consistent with 0 and less than the scale of the errors.

At first glance, this looks like it could be a flatline. However, removing the planetary signal reveals some small peaks in the periodogram of the residuals at around ~ 11 and ~ 20 days, which correspond to peaks in the activity indicator periodograms, and it is likely that these are harmonics of each other. This leads me to believe that the planetary signal is being complicated by an activity signal. To explore this, keeping the planetary Keplerian the same as before where only semi-amplitude is allowed to vary, I add another simple Keplerian with eccentricity and argument of periastron fixed to 0 and let period, semi-amplitude and epoch vary. This results in a better fit to the planetary signal (Fig. 4.13), with a now larger semi-amplitude of $10.077 \pm 5.126 \text{ m s}^{-1}$. The activity signal has a semi-amplitude of $7.792 \pm 4.397 \text{ m s}^{-1}$ and period of 11.264 ± 1.067 days. Due to the large errors on the RV points, it may be difficult to get a precise mass determination, but the next step for this target will be to treat the activity in a more complex way, e.g., detrended with a Gaussian process (see Section 3.3.2 for an example of this).

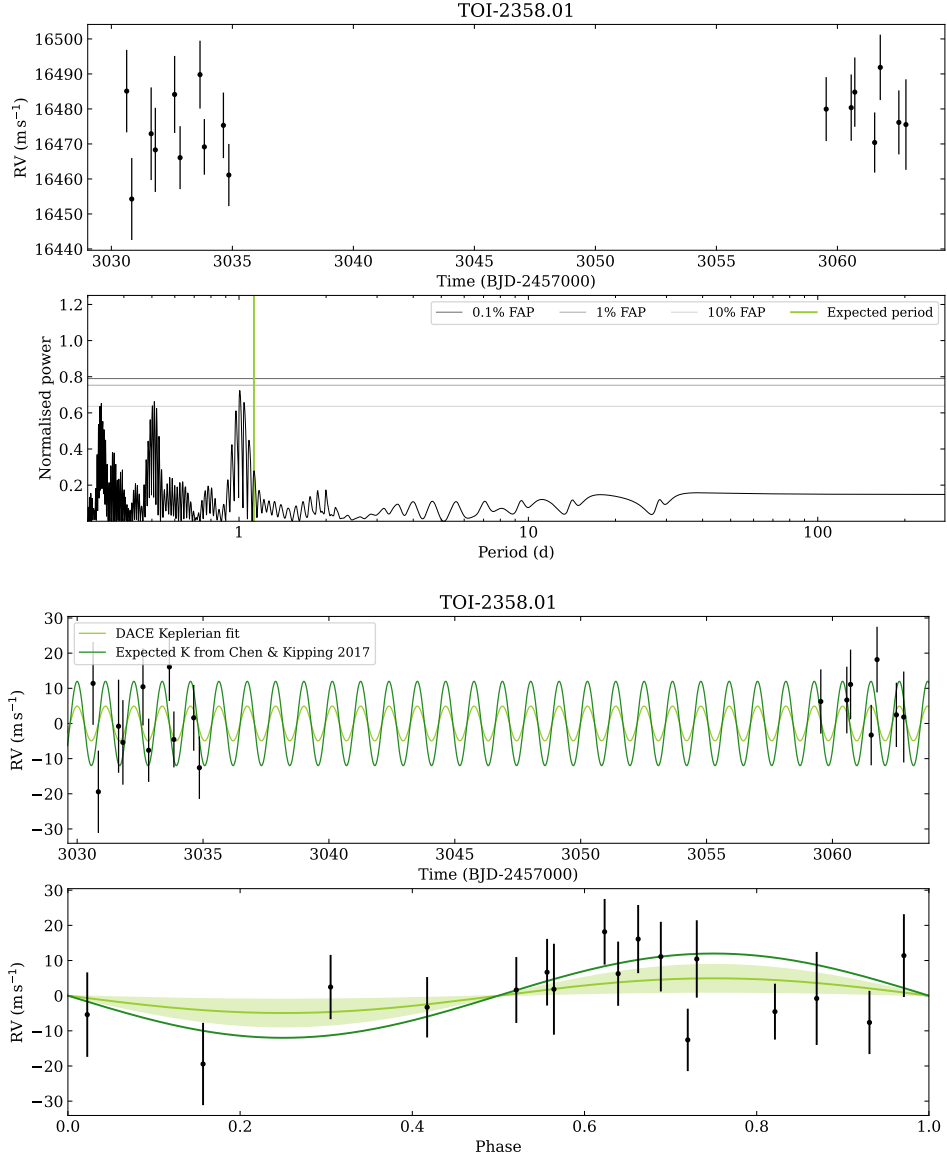


Figure 4.12: TOI-2358.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

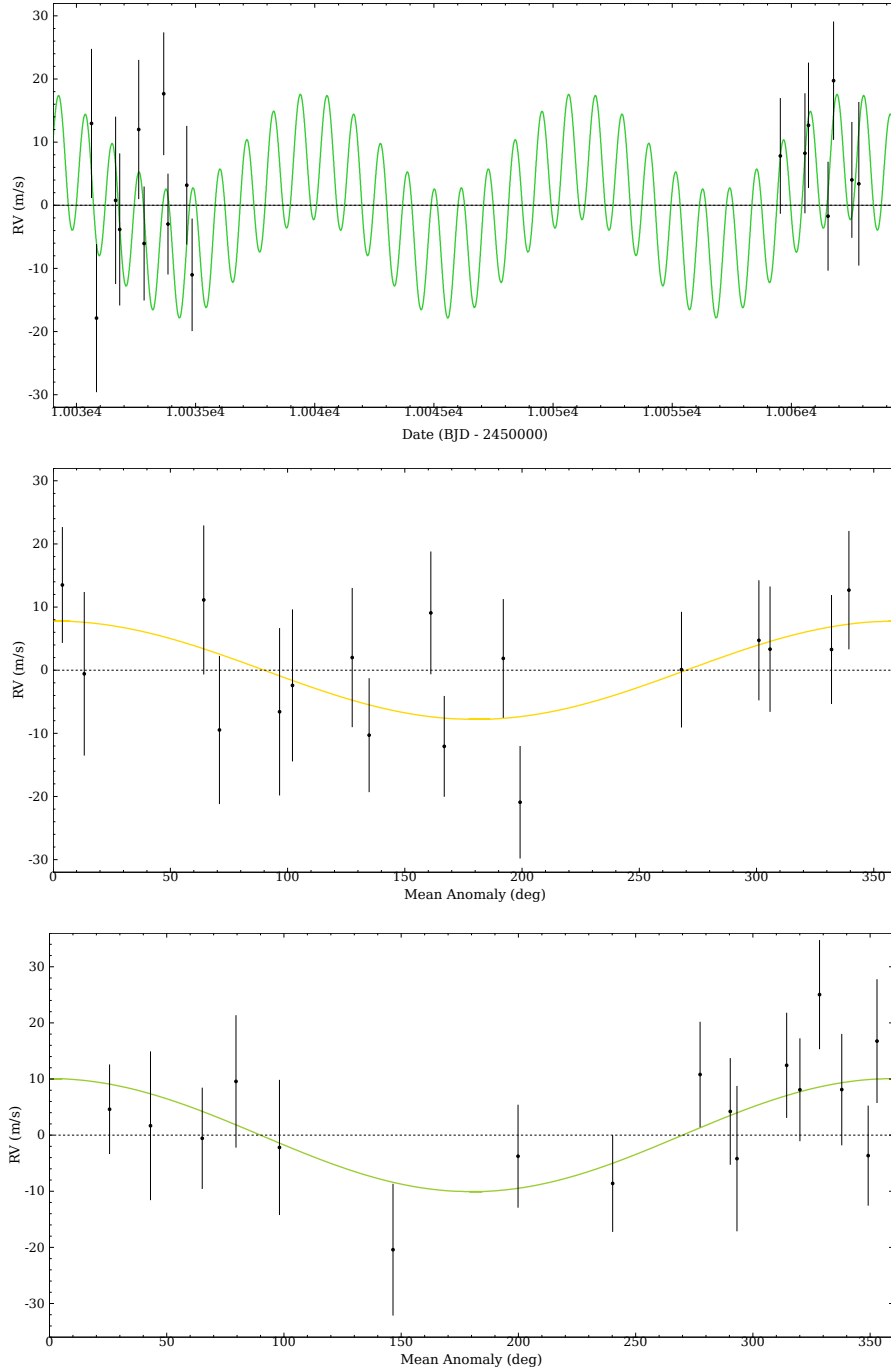


Figure 4.13: TOI-2358.01 RV data fit including a stellar activity signal. Top: the fit to the RV data, a combination of two sinusoids that fit the periodic activity signal and the planet signal (green). Middle: phase fold on the period of the activity signal, with the signal of the planet removed. Bottom: phase fold on the period of the planet signal, with the signal of the activity removed.

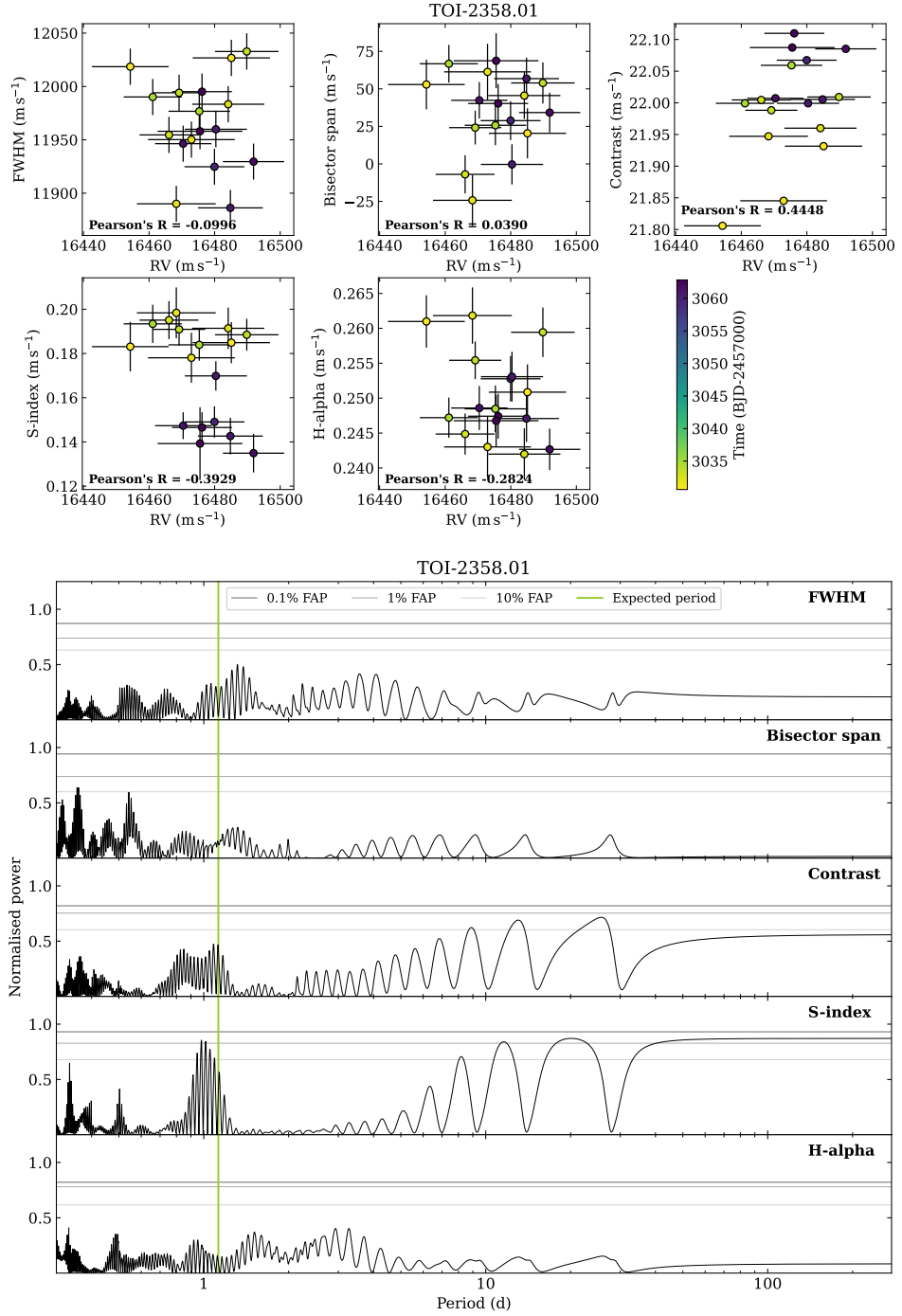


Figure 4.14: TOI-2358.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

TOI-426.01

Release values: $P = 1.32$ days; $R_P = 2.16 R_\oplus$; $M_P = 5.32 M_\oplus$; $K = 3.06 \text{ m s}^{-1}$

TOI-426 is complicated by stellar activity and drift. The CCFs show no abnormalities (Fig. 4.7). The raw RV data periodogram shows a tiny peak at the expected planetary period, but also has significant peaks beyond 10 days (Fig. 4.15, top). When the drift is modelled as a quadratic and removed, the periodogram changes, now showing a close-to-significant peak at the expected planetary period, and a significant peak at a period of around 12 days (Fig. 4.15, bottom). This latter peak correlates with strong peaks in all the activity indicator periodograms (Fig. 4.18). There are also strong correlations with the activity indicators both before and after the drift is removed (Fig. 4.17): before, with bisector span; after, with FWHM, bisector span, S-index and H-alpha. This all points to a periodic activity signal contaminating the RVs. As a preliminary analysis, to see whether the planet signal can be isolated, I model the activity signal as a simple Keplerian, fixing eccentricity and argument of periastron to 0, and allowing period, epoch, and semi-amplitude to vary (Fig. 4.16). This results in a signal fit with a period of 12.5653 ± 0.0141 days and a semi-amplitude of $17.311 \pm 0.334 \text{ m s}^{-1}$. The planet is also fit, fixing period and epoch to the TOI values and resulting in a semi-amplitude of $8.254 \pm 0.333 \text{ m s}^{-1}$, which is not consistent with 0 and more than the expected semi-amplitude of $\sim 3 \text{ m s}^{-1}$. However, the scatter is quite prominent. I am fairly confident that the planet signal is present, but in order to progress further in the analysis, the activity needs to be treated in a more complex way, e.g., detrended with a Gaussian process (see Section 3.3.2 for an example of this).

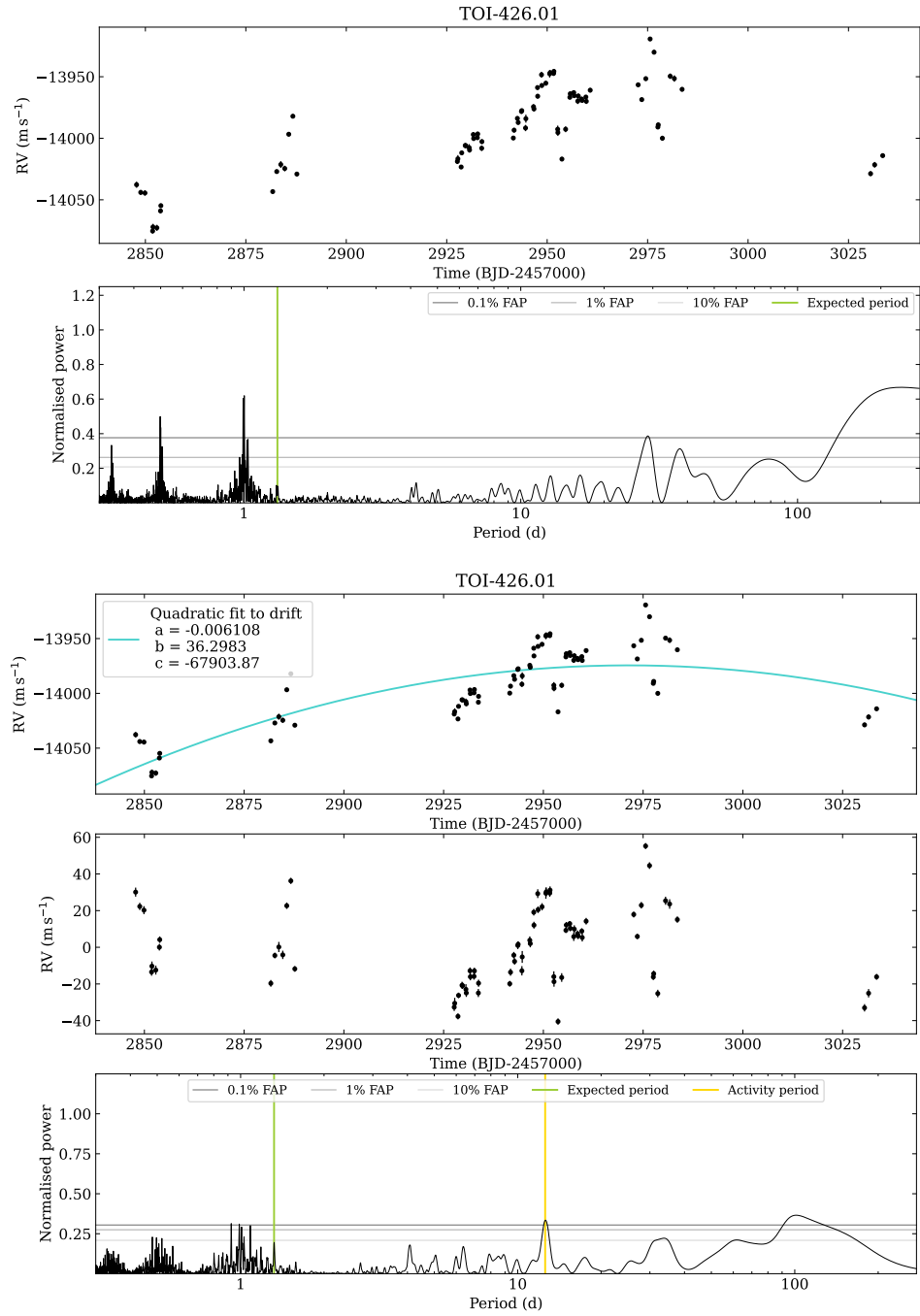


Figure 4.15: TOI-426.01 RV data. Top: the raw RV data. Top middle: periodogram of the raw RV data, with the expected planetary period shown as a green line. Middle: quadratic fit to the drift, with fit coefficients shown. Bottom middle: the RVs with the drift subtracted. Bottom: the periodogram of the RVs with the drift subtracted, with the expected planetary period shown as a green line, and the period of the stellar activity cycle shown as a yellow line.

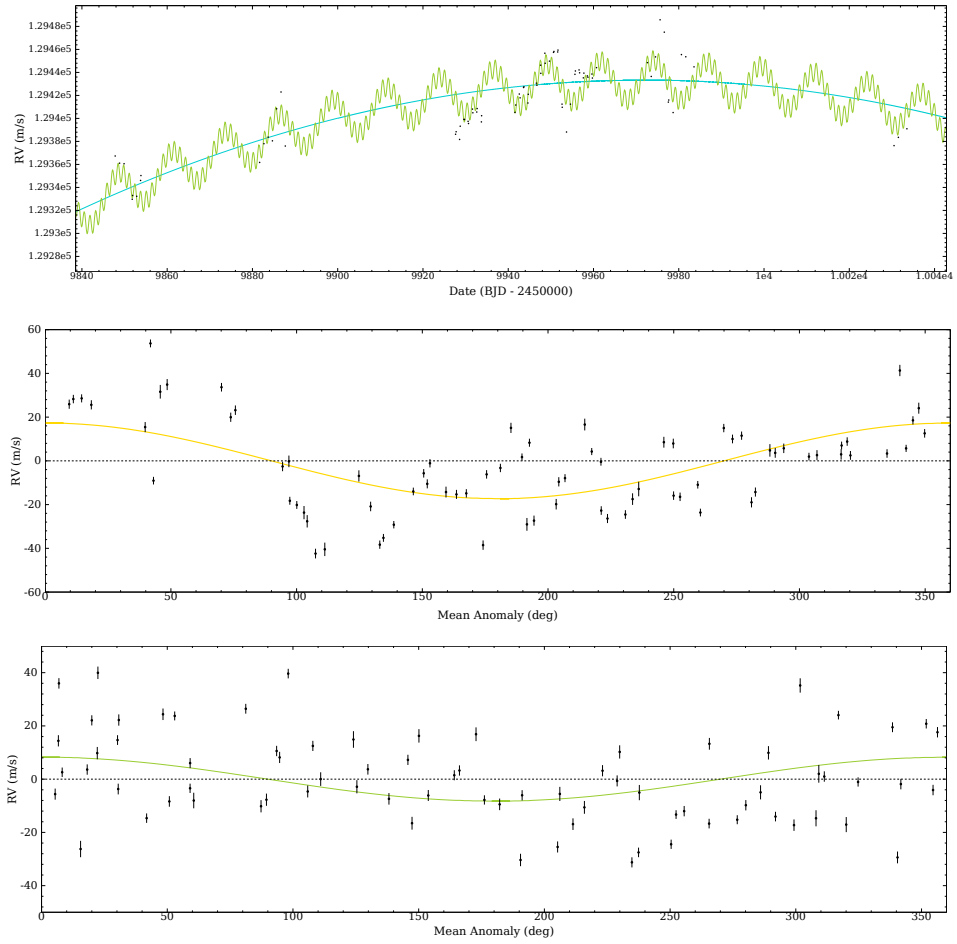


Figure 4.16: TOI-426.01 RV data fits. Top: the fit to the drift (turquoise), alongside a combination of two sinusoids that fit the periodic activity signal and the planet signal (green). Middle: phase fold on the period of the activity signal, with the signal of the planet removed. Bottom: phase fold on the period of the planet signal, with the signal of the activity removed.

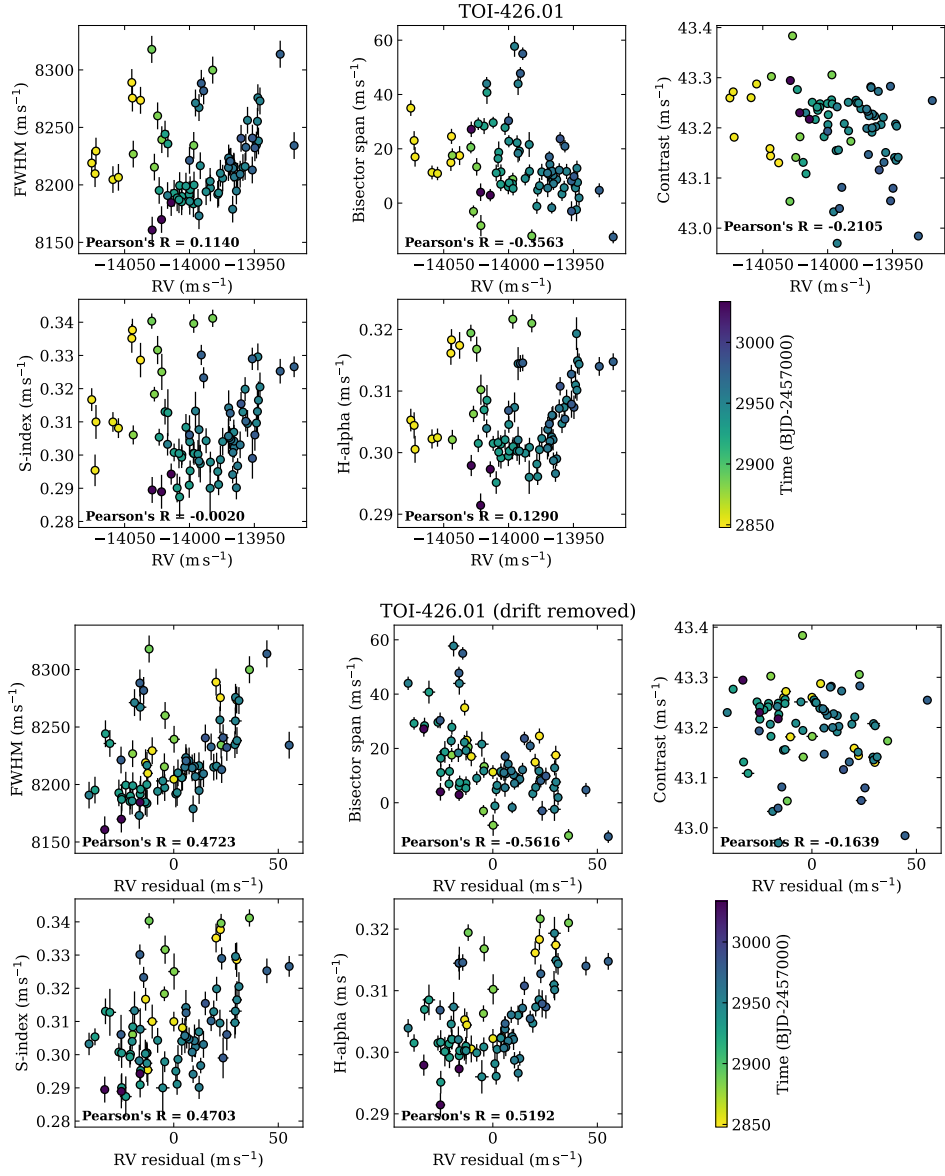


Figure 4.17: TOI-426.01: each activity indicator plot against RV, where the top panels show the raw RV data, and the bottom panels show the RV data after the drift has been subtracted. These show any potential correlation, where Pearson's R is given for each.

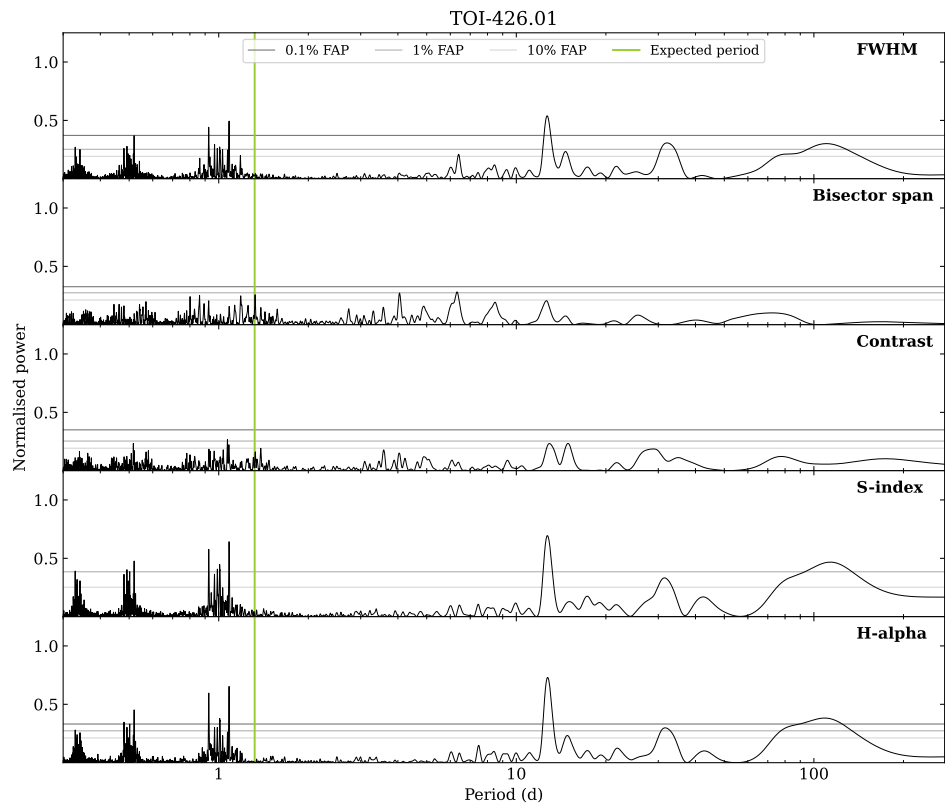


Figure 4.18: TOI-426.01: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

TOI-1839.01

Release values: $P = 1.42$ days; $R_P = 2.24 R_\oplus$; $M_P = 5.66 M_\oplus$; $K = 3.36 \text{ m s}^{-1}$

The CCFs show no abnormalities (Fig. 4.7). This target is markedly complicated by stellar activity. There is a strong correlation with H-alpha (Fig. 4.21). There are many signals in the activity indicator periodograms: FWHM shows a significant peak at ~ 23 days; bisector span at $\sim 12, 22$, and 44 days; S-index at ~ 24 days; H-alpha at $\sim 15, 40$, and 660 days. No peaks are seen at the expected planetary period.

Due to this multiplicity of activity signals, it becomes very hard to isolate the planetary signal without a proper detrending method. I, however, attempt to do this through the addition of multiple sinusoids at each significant periodicity which shows up in the RV periodogram (Fig. 4.19). The eccentricity and argument of periastron of each is fixed to 0, and the period, semi-amplitude, and epoch are allowed to vary. I start by fitting the longest period peak at ~ 790 days, then adding the signal which consequently becomes significant at ~ 32.5 days, and finally the last signal that pops up at ~ 12.2 days. This is shown in Fig. 4.20. This leaves a periodogram with a signal at the expected planetary period which is above the 1 per cent FAP threshold (alongside its harmonics). Adding this as the final element to the fit (where the period and epoch are fixed to the values from the TOI release, and semi-amplitude left to vary) results in a semi-amplitude of $4.085 \pm 0.334 \text{ m s}^{-1}$, which is consistent with, if not a little larger, than the expected semi-amplitude of $\sim 3.4 \text{ m s}^{-1}$. There are no further significant signals in the periodogram of the residuals. This bodes well for the existence of the planet, and a proper detrending of the stellar activity in a joint fit model should be able to obtain a precise mass measurement.

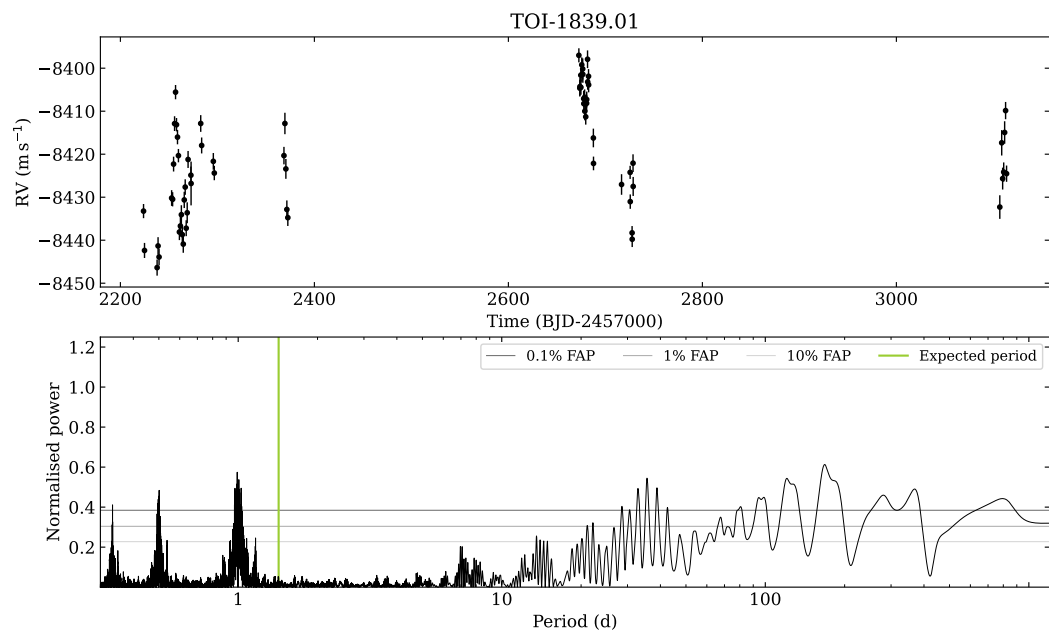


Figure 4.19: TOI-1839.01 RV data. Top: the raw RV data. Bottom: periodogram of the RV data, with the expected planetary period shown as a green line.

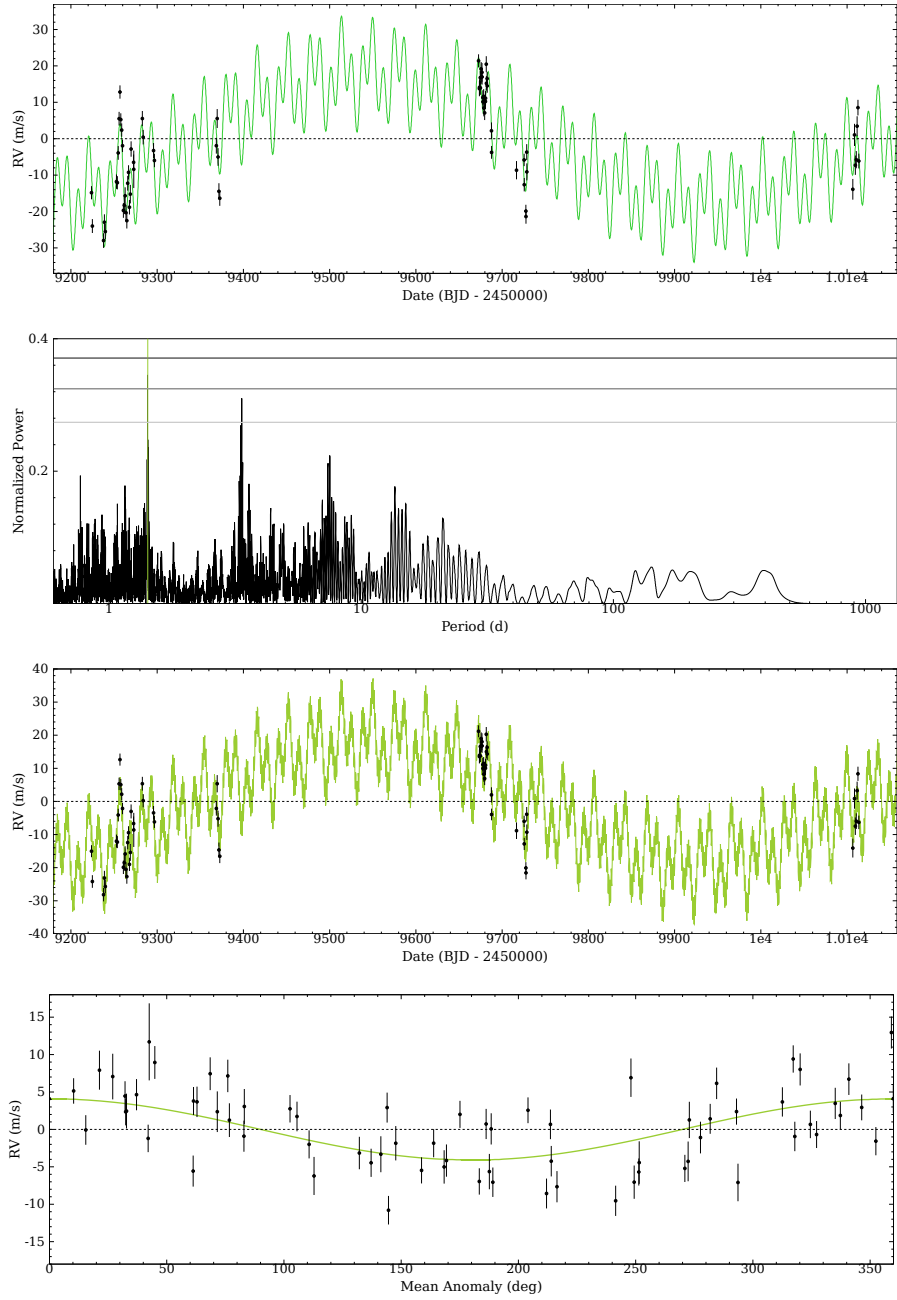


Figure 4.20: TOI-1839.01 RV data and activity fits. Top: three combined sinusoids to detrend the activity signals present in the RV data. Top middle: the periodogram of the residuals after the activity signals are removed, where the expected periodicity of the planet is marked in green and 0.1, 1, and 10 percent FAPs are marked from dark to light grey. Bottom middle: the three sinusoids detrending the activity plus the fit to the planet period. Bottom: the phase fold on the period of the planet, with the contribution of the stellar activity removed.

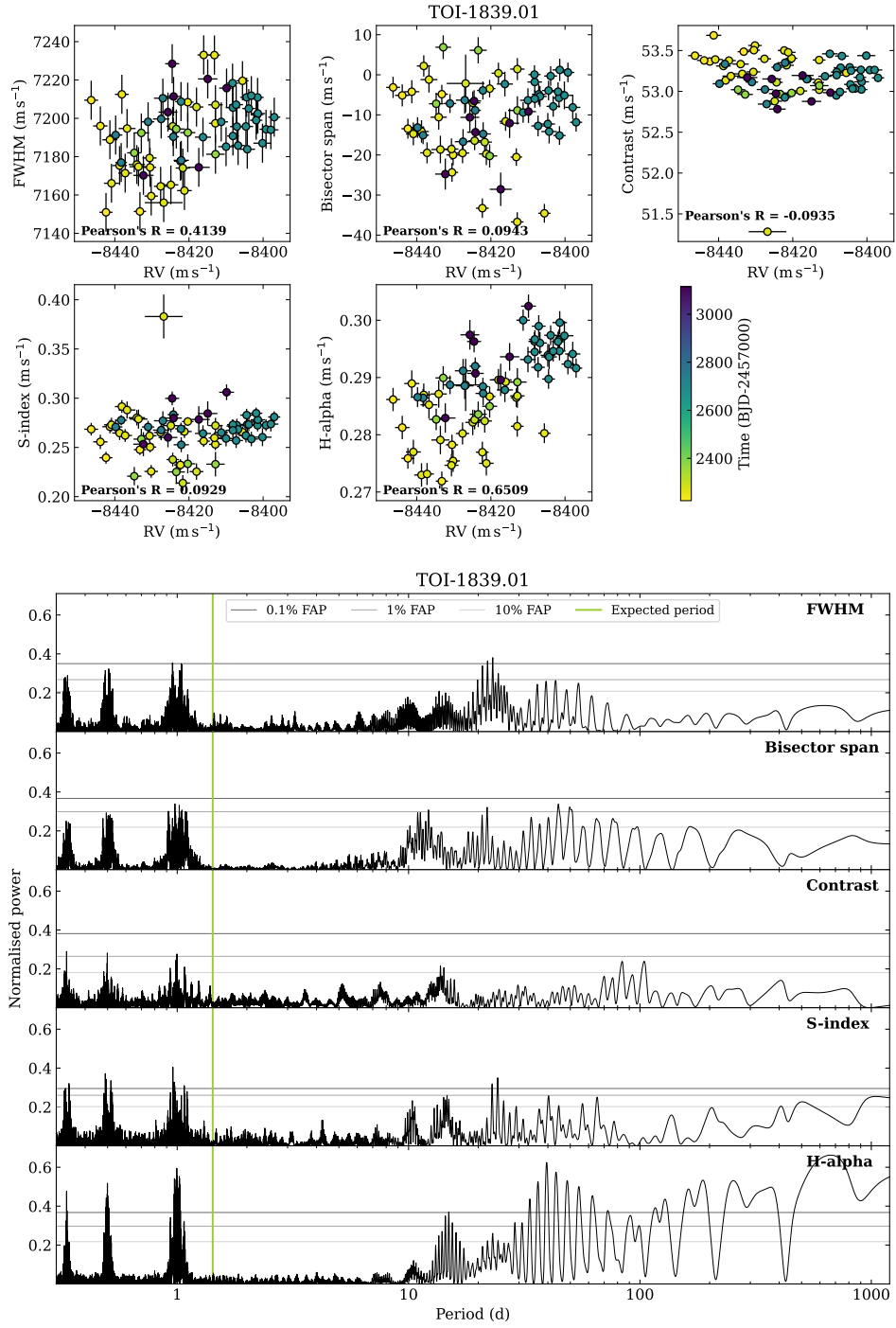


Figure 4.21: TOI-1839.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

4.5.5 Systems that are double-lined spectroscopic binaries - TOIs 2539, 2224

As seen in Fig. 4.7, both TOI-2539.01 and TOI-2224.01 show the double-lined CCF structure that is indicative of a binary star system, a common false positive scenario. As such, these targets were discontinued. I do not further analyse the RV data here, as the focus of Nomads is on planet discovery, but it could in future be used to characterise the binary systems.

4.5.6 Systems that show a null result - TOIs 422, 570, 855, 728, 355, 1975, 4537, 2365, 502, 835, 499, 2227

TOI-422.01

Release values: $P = 0.63$ days; $R_P = 4.50 R_\oplus$; $M_P = 18.49 M_\oplus$; $K = 12.87 \text{ m s}^{-1}$

The CCFs show no abnormalities (Fig. 4.7). There is a potential correlation with contrast, but no other activity indicator shows a correlation with RV (Fig. 4.23). There are not enough data points for the periodograms to be especially useful. Unfortunately, the RVs show no variation and so the target was discontinued - the Keplerian fit to the data results in a semi-amplitude of $1.664 \pm 1.341 \text{ m s}^{-1}$ (Fig. 4.22), which is comparable to 0, much less than the expected value of $\sim 13 \text{ m s}^{-1}$ and on the scale of the errors (with an average of 2.51 m s^{-1}). I therefore conclude this to be a flatline, and only an upper limit to the planet candidate could be obtained from this data.

TOI-570.01

Release values: $P = 1.47$ days; $R_P = 7.28 R_\oplus$; $M_P = 41.99 M_\oplus$; $K = 20.14 \text{ m s}^{-1}$

The CCFs show no abnormalities (Fig. 4.7). There is a potential correlation with FWHM, but no other indicator shows a correlation with RV (Fig. 4.25). While a peak is starting to show in the periodogram of the RVs at the expected periodicity of the planet, this is also mirrored in the periodograms for FWHM, bisector span, S-index and H-alpha, making it likely that this planet is a false positive due to activity. Additionally, the RVs show little variation and so the target was discontinued - the Keplerian fit to the data results in a semi-amplitude of $1.038 \pm 0.800 \text{ m s}^{-1}$ (Fig. 4.24), which is comparable to 0, much less than the expected value of $\sim 20 \text{ m s}^{-1}$, and on the scale of the errors (with an average of 1.76 m s^{-1}). I therefore conclude that this is a flatline, and a likely false positive due to activity.

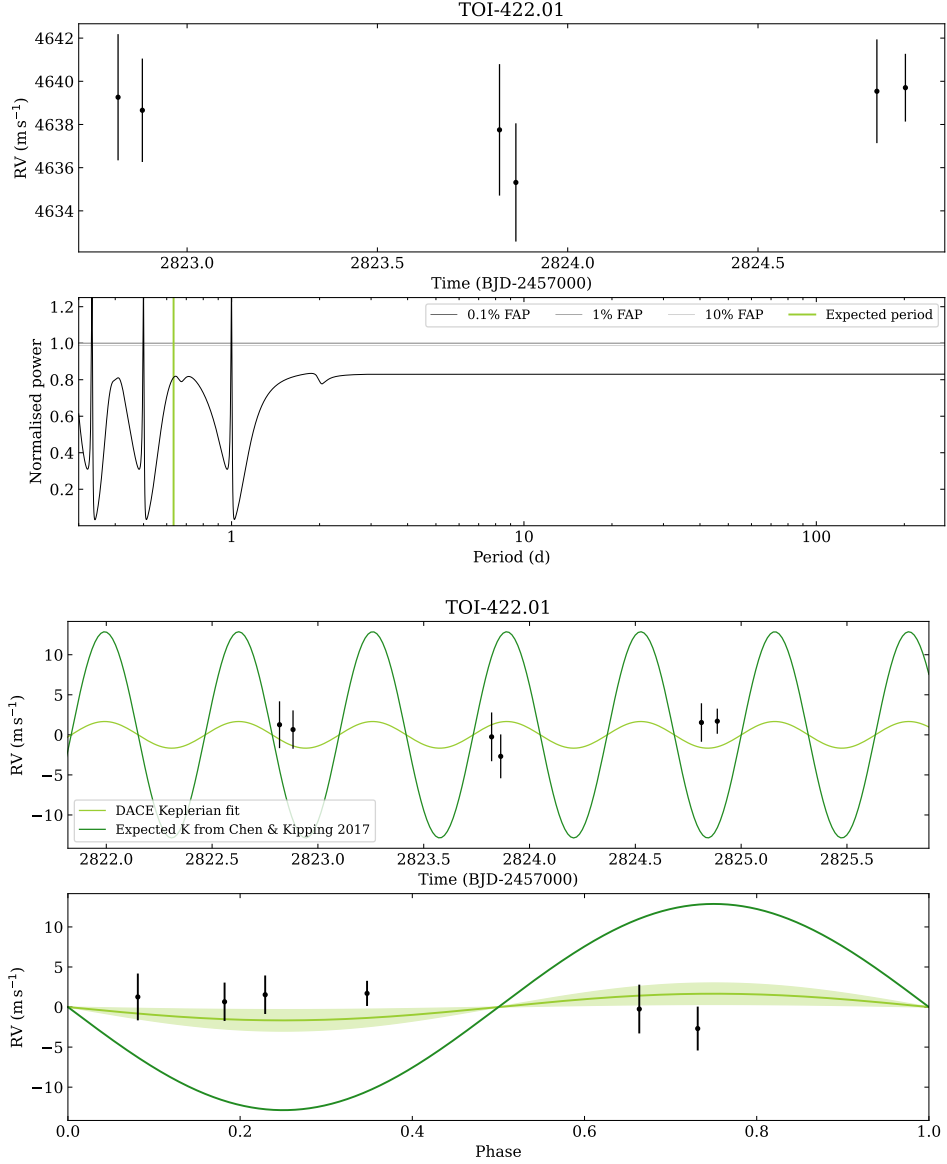


Figure 4.22: TOI-422.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

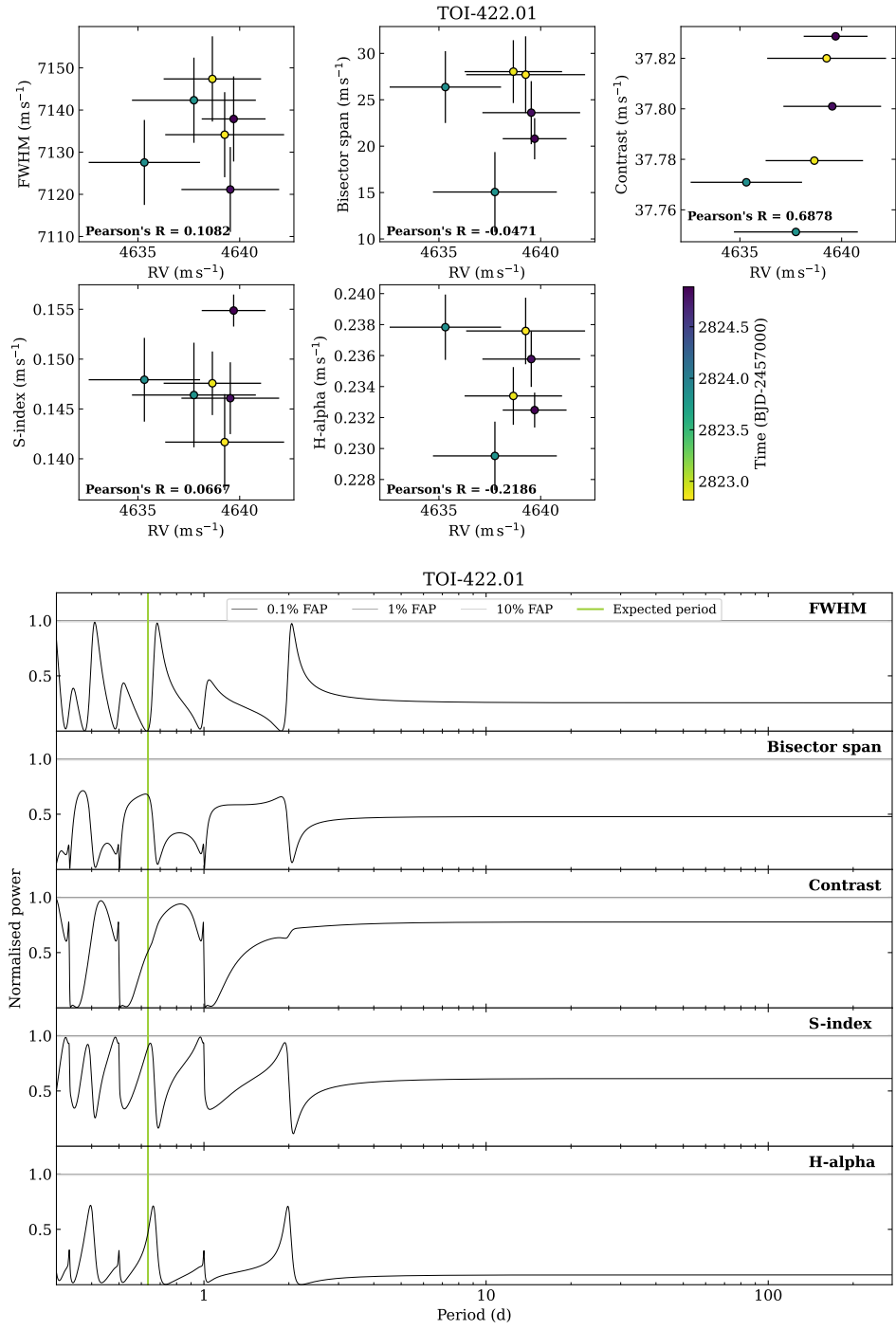


Figure 4.23: TOI-422.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

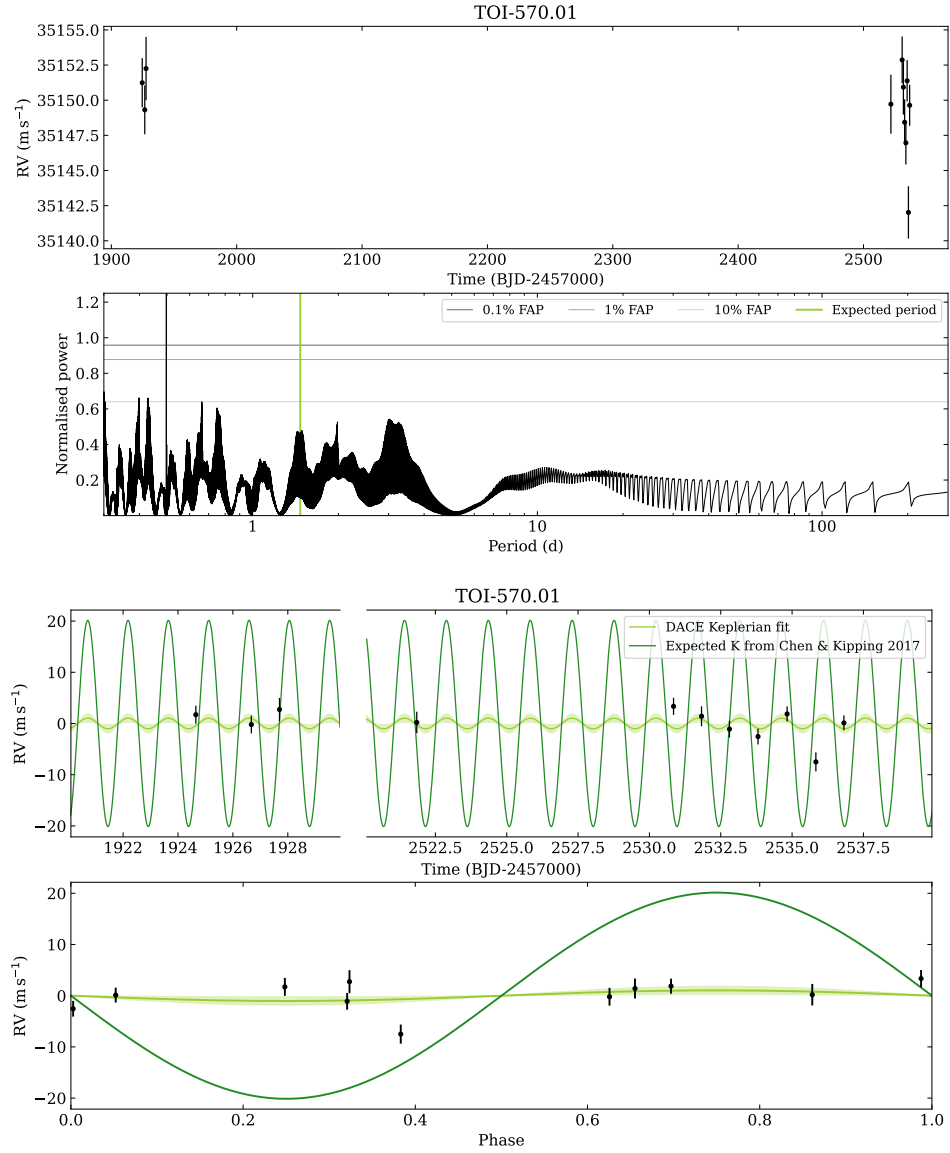


Figure 4.24: TOI-570.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

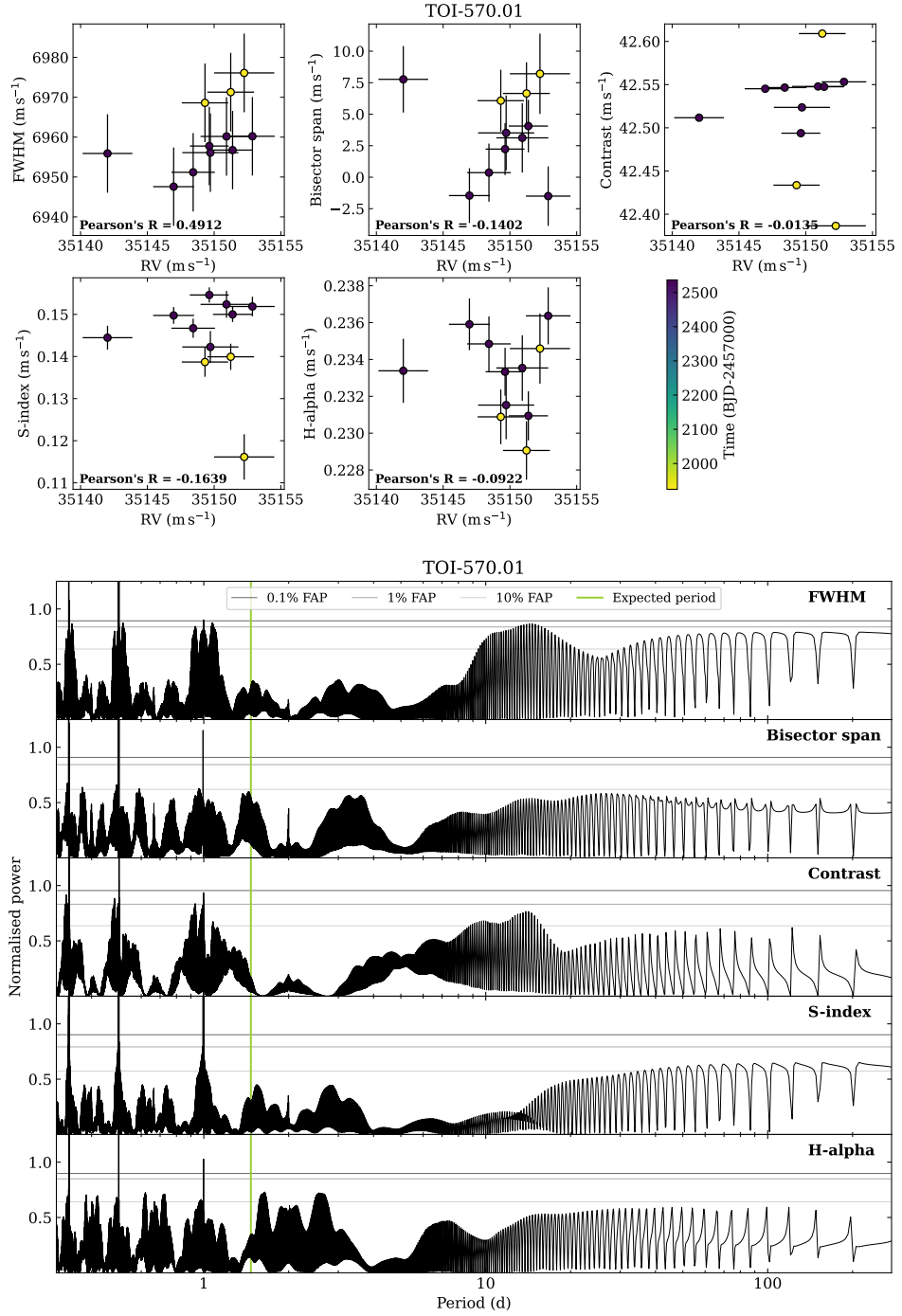


Figure 4.25: TOI-570.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

TOI-855.01

Release values: $P = 1.83$ days; $R_P = 6.66 R_\oplus$; $M_P = 36.08 M_\oplus$; $K = 15.00 \text{ m s}^{-1}$

The CCFs show no abnormalities in their shape (Fig. 4.7), though there is a shift in RV - this is due to drift. Said drift can be seen clearly in Fig. 4.26 and is fit out with a quadratic (with coefficients displayed in the figure). Correlations with activity indicators are examined both before and after the drift is removed: before, there are strong correlations with FWHM and contrast, but after there are none (Fig. 4.27). Looking at the activity indicator periodograms, there is a significant peak in H-alpha at the expected periodicity of the planet, which indicates this is likely a false positive due to activity (Fig. 4.28). Unfortunately, the errors in the RVs are particularly large for this target - an average of 5.05 m s^{-1} , maybe another consequence of activity. There is a slight semi-amplitude, fitting to $3.268 \pm 1.072 \text{ m s}^{-1}$ (Fig. 4.26), but this is still consistent with 0 when considering several standard deviations, less than the expected value of $\sim 15 \text{ m s}^{-1}$, and on the scale of the errors. I therefore conclude this to be a flatline, and a likely false positive due to activity.

TOI-728.01

Release values: $P = 1.60$ days; $R_P = 4.58 R_\oplus$; $M_P = 19.05 M_\oplus$; $K = 10.21 \text{ m s}^{-1}$

The CCFs show no abnormalities (Fig. 4.7). There are strong correlations with FWHM and contrast, and though there are too few data points for the periodograms to be conclusive, peaks are developing around the expected planetary period in FWHM, bisector span, and contrast (Fig. 4.30). This is therefore likely a false positive due to activity. Additionally, the RVs show little variation and so the target was discontinued - the Keplerian fit to the data results in a semi-amplitude of $1.622 \pm 1.664 \text{ m s}^{-1}$ (Fig. 4.29), which is comparable to 0, much less than the expected value of $\sim 10 \text{ m s}^{-1}$, and on the scale of the errors (with an average of 2.61 m s^{-1}). I therefore conclude that this is a flatline, and a likely false positive due to activity.

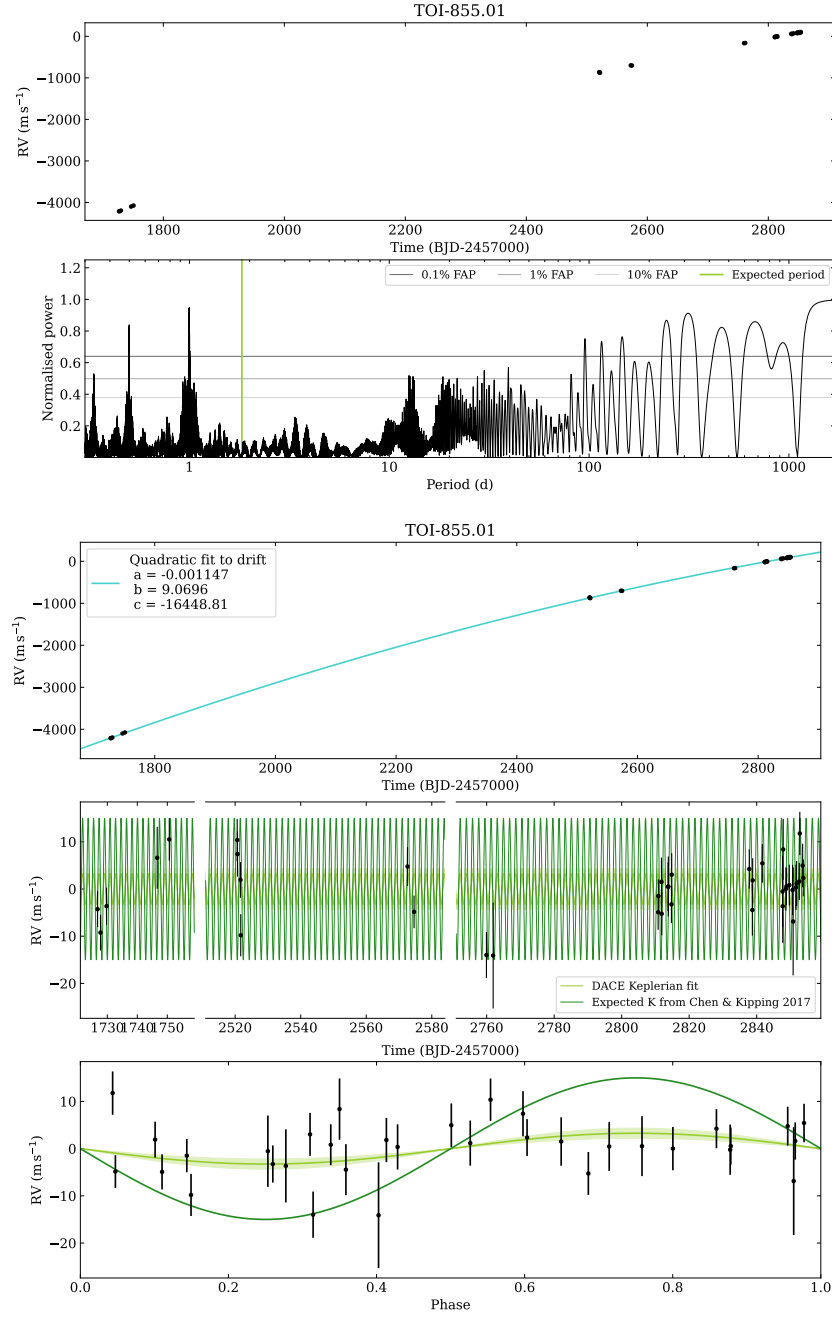


Figure 4.26: TOI-855.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Middle: quadratic fit to the drift, with fit coefficients shown. Bottom middle: a simple Keplerian fit to the RV data after the drift has been removed is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

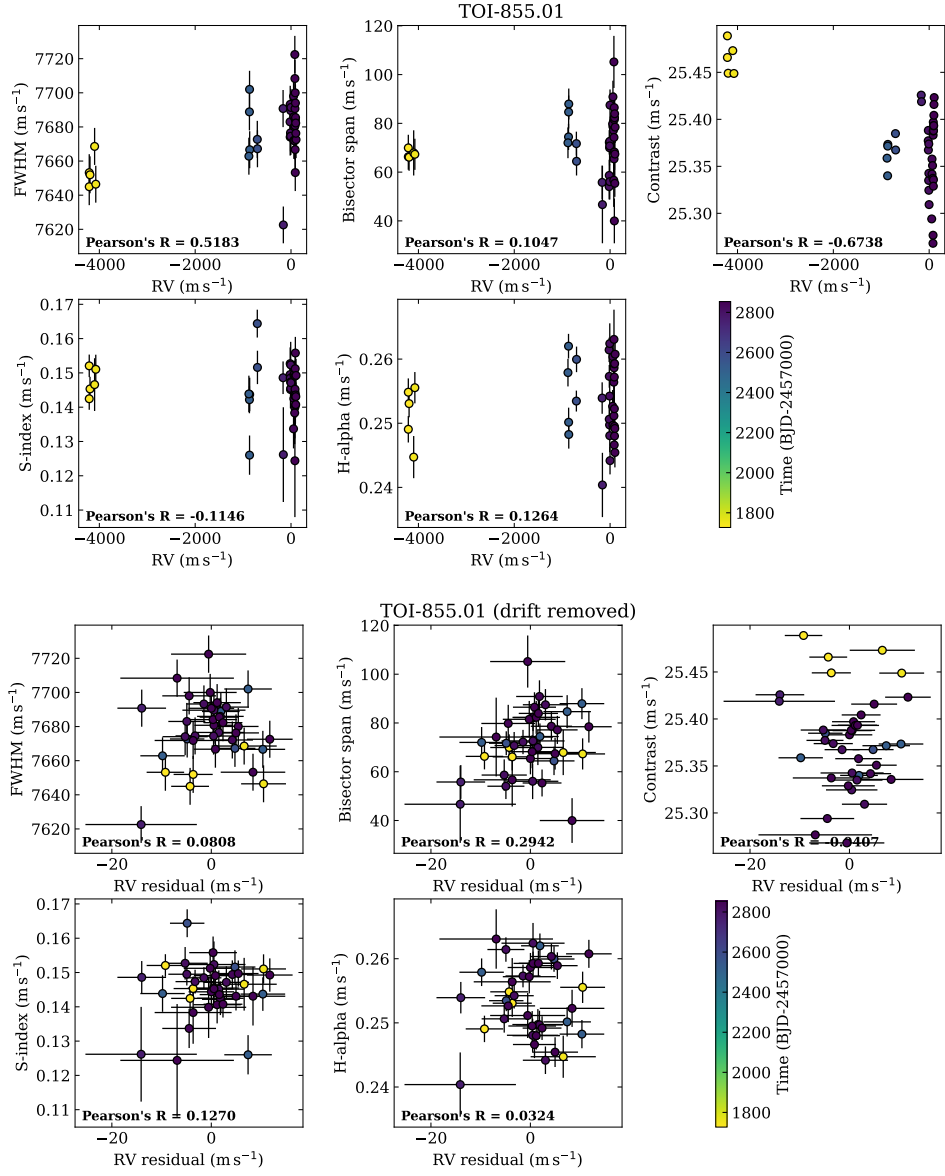


Figure 4.27: TOI-855.01: each activity indicator plot against RV, where the top panels show the raw RV data, and the bottom panels show the RV data after the drift has been subtracted. These show any potential correlation, where Pearson's R is given for each.

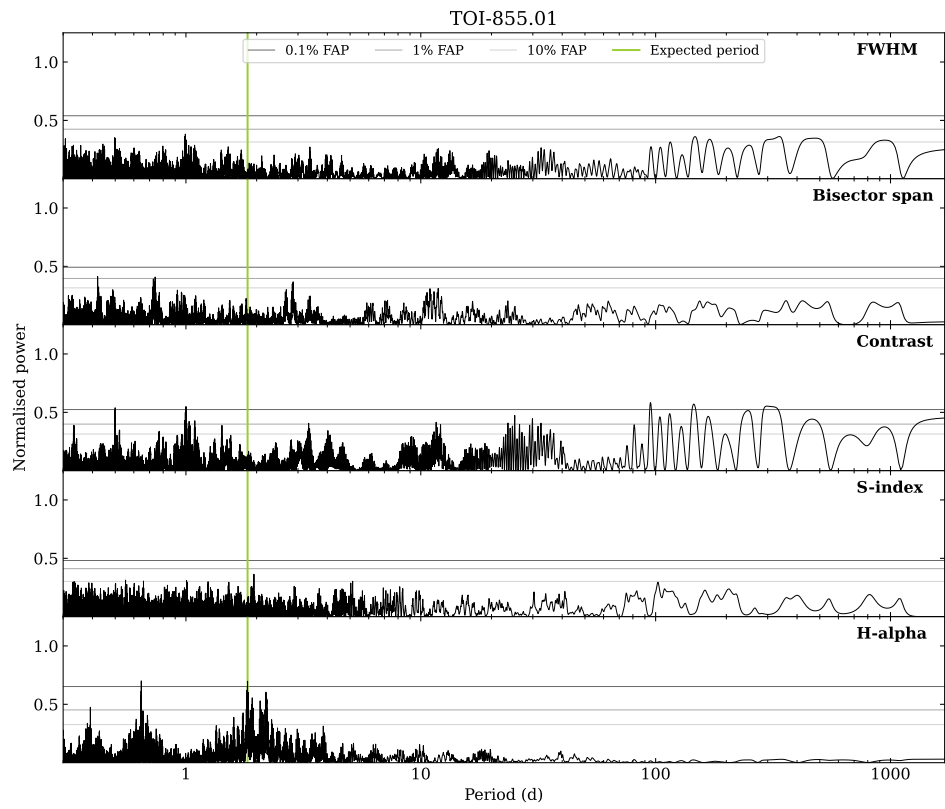


Figure 4.28: TOI-855.01: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

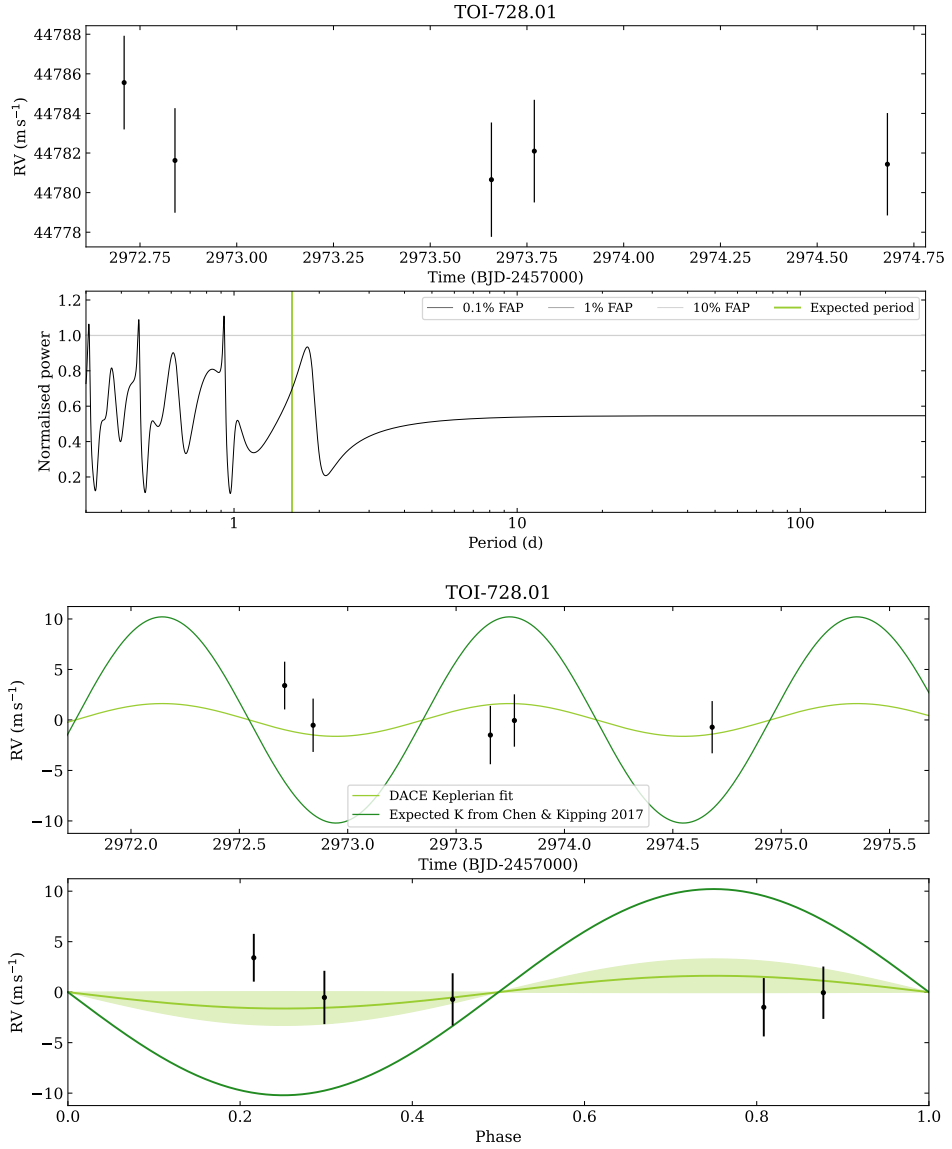


Figure 4.29: TOI-728.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

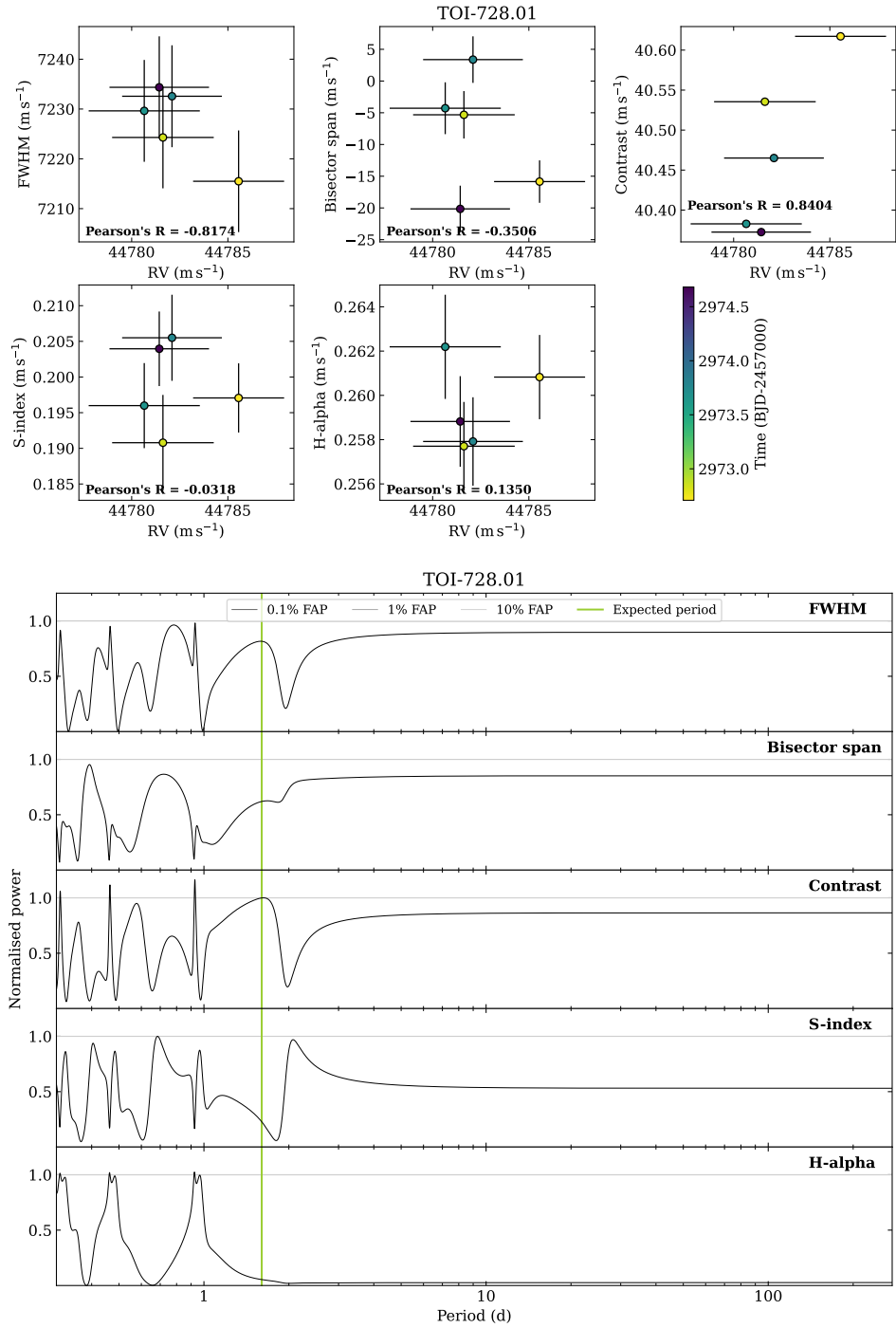


Figure 4.30: TOI-728.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

TOI-355.01

Release values: $P = 1.04$ days; $R_P = 4.60 R_\oplus$; $M_P = 19.26 M_\oplus$; $K = 11.21 \text{ m s}^{-1}$

The CCFs show no abnormalities in their shape (Fig. 4.7), though there is a shift in RV - this is due to drift. Said drift can be seen clearly in Fig. 4.26 and is fit out with a quadratic (with coefficients displayed in the figure). Correlations with activity indicators are examined both before and after the drift is removed (Fig. 4.32): there is a strong correlation with FWHM before, but no strong correlations after. There is a significant peak in the RV periodogram at the expected periodicity (Fig. 4.31), but looking at the activity indicator periodograms, there are developing peaks in FWHM and H-alpha at the expected periodicity of the planet, which indicates this is likely a false positive due to activity (Fig. 4.33). Because the expected period is close to 1 day, it becomes difficult to cover the full phase curve, but with the coverage obtained, the RVs show little variation and so the target was discontinued. The Keplerian fit to the data results in a semi-amplitude of $0.562 \pm 5.452 \text{ m s}^{-1}$ (Fig. 4.31), which is comparable to 0, less than the expected value of $\sim 11 \text{ m s}^{-1}$, and on the scale of the errors (with an average of 2.84 m s^{-1}). I therefore conclude that this is a flatline, and a likely false positive due to activity.

TOI-1975.01

Release values: $P = 2.83$ days; $R_P = 4.27 R_\oplus$; $M_P = 16.93 M_\oplus$; $K = 7.81 \text{ m s}^{-1}$

The CCFs show no abnormalities (Fig. 4.7), aside from one observation taken on the night of 12 July 2022. This was due to a power cut at La Silla, causing issues with the telescope pointing for rest of the night. This observation is therefore discarded before analysis. There are very strong correlations with FWHM, contrast, and S-index, and while a small peak shows in the RV periodogram at the expected period, this likely corresponds to peaks in the FWHM, bisector span and contrast. Unfortunately, the errors in the RVs are particularly large for this target - an average of 4.58 m s^{-1} , maybe another consequence of activity. The RVs show little variation and so due to this and the activity correlations the target was discontinued - the Keplerian fit to the data results in a semi-amplitude of $0.441 \pm 0.762 \text{ m s}^{-1}$ (Fig. 4.34), which is comparable to 0, much less than the expected value of $\sim 8 \text{ m s}^{-1}$, and on the scale of the errors. I therefore conclude that this is a flatline, and a likely false positive due to activity.

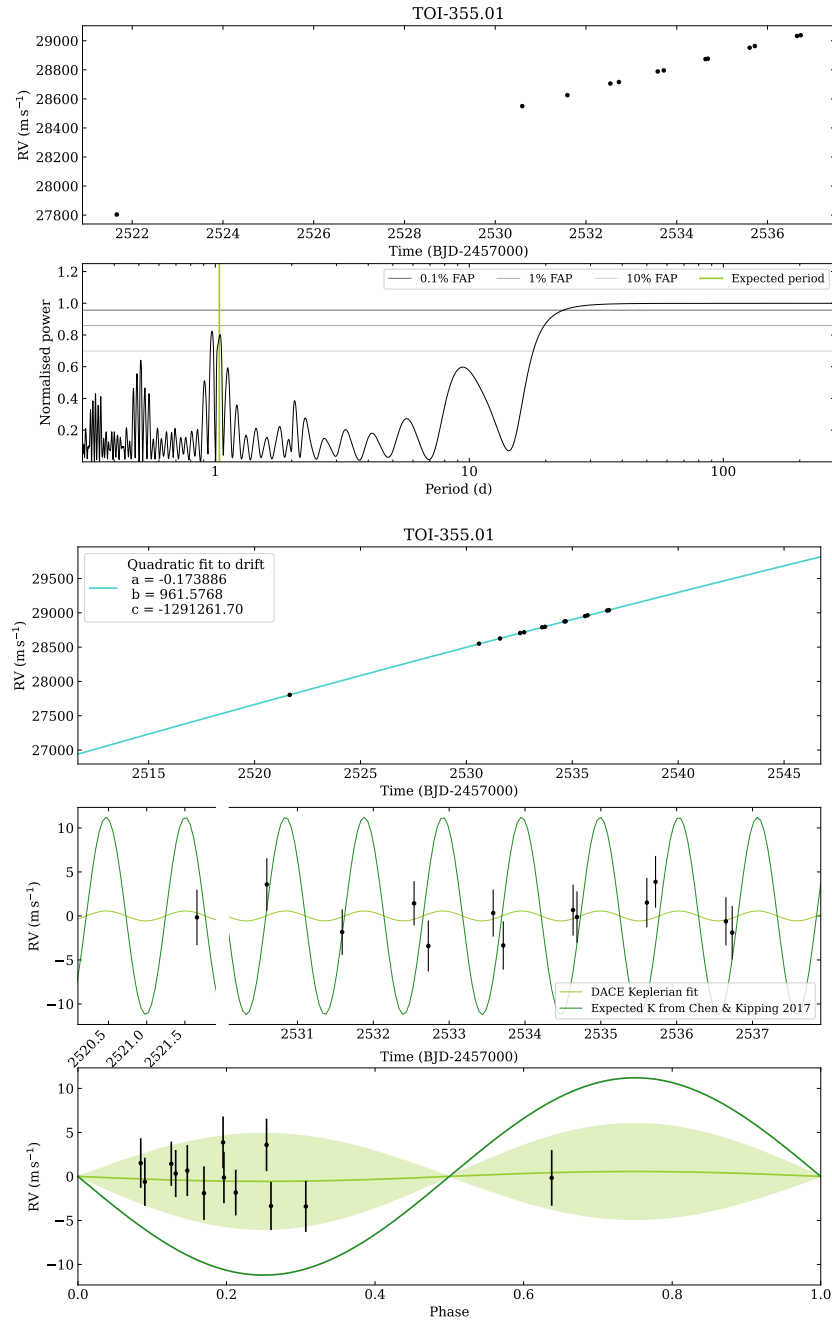


Figure 4.31: TOI-355.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Middle: quadratic fit to the drift, with fit coefficients shown. Bottom middle: a simple Keplerian fit to the RV data after the drift has been removed is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

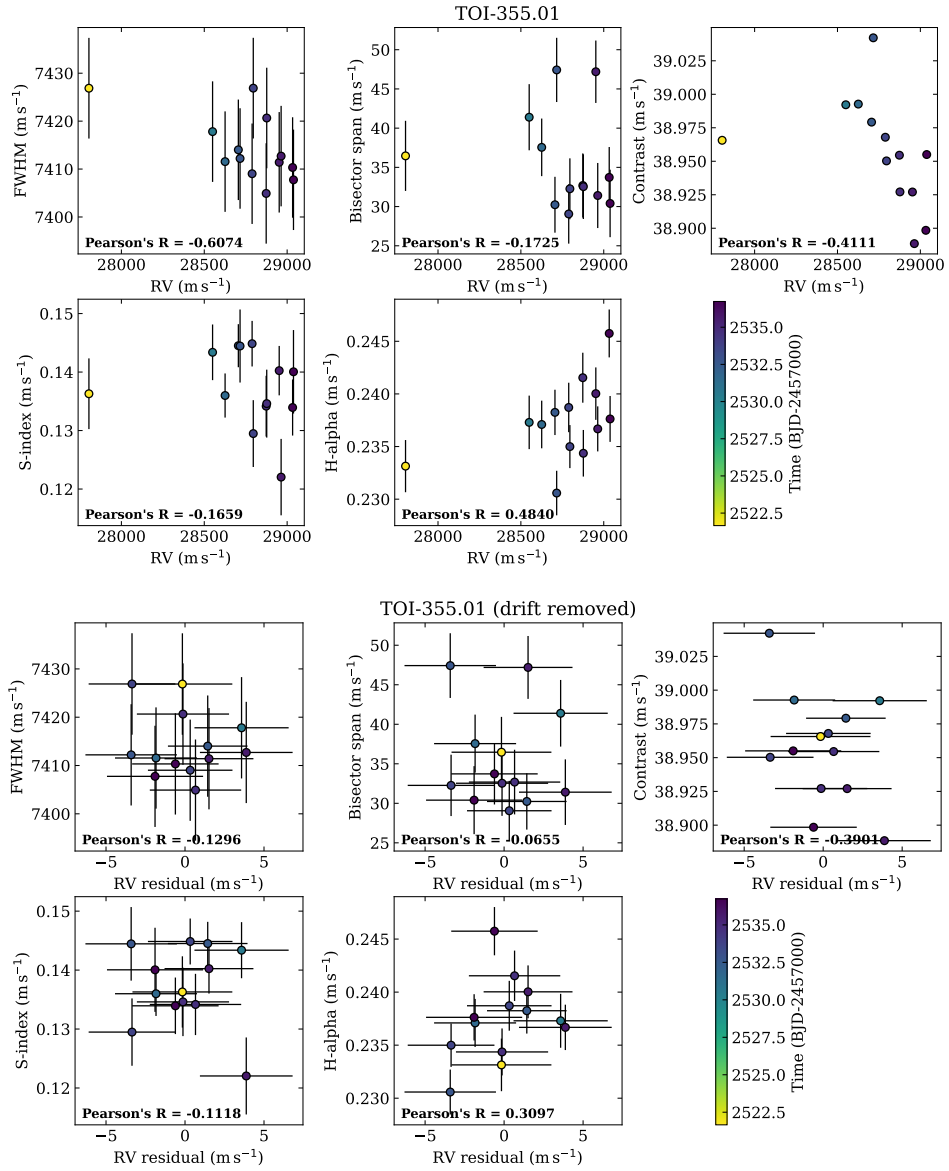


Figure 4.32: TOI-355.01: each activity indicator plot against RV, where the top panels show the raw RV data, and the bottom panels show the RV data after the drift has been subtracted. These show any potential correlation, where Pearson's R is given for each.

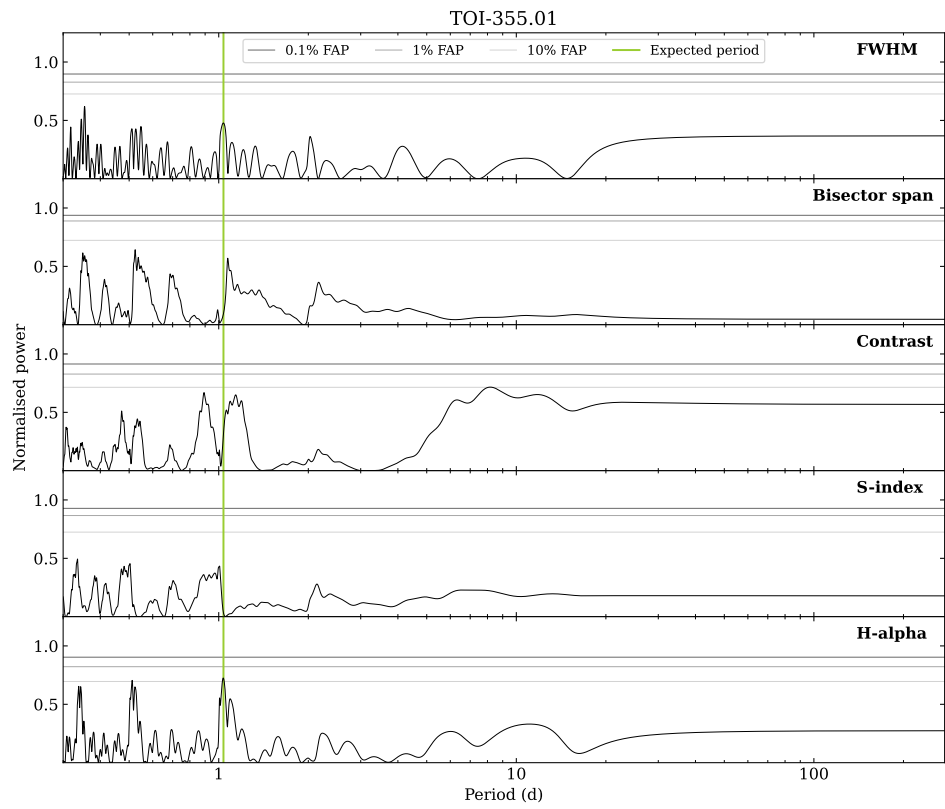


Figure 4.33: TOI-355.01: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

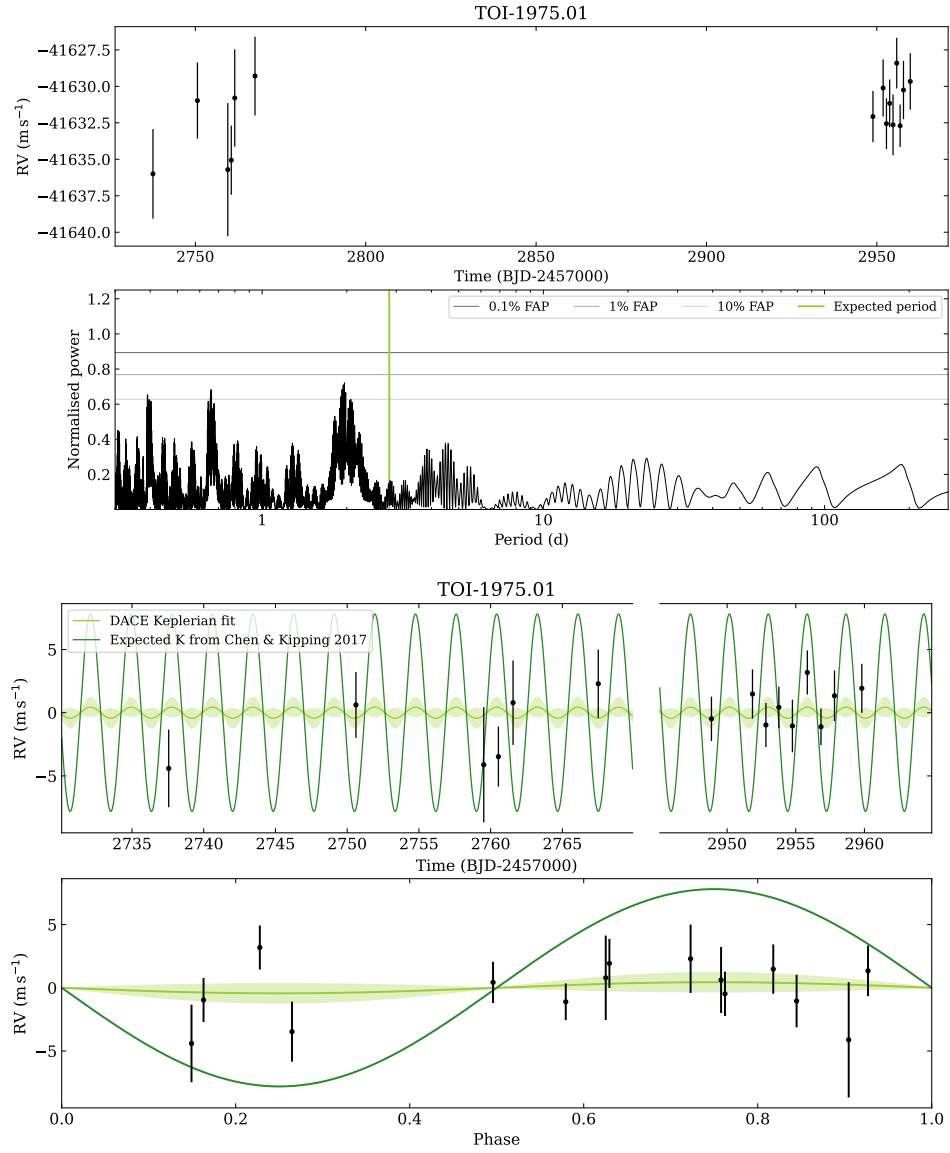


Figure 4.34: TOI-1975.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

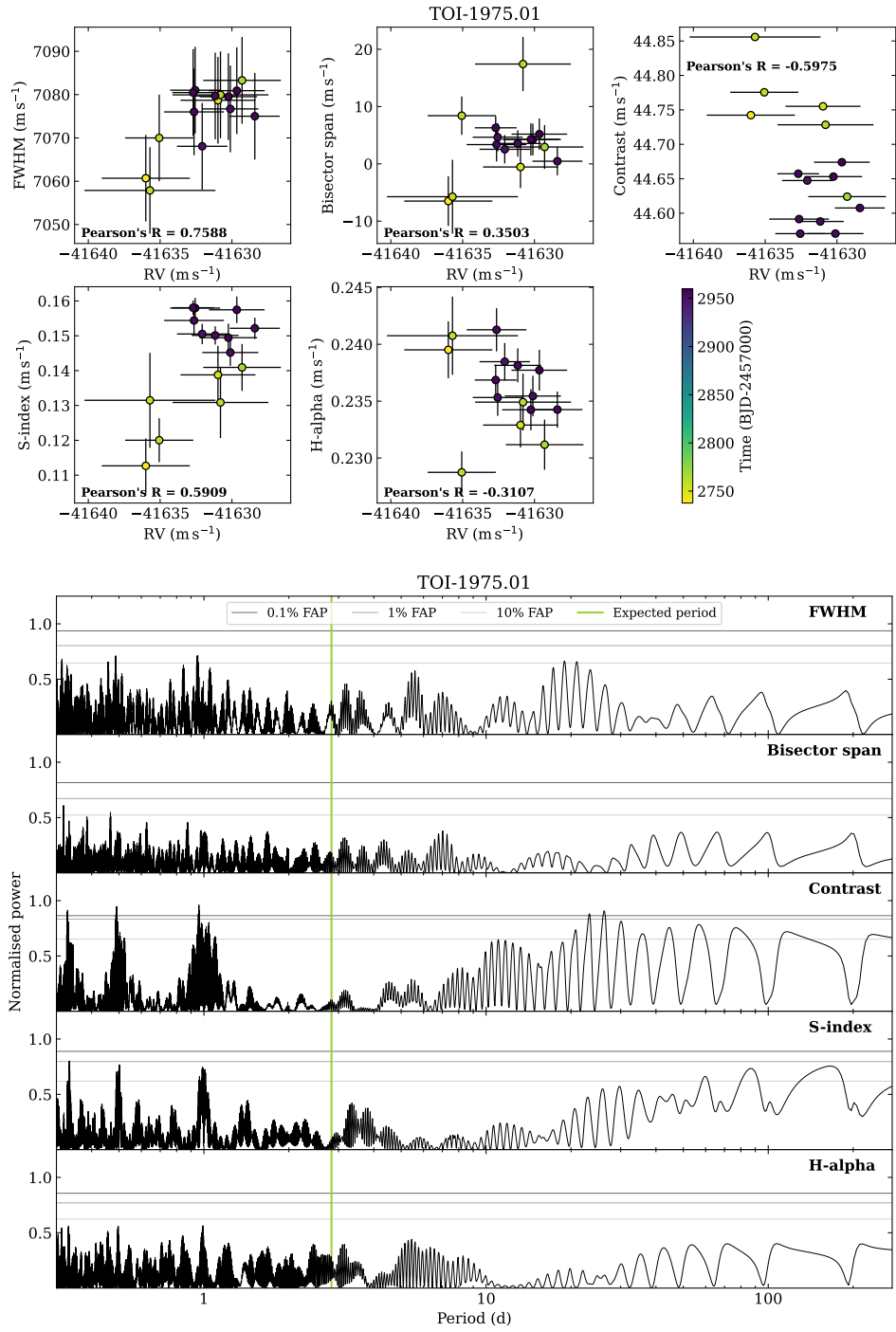


Figure 4.35: TOI-1975.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

TOI-4461.01

Release values: $P = 3.70$ days; $R_P = 4.49 R_\oplus$; $M_P = 18.48 M_\oplus$; $K = 8.61 \text{ m s}^{-1}$

It is important to note that, for this target alone, there is currently only data reduced by the old HARPS DRS 3.0.0 pipeline available, rather than the newer DRS 3.5. I use that in this analysis, but any further analysis should try and re-reduce the raw data using DRS 3.5.

There are 3 observations that have abnormally large errors; these are on the nights of 18, 19, and 25 May 2023. This is likely caused by bad weather conditions and these data points are discarded before analysis.

The CCFs show no abnormalities (Fig. 4.7). There are no strong correlations with the activity indicators. There are no peaks in the activity indicator periodograms at the expected period of the planet, and no significant peaks overall, though there are peaks forming at and beyond 10 days indicative of an activity cycle. This is reflected in a ~ 10.3 day peak present also in the RV periodogram, which also does not show a significant peak at the expected planet period. The Keplerian fit to the data results in a semi-amplitude of $0.851 \pm 1.053 \text{ m s}^{-1}$ (Fig. 4.36), which is comparable to 0, much less than the expected value of $\sim 8.6 \text{ m s}^{-1}$, and on the scale of the errors (with an average of 6.11 m s^{-1}).

This is marginally improved by also adding a Keplerian to simultaneously fit the likely activity period at ~ 10.3 days; this boosts the semi-amplitude of the planet fit to $1.801 \pm 1.138 \text{ m s}^{-1}$, but this is still the same overall result as previous. I conclude that this is a flatline, though the existence of a planet is not precluded and an upper mass limit could be obtained, though the activity would also need to be simultaneously detrended. More precise RVs would be needed to obtain a precise mass measurement.

TOI-4537.01

Release values: $P = 6.66$ days; $R_P = 3.88 R_\oplus$; $M_P = 14.42 M_\oplus$; $K = 4.60 \text{ m s}^{-1}$

The CCFs show no abnormalities (Fig. 4.7). There are no correlations with the activity indicators. The periodograms are not necessarily helpful due to the handful of data points obtained, but there is no peak in the RV periodogram at the expected period (Fig. 4.39). There are also no significant peaks in the activity indicator periodograms. While the variation in the RV points is on the same scale as the expected small semi-amplitude of $\sim 4.6 \text{ m s}^{-1}$, there are several points that are definitely out of phase with the expected curve, and the size of the errors (with an average of 2.15 m s^{-1}) are comparable with the expected semi-amplitude (Fig. 4.38). The semi-amplitude fits to $0.734 \pm 1.216 \text{ m s}^{-1}$. I conclude that this is a flatline, though the existence of a planet is not precluded and an upper mass limit could be obtained. More precise RVs would be needed to obtain a precise mass measurement.

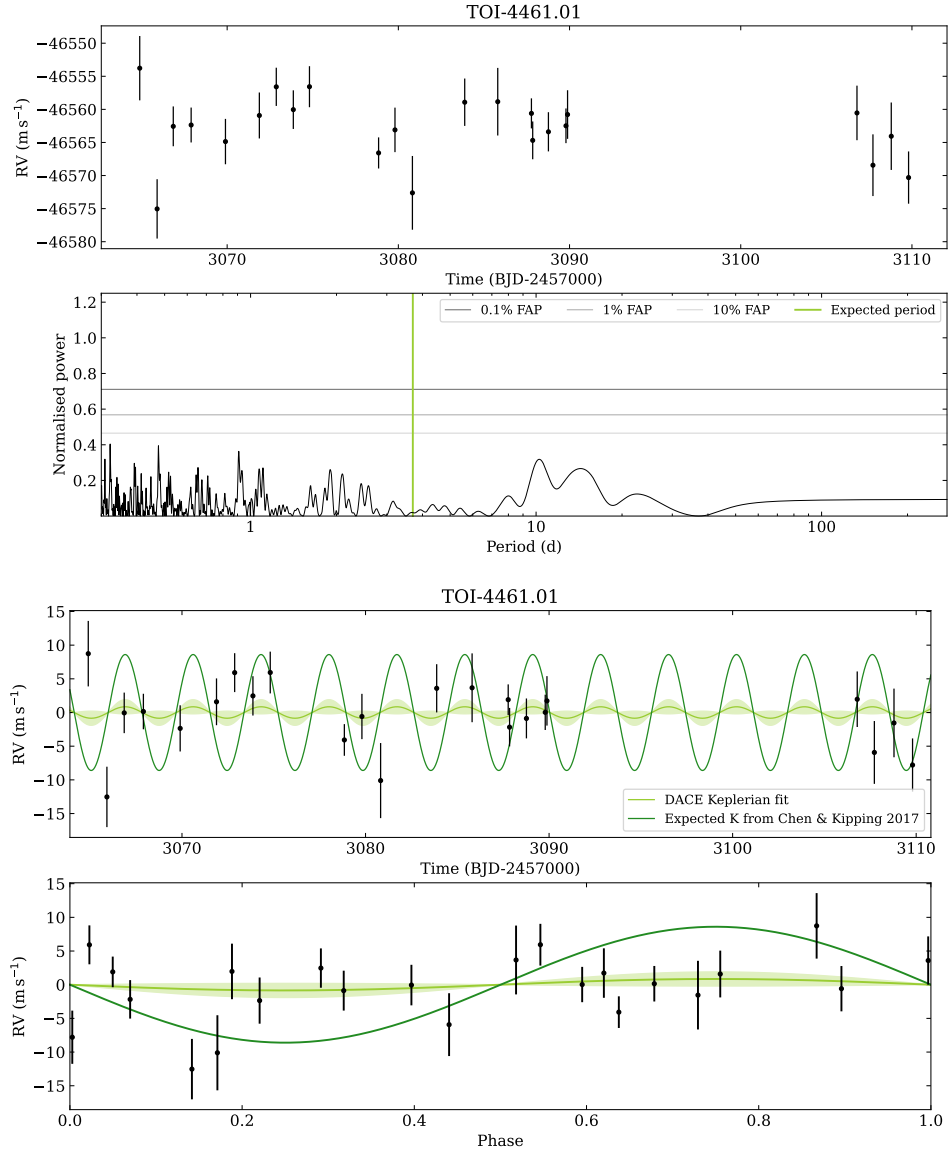


Figure 4.36: TOI-4461.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

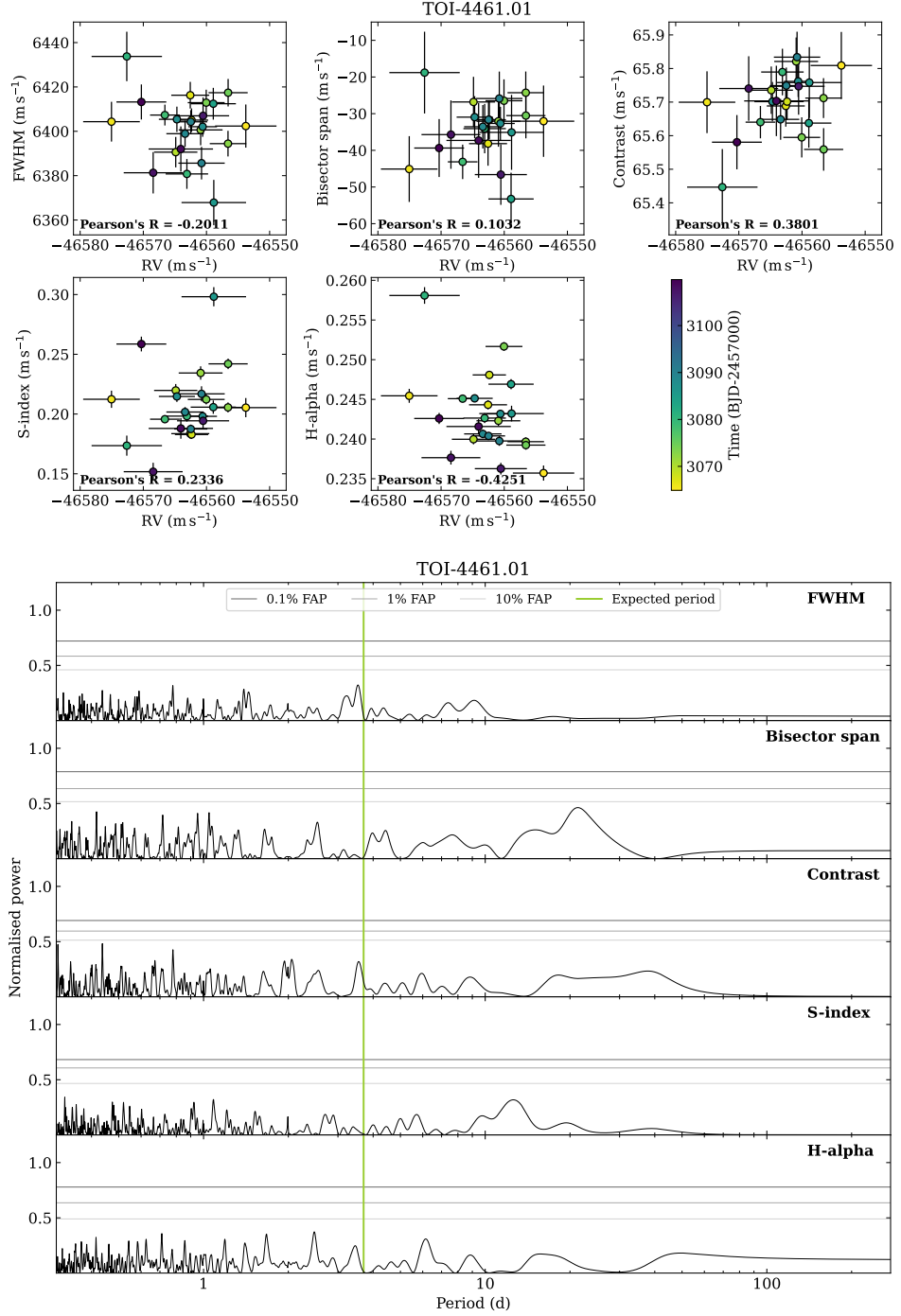


Figure 4.37: TOI-4461.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

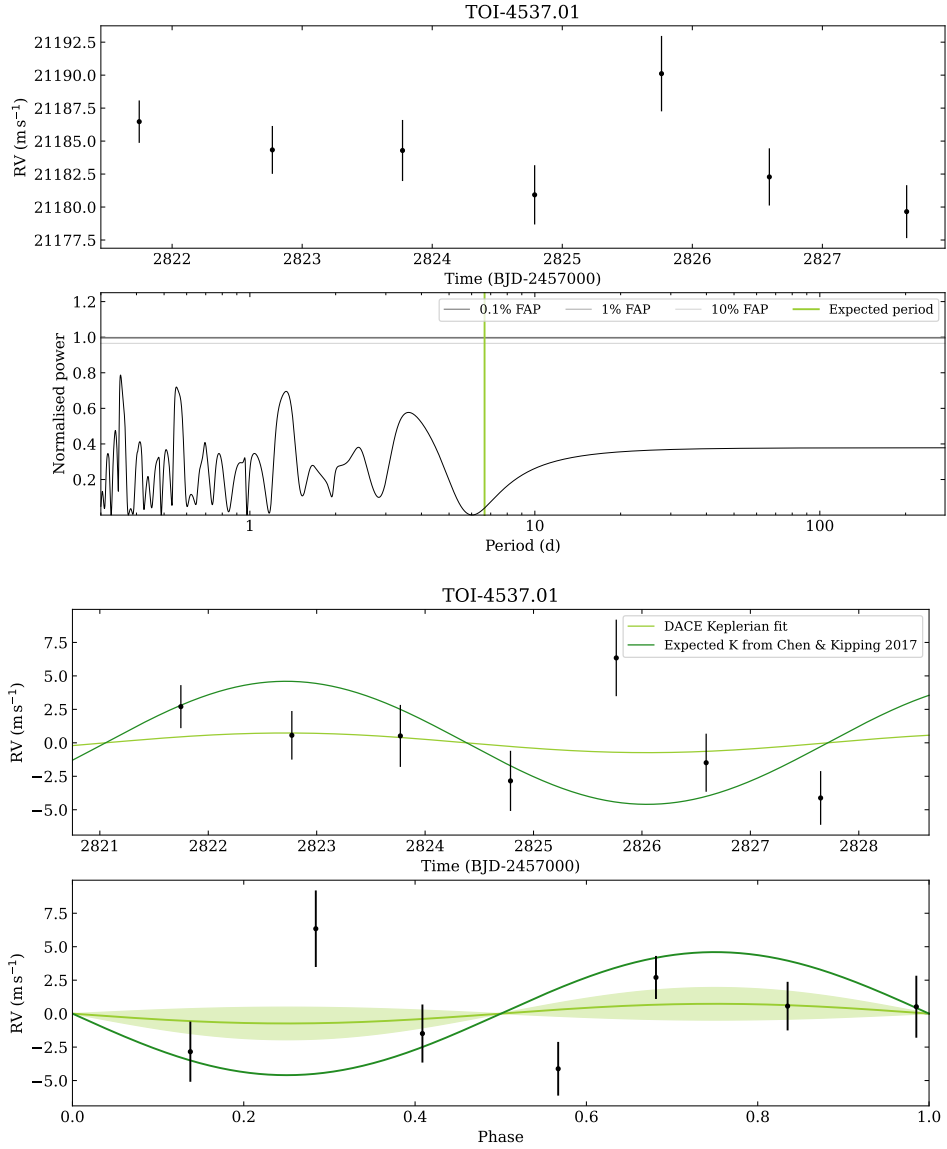


Figure 4.38: TOI-4537.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

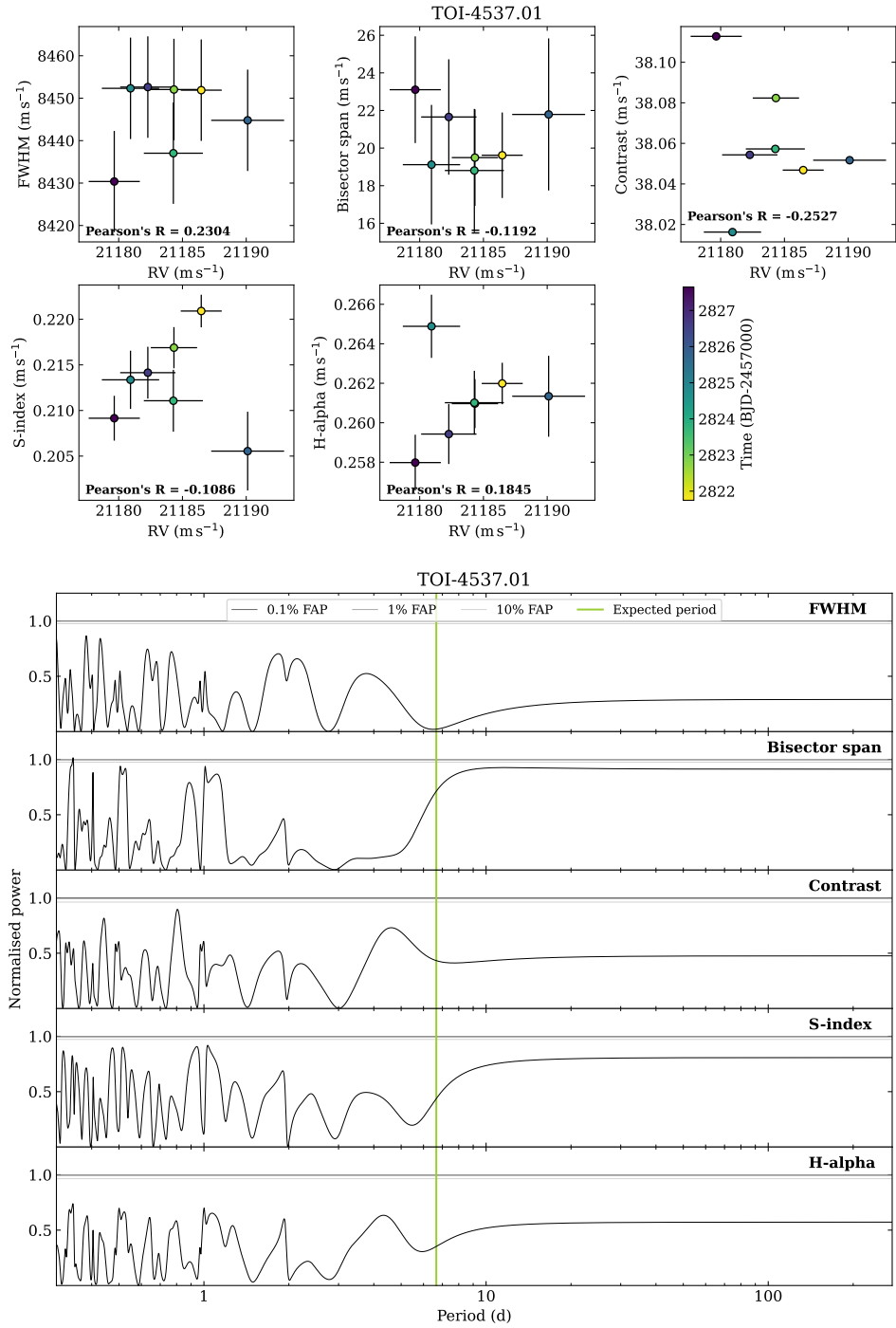


Figure 4.39: TOI-4537.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

TOI-2365.01

Release values: $P = 0.66$ days; $R_P = 6.08 R_\oplus$; $M_P = 30.92 M_\oplus$; $K = 21.98 \text{ m s}^{-1}$

The CCFs show no abnormalities (Fig. 4.7). There is a slight correlation with FWHM, but no correlation with the other activity indicators. There are no significant peaks in the activity indicator periodograms, and definitely none at the expected planet period (Fig. 4.41). There is also no peak in the RV periodogram at the expected period. The RVs show little variation and so the target was discontinued - the Keplerian fit to the data results in a semi-amplitude of $0.209 \pm 2.394 \text{ m s}^{-1}$ (Fig. 4.40), which is comparable to 0, much less than the expected value of $\sim 22 \text{ m s}^{-1}$, and on the scale of the errors (with a large average of 6.20 m s^{-1}). I therefore conclude that this is a flatline, though the existence of a planet is not precluded and an upper mass limit could be obtained. More precise RVs would be needed to obtain a precise mass measurement.

TOI-502.01

Release values: $P = 2.94$ days; $R_P = 8.83 R_\oplus$; $M_P = 58.29 M_\oplus$; $K = 24.30 \text{ m s}^{-1}$

The CCFs show no abnormalities in shape (Fig. 4.7), though there is a shift in RV due to drift. Said drift can be seen clearly in Fig. 4.42 and is fit out with a quadratic (with coefficients displayed in the figure). Correlations with activity indicators are examined before and after the drift is removed (Fig. 4.43): before, there are correlations with FWHM and H-alpha; after, there is still a strong correlation with FWHM. While there is a small peak in the RV periodogram at the expected periodicity, this is reflected as a peak in every activity indicator (Fig. 4.44), making a strong case that this is a false positive due to activity. The RVs show little variation and so the target was discontinued - the Keplerian fit to the data results in a semi-amplitude of $0.667 \pm 2.299 \text{ m s}^{-1}$ (Fig. 4.42), which is comparable to 0, much less than the expected value of $\sim 24 \text{ m s}^{-1}$, and on the scale of the errors (with an average of 3.08 m s^{-1}). I conclude that this is a flatline, and a likely false positive due to activity.

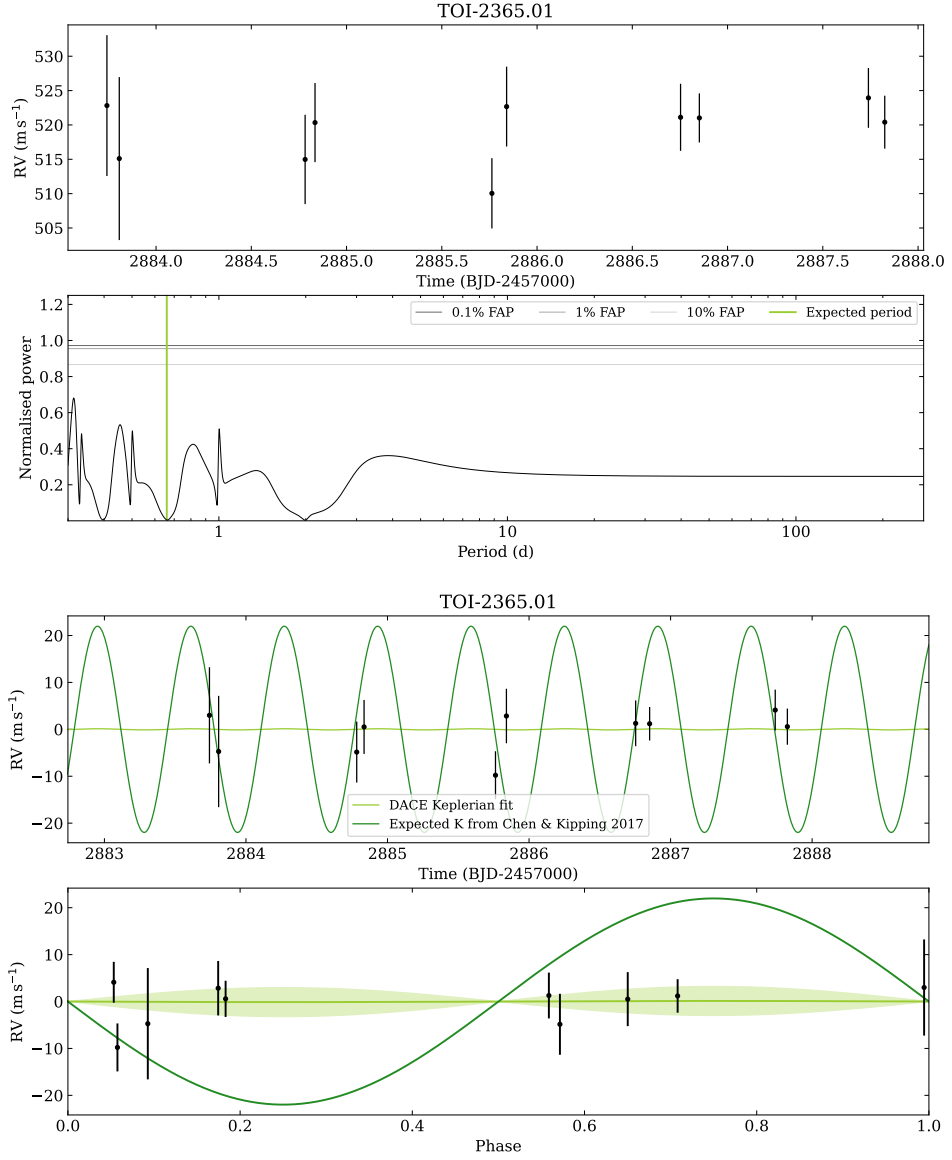


Figure 4.40: TOI-2365.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

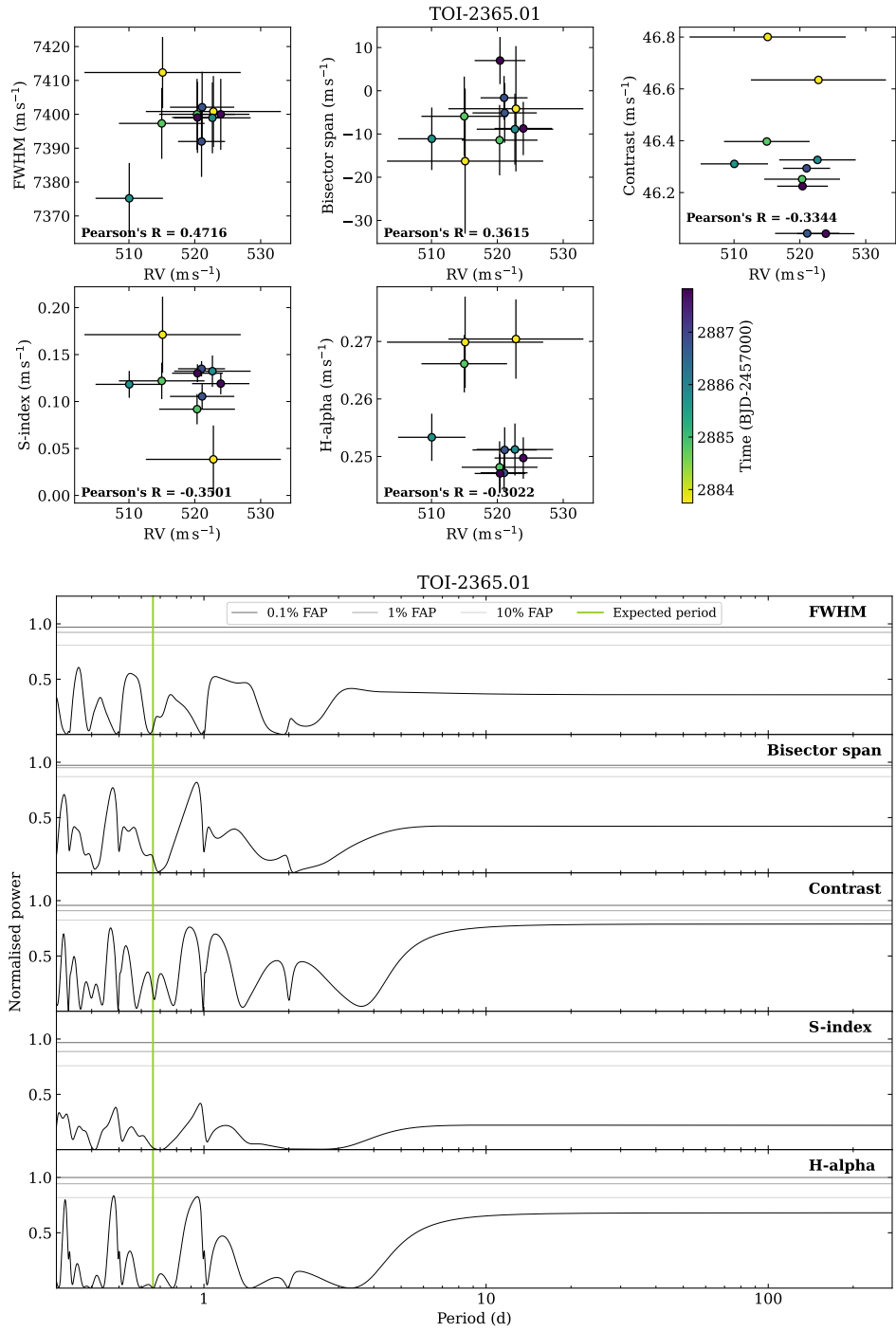


Figure 4.41: TOI-2365.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

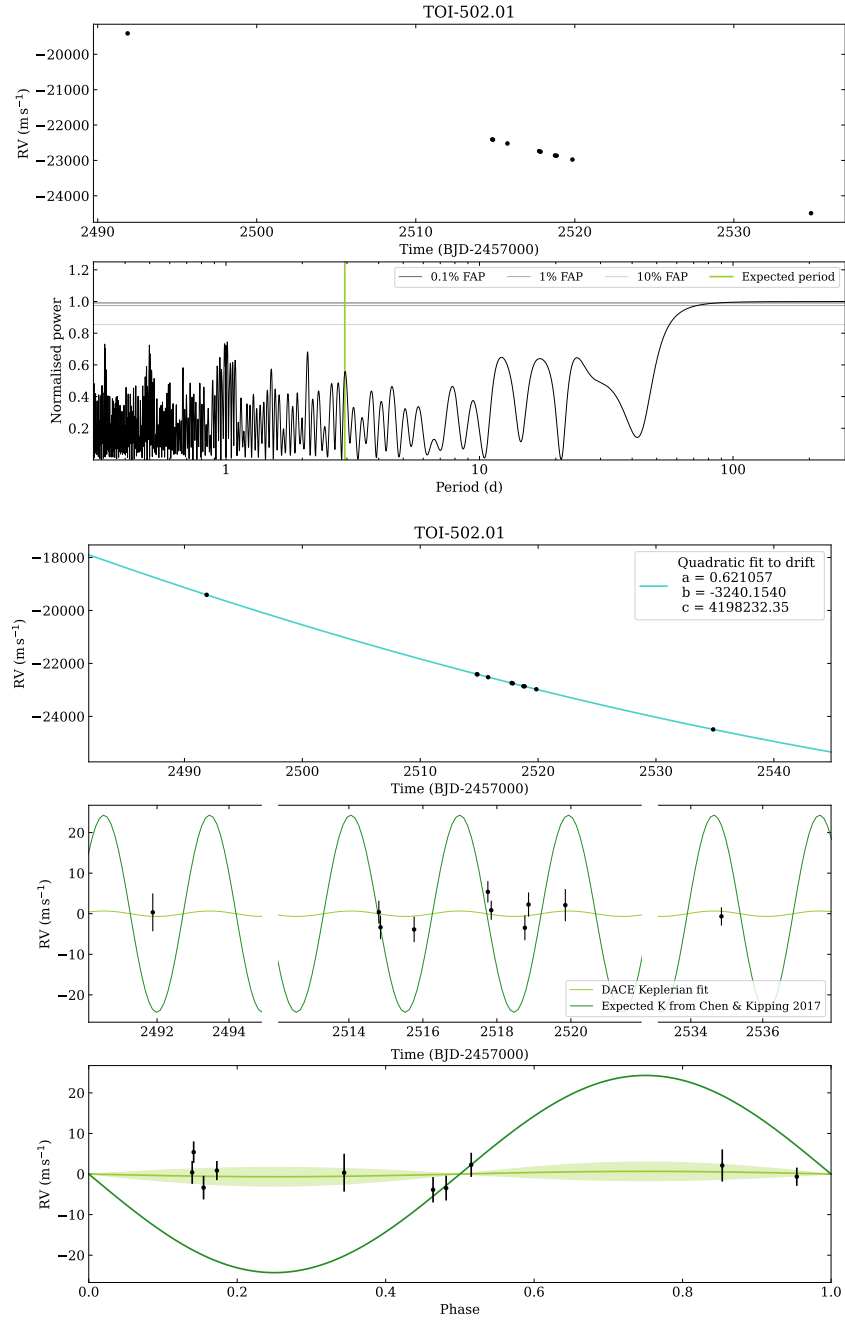


Figure 4.42: TOI-502.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Middle: quadratic fit to the drift, with fit coefficients shown. Bottom middle: a simple Keplerian fit to the RV data after the drift has been removed is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

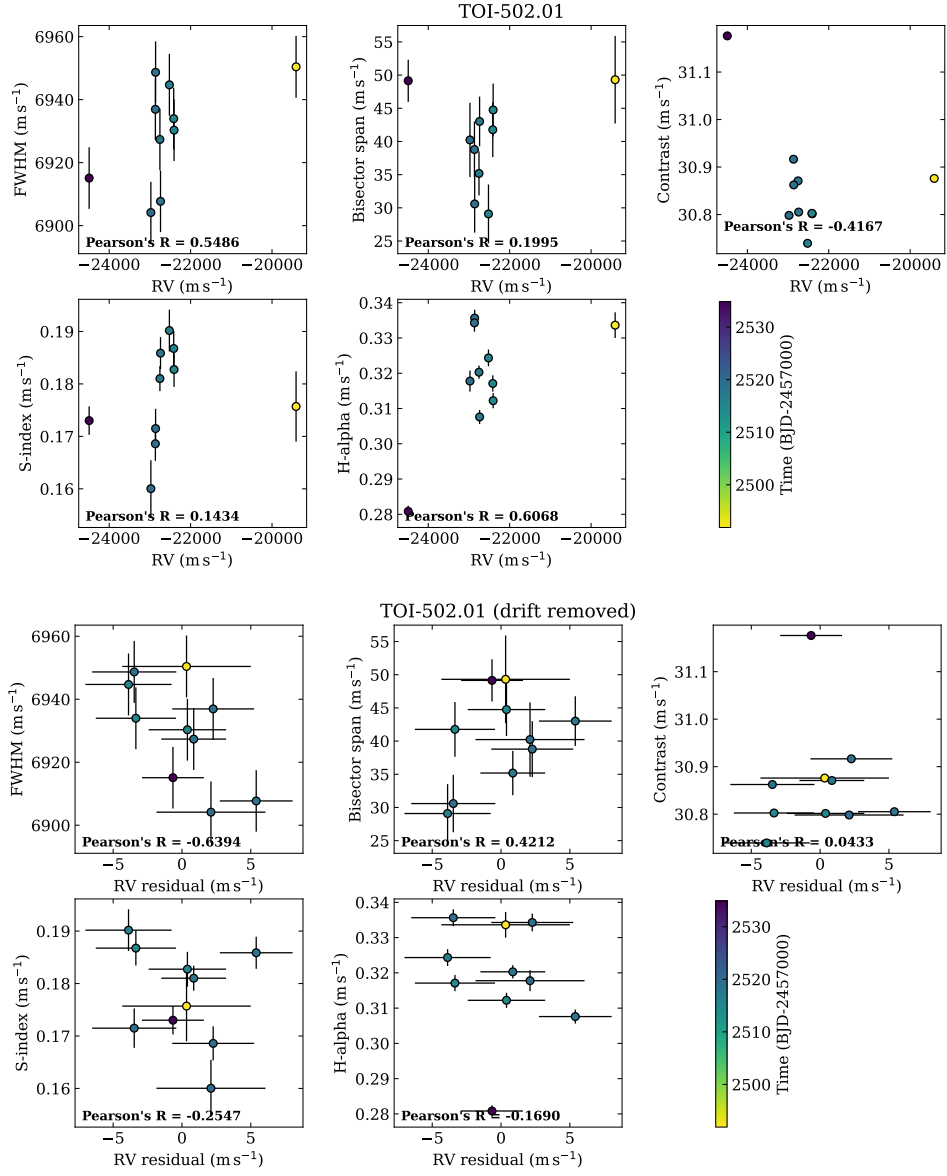


Figure 4.43: TOI-502.01: each activity indicator plot against RV, where the top panels show the raw RV data, and the bottom panels show the RV data after the drift has been subtracted. These show any potential correlation, where Pearson's R is given for each.

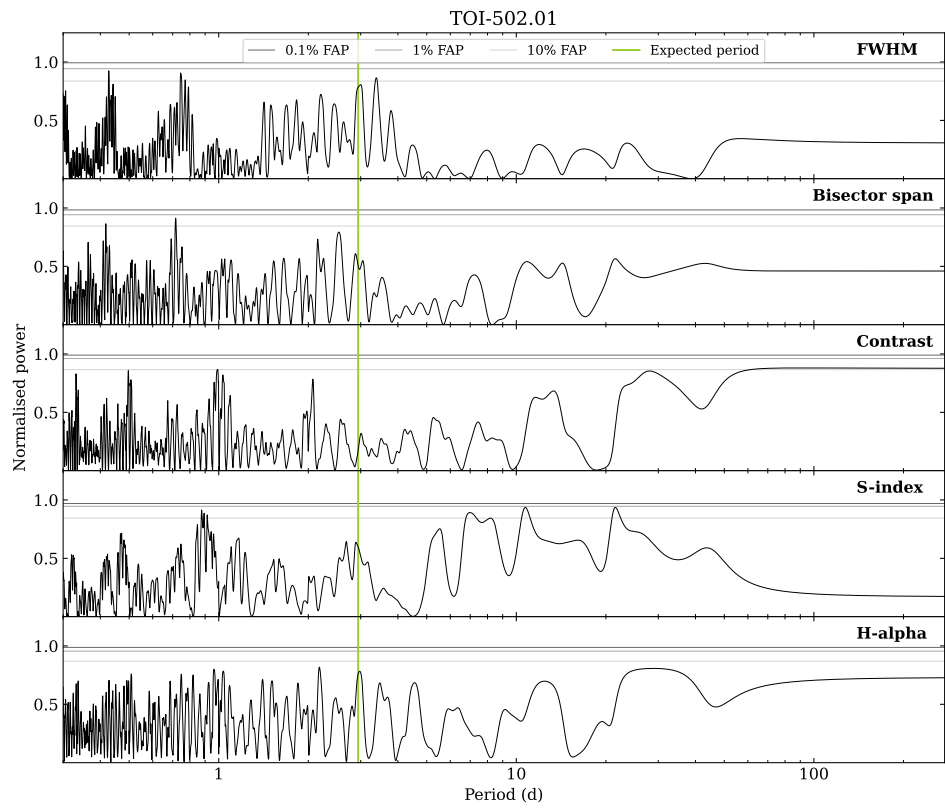


Figure 4.44: TOI-502.01: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

TOI-835.01

Release values: $P = 4.79$ days; $R_P = 4.20 R_\oplus$; $M_P = 16.46 M_\oplus$; $K = 6.24 \text{ m s}^{-1}$

The CCFs show no abnormalities (Fig. 4.7). There are strong correlations with every activity indicator (Fig. 4.46, and small peaks in the FWHM and S-index periodograms at the expected periodicity. However, there are also significant peaks in FWHM, bisector span, and S-index at longer periods, which also show as a significant peak at ~ 9.3 days in the RV periodogram (Fig. 4.45). Indeed, fitting a simple Keplerian at this period explains the large scatter (much larger than the expected semi-amplitude of $\sim 6.24 \text{ m s}^{-1}$) in the data – it is caused by the stellar activity. So while this is not a flatline at first glance, removing the activity leaves a flatline with no periodicity. I conclude that this is a likely false positive due to activity.

TOI-499.01

Release values: $P = 8.53$ days; $R_P = 6.59 R_\oplus$; $M_P = 35.39 M_\oplus$; $K = 9.75 \text{ m s}^{-1}$

The CCFs show no abnormalities (Fig. 4.7). There is a strong correlation with bisector span (Fig. 4.48), and a veritable forest of peaks showing up in the activity indicator periodograms, with some clustered around/near the expected periodicity in FWHM, bisector span and contrast. Some of the RVs are decidedly out of phase with the expected curve (Fig. 4.47), though the scatter is on a similar scale to the expected semi-amplitude of $\sim 10 \text{ m s}^{-1}$. This combination of factors led to the target being discontinued. I conclude that this is a likely false positive due to activity.

TOI-2227.01

Release values: $P = 4.22$ days; $R_P = 3.01 R_\oplus$; $M_P = 9.34 M_\oplus$; $K = 3.62 \text{ m s}^{-1}$

I have now reached the last target in the sample. The CCFs show no abnormalities (Fig. 4.7). There is a correlation with S-index and a slight correlation with H-alpha. There are no significant peaks in the activity indicator periodograms, though lots of small peaks cluster around the expected planetary periodicity, but not directly at the expected value (Fig. 4.50). Indeed, these peaks show in the RV periodogram, but again there is no peak directly at the expected value (Fig. 4.49). The scatter of the data points is comparable to the expected semi-amplitude of $\sim 9 \text{ m s}^{-1}$, but some of the points are decidedly out of phase with the expected curve. The Keplerian fit to the data results in a semi-amplitude of $0.118 \pm 1.573 \text{ m s}^{-1}$ (Fig. 4.42), which is comparable to 0, and on the scale of the errors (with an average of 2.67 m s^{-1}). I therefore conclude that this is a flatline, though the existence of a planet is not precluded and an upper mass limit could be obtained. More precise RVs would be needed to obtain a precise mass measurement if a planet were to be present.

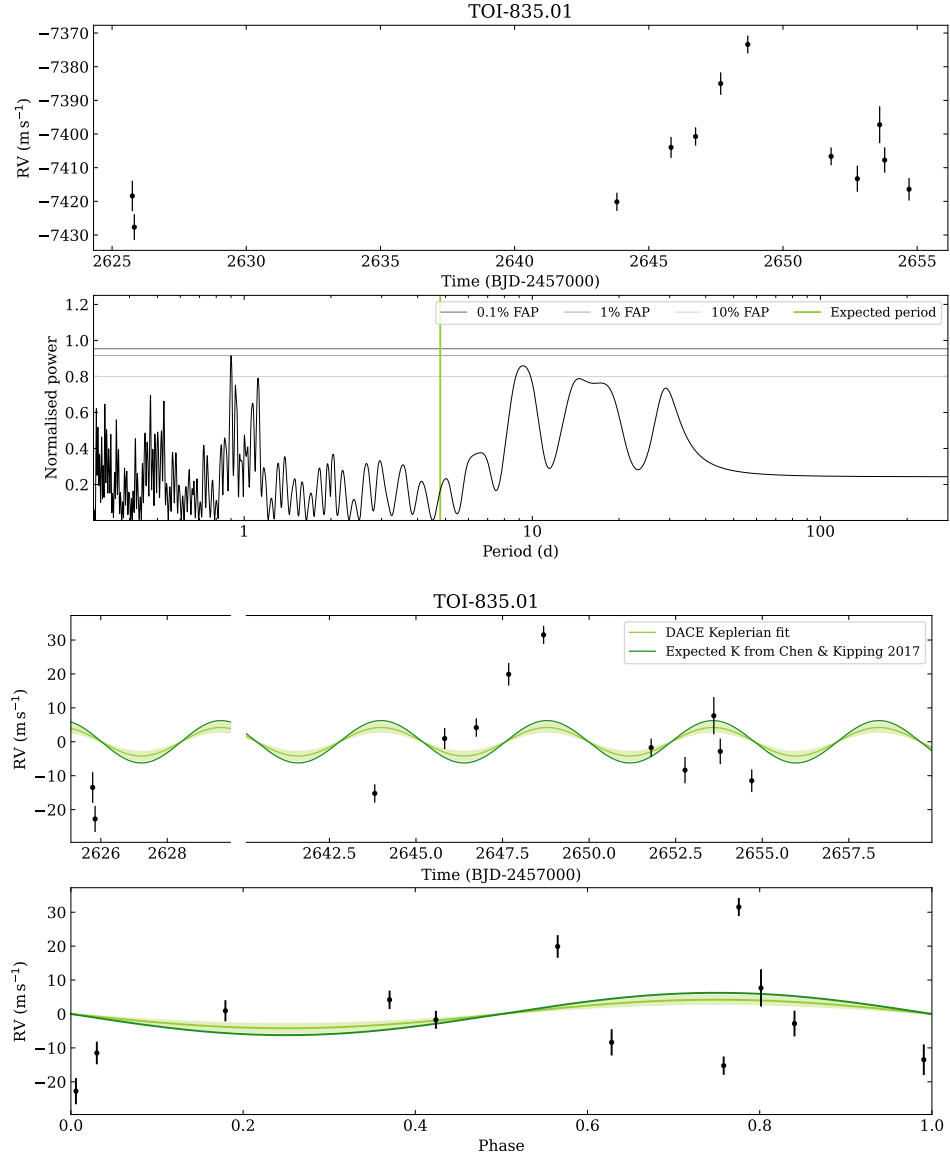


Figure 4.45: TOI-835.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

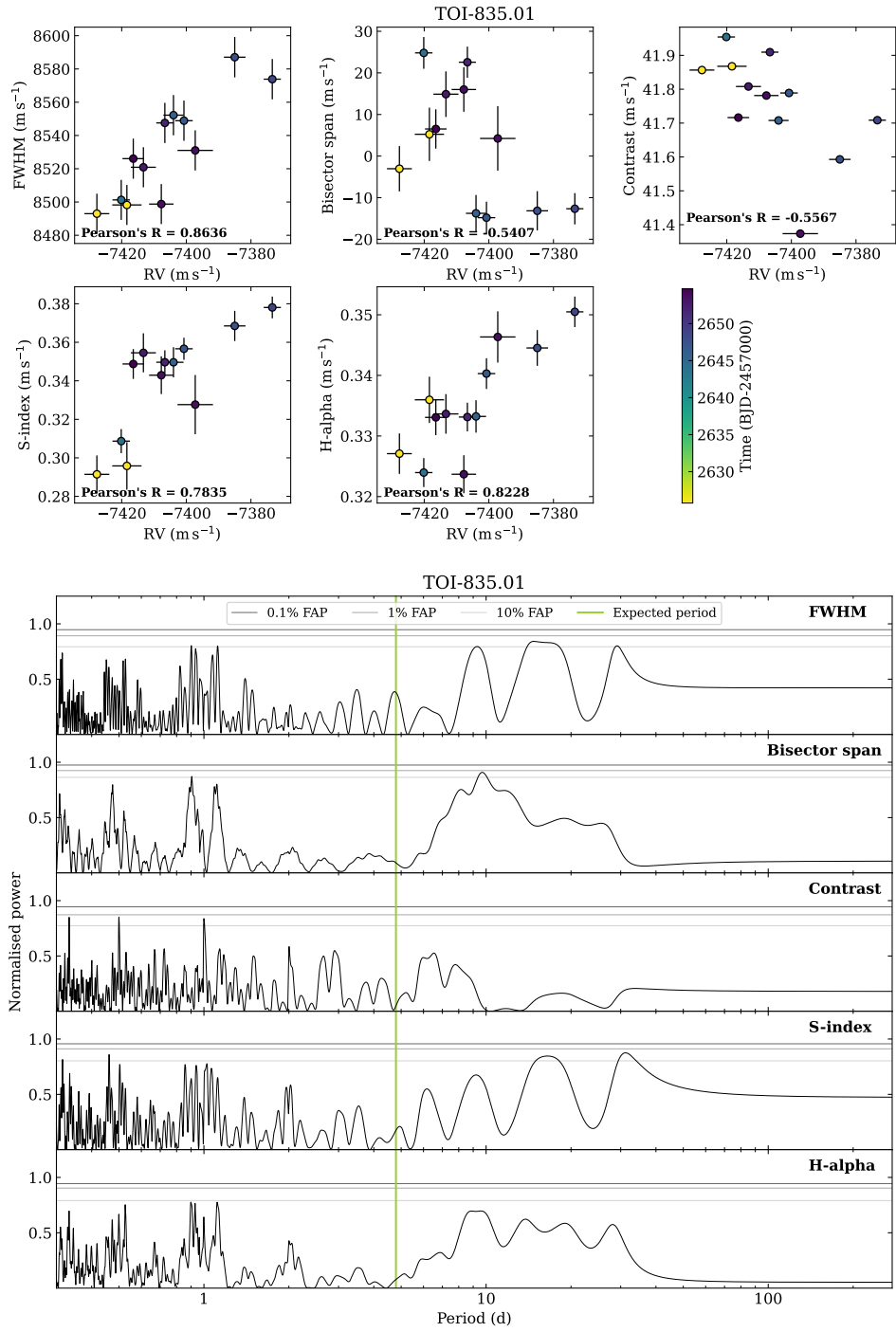


Figure 4.46: TOI-835.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

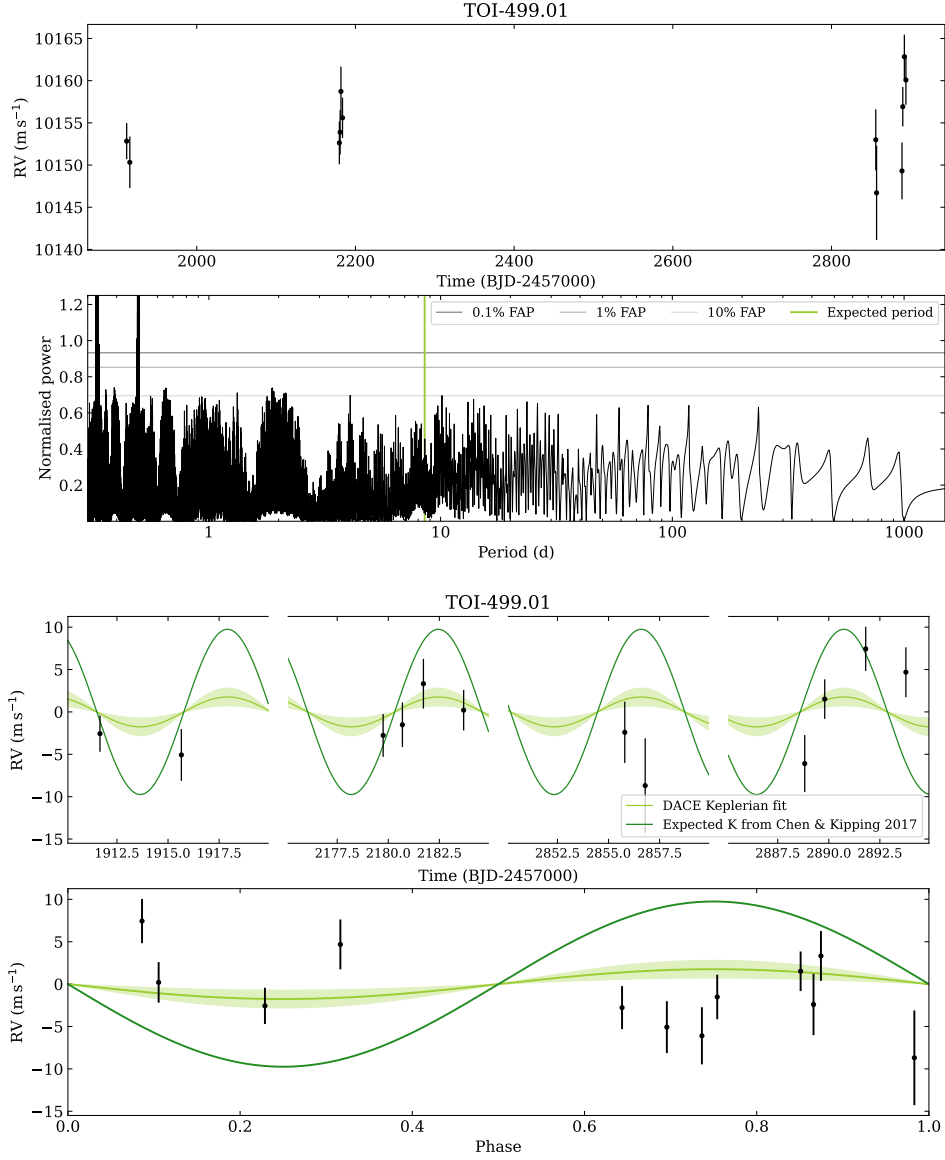


Figure 4.47: TOI-499.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

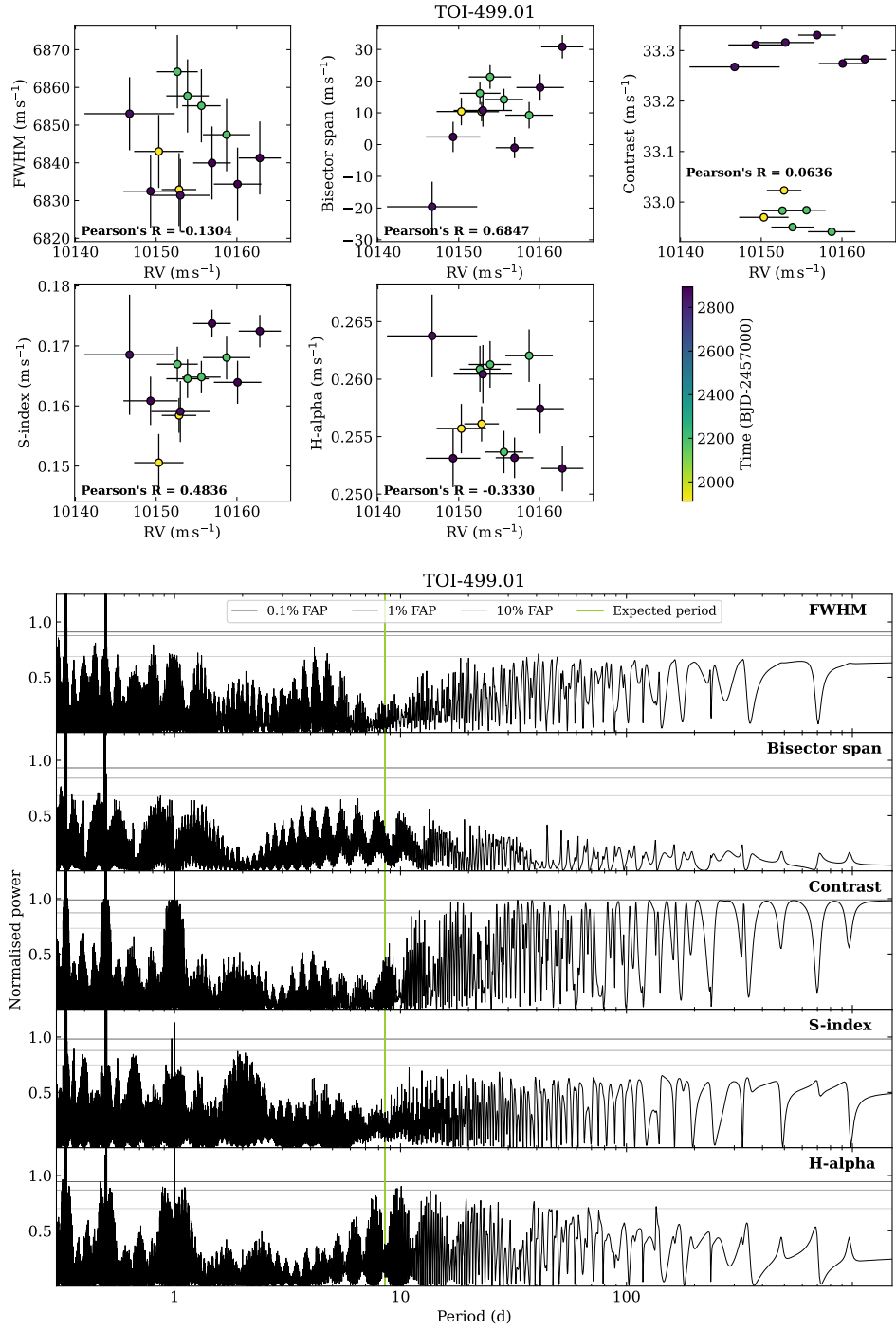


Figure 4.48: TOI-499.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

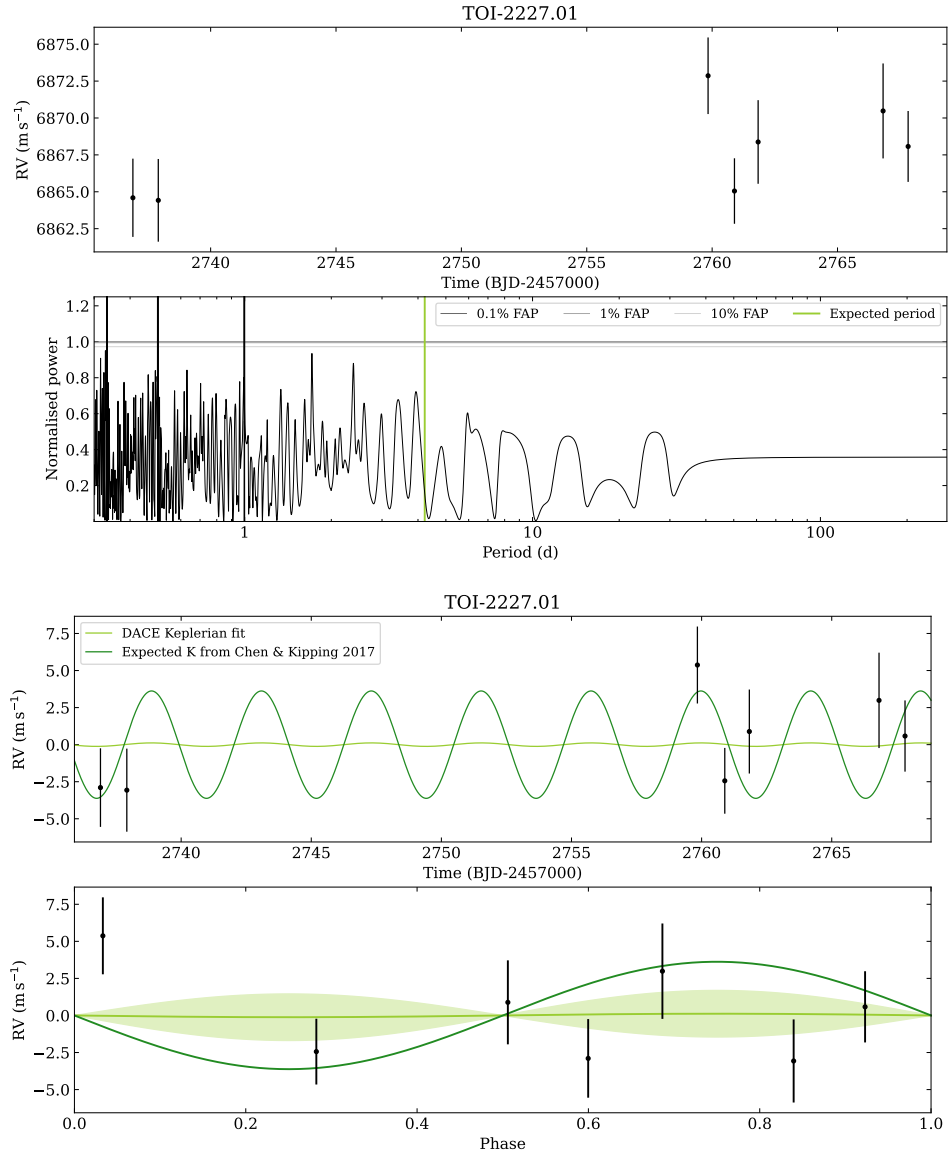


Figure 4.49: TOI-2227.01 RV data. Top: the raw RV data. Top middle: periodogram of the RV data, with the expected planetary period shown as a green line. Bottom middle: a simple Keplerian fit to the RV data is shown in pale green, with the shaded area representing 1 standard deviation from the model. The expected semi-amplitude of the planet signal from Chen and Kipping (2017) is shown in dark green. Bottom: same as bottom middle, but phase-folded on the expected planetary period.

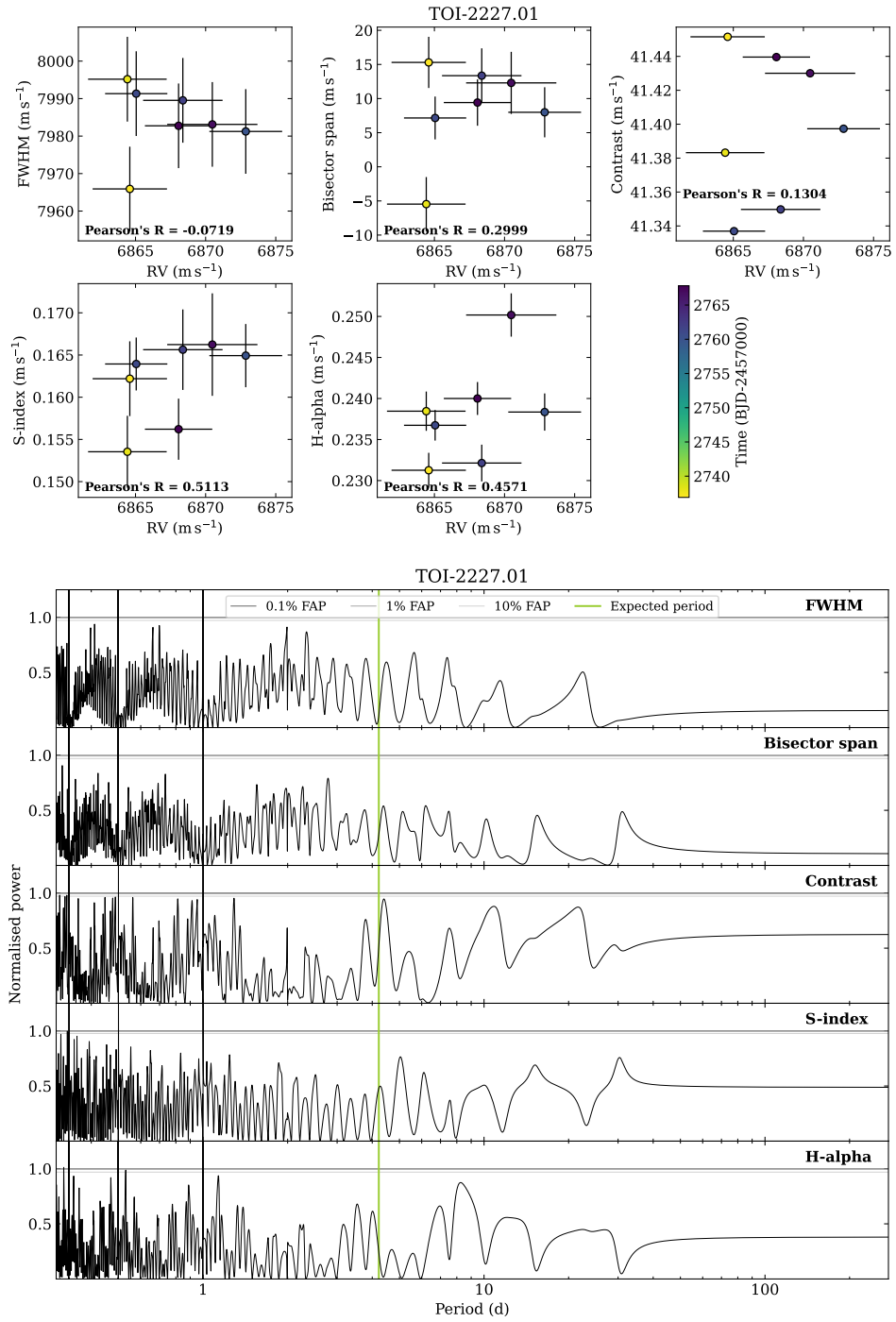


Figure 4.50: TOI-2227.01 activity indicator data. Top: each activity indicator plot against RV, showing any potential correlation, where Pearson's R is given for each. Bottom: periodograms of each activity indicator, where the expected planetary period is indicated with a green line.

4.6 Discussion

While it is hard to draw many conclusions at this stage, with many of the Nomads targets requiring further, full analyses, I can take a preliminary look at the sample and make some inferences.

4.6.1 Outcomes of the observation sample

First, the outcomes of the observations that were made are shown in Fig. 4.52 and summarised in Table 4.6.

2 targets are spectroscopic double-lined binaries, and thus confirmed false positives. 13 further targets are null results with flatlining planet signals, some likely false positives due to stellar activity. Some could still be planets, but are less massive than they are predicted to be, and only upper bounds on mass could be calculated.

3 planets are possible, but need dedicated stellar activity detrending to determine their mass. The rest of the sample is confirmed planets, made up of 1 planet that awaits further analysis, 4 that have publications in preparation, and 3 that are published, for a total of 8.

Overall, this is 2 definitely negative, 13 likely negative, and 11 positive. Looking at the overall sample, the numbers change to 2 definitely negative, 19 likely negative, and 52 positive. The hit rate in the observation sample is lower by virtue of essentially “mopping up” targets that were either alerted later, or not already followed up due to interest from other teams. However, confirming a planet is real doesn’t necessarily guarantee that it is actually within the desert, as values for radius can vary between the TOI release and publication.

It is also important to note that planets within the deep-desert remain rare. Most of the targets in this regime end up being false positives or null results, though there are a few key discoveries now confirmed.

4.6.2 Published planets in the sample

While only 3 of the Nomads targets have published values for period, radius, and mass thus far, there is the wider sample (Table 4.4) which provides us a further 29 targets to take an initial look at. These are presented in Fig. 4.52.

For this published Neptunian desert sample thus far, the period value from the TOI release is generally accurate to the publication value. However, this is not true for radius. Additionally, predicted values for mass are generally not accurate at all.

The targets that have ended up with planets outside the desert in period-radius space (Fig. 4.52, left) are TOI-564 (Davis et al., 2020), TOI-129 (Nielsen et al., 2020b), TOI-4599 (Luque et al., 2022), and TOI-2641 (Psaridi et al., 2023). The former two are more accurately

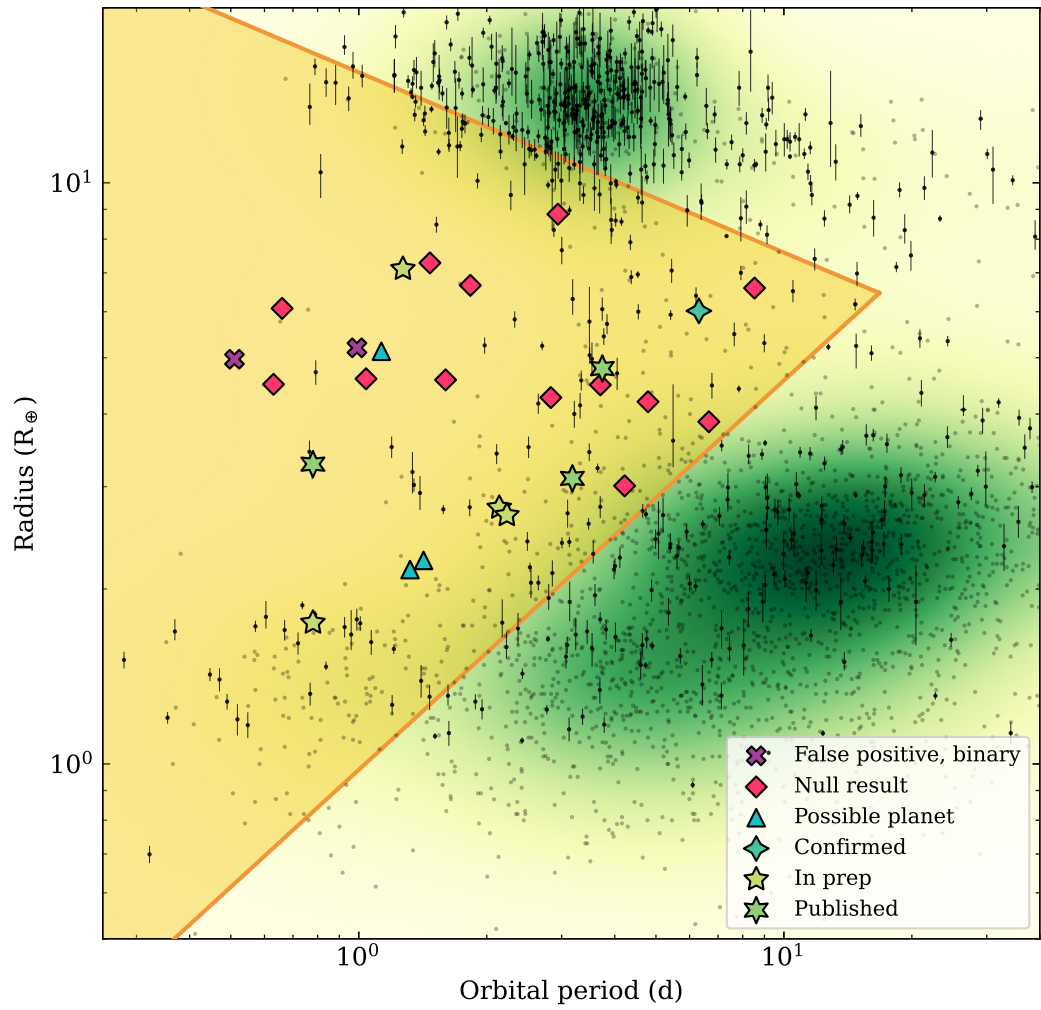


Figure 4.51: The Nomads observation sample, presented earlier in Fig. 4.52, and now with the outcomes of the observations depicted in this updated plot.

hot Jupiters; the latter are rocky. They have ended up outside the desert as their radii are wildly different to the TOI release values. There are a few other targets as well where TOI release radius has been particularly unreliable: TOI-731 (Lam et al., 2021) and TOI-969 (Lillo-Box et al., 2023). Why has this happened? TOI-564 b, TOI-129 b, and TOI-2641 b have grazing transits. In the case of TOI-969 b, it is probably due to an inaccurate stellar radius: the TOI release value for this was $0.82 \pm 0.08 R_{\odot}$, whereas the publication value from spectroscopic analysis is $0.671 \pm 0.015 R_{\odot}$ (Lillo-Box et al., 2023). This is similar for TOI-4599 c, where the TOI release value is $0.47 \pm 0.01 R_{\odot}$ and the publication value is $0.439 \pm 0.003 R_{\odot}$. TOI-731 b is a bizarre case: the TOI release value displayed on ExoFOP-TESS does not match up with the quoted TOI release value in the publication – 2.66 versus $0.75 R_{\oplus}$, respectively (Lam et al., 2021). I believe this might be due to a clerical error, as the *TESS* DVRs for the target also list values of the latter. This means that the release value is in actuality fairly reliable, though it was found to be slightly off due to dilution in the aperture.

The targets that have ended up with planets outside the desert in period-mass space (Fig. 4.52, right) are all of those outside the period-radius bounds (TOI-564, TOI-129, TOI-4599, TOI-2641), plus TOI-731, TOI-118, TOI-500, TOI-4517, and TOI-5094. It is evidently shown in the plot that a predicted value for mass from Chen and Kipping (2017) is, in most cases, wildly different from the actual mass obtained, even when the radius from the TOI release is accurate to the real radius value. This justifies not making any decisions on what to include in the sample based on a predicted mass.

So, out of the 32 targets in our sample that have published values, 9 are found to not be in the desert in either period-radius or period-mass space, almost 30 per cent. This hit rate will need to be updated when more of the Nomads targets have been published.

Finally, it's worth noting that our published targets thus far cluster towards the lower boundary of the desert. It will be interesting to see whether this remains true as the rest of the targets are analysed.

I also plot the published planets in the Nomads sample on a mass-radius diagram, shown in Fig. 4.53. There is an interesting trend that I note: planets with shorter periods cluster along the Earth-like and water lines, whereas longer period planets preferentially sit above the H-He line. However, this could be disrupted as more planets in the sample are fully analysed and published, so it cannot be drawn as a conclusion quite yet.

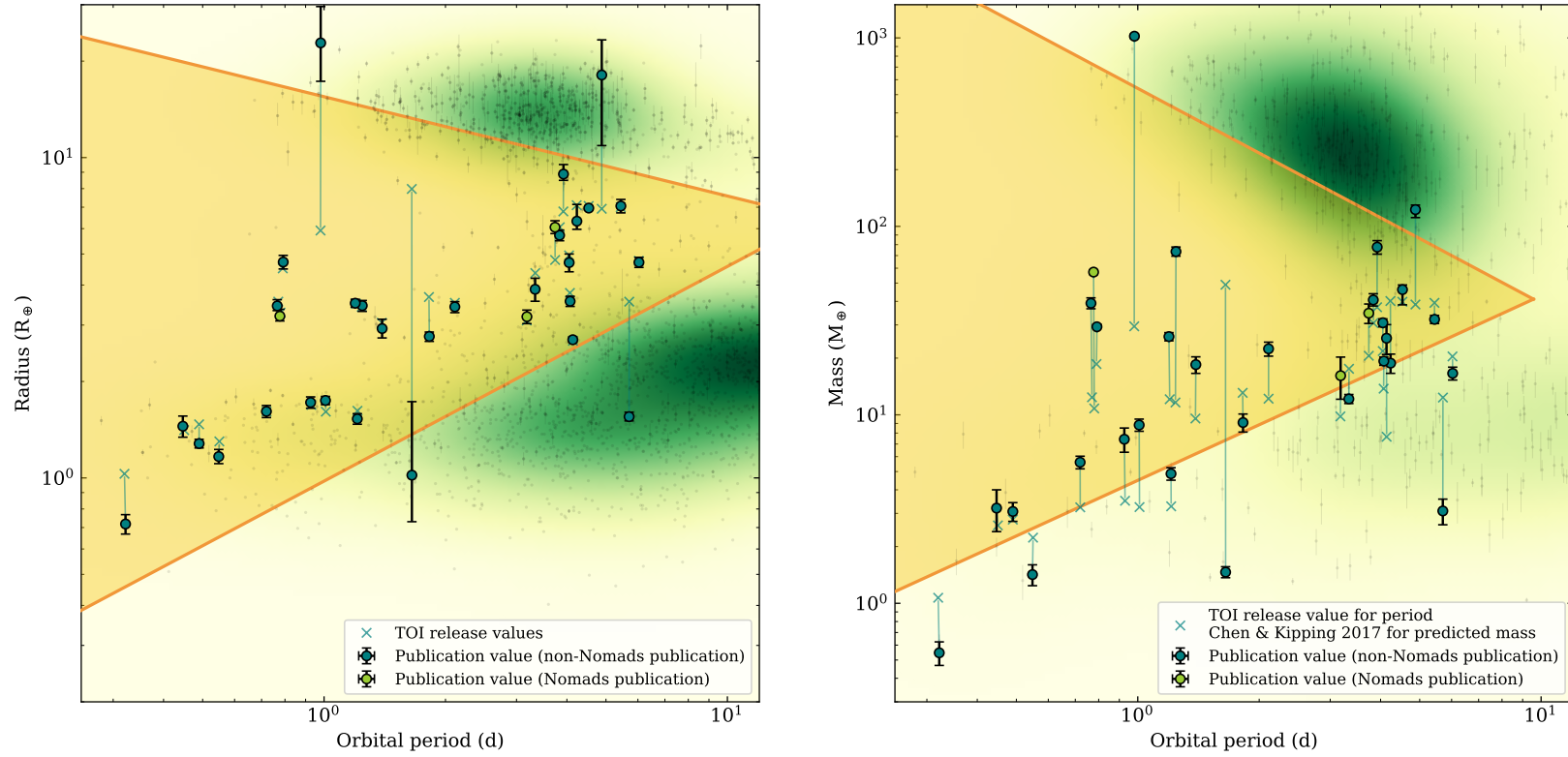


Figure 4.52: The currently published planets within the Nomads sample, shown in period-radius space (left) and period-mass space (right). The 3 targets observed by the Nomads programme are coloured green. For each target, both the TOI release values (blue crosses) and the published values (coloured circles with errorbars) are displayed and connected by a line.

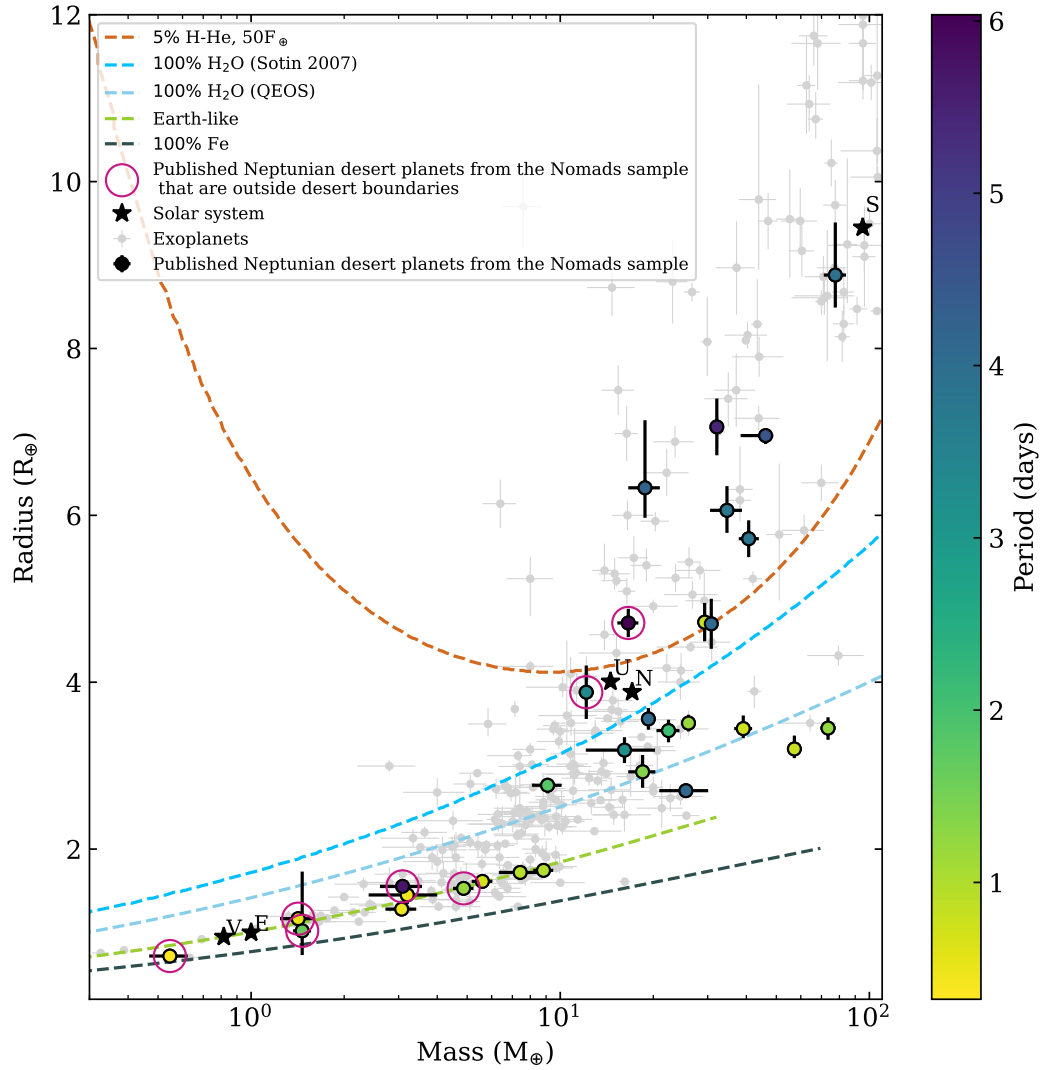


Figure 4.53: The published planets shown in Fig. 4.52, this time shown in mass-radius space with composition lines shown. The points are colour-coded by period. Planets that ended up being outside the Neptunian desert boundaries in Fig. 4.52 are circled in purple. Plot based upon Osborn et al. (2021a).

4.7 Future outlook

There is a lot of further work to be done on this Nomads project. A lot of it has been mentioned throughout this chapter, but I summarise here.

Most important is the full analysis of planets within the sample. A few are already in progress, but TOI-2358, TOI-426, and TOI-1839 all need activity detrending to determine their masses. There is also the matter of the null results - each one can have an upper mass limit determined, and then they should be investigated on a case-by-case basis as to why they may be a null result. If we find a general trend in this, it may help inform future surveys. We can then try to quantify the “success rate” of the sample, i.e. what proportion of candidates were confirmed.

Once every planet has period, radius, and mass values, we can return to the overall analysis of the sample, quantifying how many planets fall out of the desert due to their TOI release value for radius being incorrect, or their mass far from expected. We can examine the densities and compositions of planets within the desert for any trends like noted previously - e.g., are there any correlations between period and composition of planets within the desert?

As noted at the very beginning of this chapter, an unbiased and completeness-corrected search of the *TESS* FFIs is required to do a true statistical study of Neptunian desert planets. This would likely necessitate further RV observations on targets that still do not have masses, but this current sample and the observations made should hopefully contribute significantly to this goal.

Chapter 5

TOI-332 b: a super dense Neptune found deep within the Neptunian desert

Declaration and data availability

This chapter is a reproduction of the paper “TOI-332 b: a super dense Neptune found deep within the Neptunian desert” (Osborn et al., 2023), published in the Monthly Notices of the Royal Astronomical Society in August 2023. The majority of the analysis and writing was performed by myself. Contributions from co-authors, including observational data, analysis, and partial writing of sections, are found in Sections 5.2.1 (specifically photometry from LCOGT by Karen A. Collins, PEST by Thiam-Guan Tan, and WASP by Coel Hellier), 5.10 (by Rachel A. Matson and Elisabeth C. Matthews), 5.3 (by Nuno C. Santos, Sérgio G. Sousa, Vardan Adibekyan, Elisa Delgado-Mena, João Gomes da Silva, Keivan G. Stassun, and Malcolm Fridlund), 5.5.1 (by Henrik Knierim, Caroline Dorn, and Ravit Helled), 5.6 (by Jorge Lillo-Box), and 5.5.3 (by Jorge Fernández Fernández and George W. King).

The *TESS* data are available from the Mikulski Archive for Space Telescopes (MAST), at <https://heasarc.gsfc.nasa.gov/docs/tess/data-access.html>. The other photometry from LCOGT and PEST, and the high-resolution imaging data, are available for public download from the ExoFOP-TESS archive at <https://exofop.ipac.caltech.edu/tess/target.php?id=139285832>. The full HARPS RV data products are publicly available from the ESO archive, at http://archive.eso.org/wdb/wdb/adp/phase3_main/form. The model code underlying this article will be shared on reasonable request. The MCMC chains are available from Zenodo, at <https://doi.org/10.5281/zenodo.8199962>.

Abstract

To date, thousands of planets have been discovered, but there are regions of the orbital parameter space that are still bare. An example is the short period and intermediate mass/radius space known as the “Neptunian desert”, where planets should be easy to find but discoveries remain few. This suggests unusual formation and evolution processes are responsible for the planets residing here. We present the discovery of TOI-332 b, a planet with an ultra-short period of 0.78 d that sits firmly within the desert. It orbits a K0 dwarf with an effective temperature of 5251 ± 71 K. TOI-332 b has a radius of $3.20^{+0.16}_{-0.12} R_{\oplus}$, smaller than that of Neptune, but an unusually large mass of $57.2 \pm 1.6 M_{\oplus}$. It has one of the highest densities of any Neptune-sized planet discovered thus far at $9.6^{+1.1}_{-1.3} \text{ g cm}^{-3}$. A 4-layer internal structure model indicates it likely has a negligible hydrogen-helium envelope, something only found for a small handful of planets this massive, and so TOI-332 b presents an interesting challenge to planetary formation theories. We find that photoevaporation cannot account for the mass loss required to strip this planet of the Jupiter-like envelope it would have been expected to accrete. We need to look towards other scenarios, such as high-eccentricity migration, giant impacts, or gap opening in the protoplanetary disc, to try and explain this unusual discovery.

5.1 Introduction

One of the key outcomes of the *Kepler* mission (Borucki et al., 2010) was the population studies performed on the planets it discovered. This led to the identification of the “Neptunian desert” (also known as the “hot Neptune desert”, “sub-Jovian desert”, and “evaporation desert”), a region of period-radius and period-mass parameter space where planets have, so far, been rarely found. The desert was first noted by Szabó and Kiss (2011), and has been the subject of many studies in the years since (e.g., Boué et al., 2012; Beaugé and Nesvorný, 2013; Helled et al., 2016; Lundkvist et al., 2016), and its boundaries were first formally defined by Mazeh et al. (2016). As shown in Fig. 5.1, it is a wedge shaped region where the upper boundary at large radii (or mass) decreases with increasing semi-major axis, and a lower boundary at small radii (or mass) which increases with increasing semi-major axis. The desert roughly encompasses intermediately-sized planets (approximately $2 R_{\oplus} < R_p < 9 R_{\oplus}$ and $10 M_{\oplus} < M_p < 250 M_{\oplus}$) with periods out to ~ 5 days.

This should not be due to an observational bias, as Neptune-sized planets with short periods are readily discovered by transit surveys like *Kepler* and, more recently, the Transiting Exoplanet Survey Satellite (*TESS*, Ricker et al., 2015). Theories have been put forward to explain the desert’s existence and boundaries (e.g., Owen and Lai, 2018; Vissapragada et al.,

2022). The lower boundary could be caused by photoevaporation of planets above the boundary, stripping their envelopes and reducing their radii/mass; while the upper boundary seems to be stable against photoevaporation, and may instead be understood as a “tidal disruption barrier”, where planets below and left of the boundary migrating inwards can no longer successfully circularise and stabilise (see review by Dawson and Johnson (2018)).

In the years since its discovery, the desert has become more populated with planet discoveries, especially around its boundaries. However, there are so far only four planets with precisely determined masses (i.e. an error on their mass of better than 20 per cent) found deep within the desert, far from the boundaries set by Mazeh et al. (2016): NGTS-4 b (West et al., 2019); LTT-9779 b (Jenkins et al., 2020); TOI-849 b (Armstrong et al., 2020); and TOI-2196 b (Persson et al., 2022). They are annotated in Fig. 5.1. There are an additional few without precise masses: K2-100 b (Barragán et al., 2019); K2-278 b (Livingston et al., 2018), and Kepler-644 b (Morton et al., 2016), the latter two being validated and having no mass determination. The few planets found in this barren desert are likely to have undergone unusual formation and/or evolutionary processes compared to those in more populated parameter spaces. There are now concerted efforts to find planets in and around the desert (e.g. Magliano et al., 2023; Bourrier et al., 2023) to determine what sculpts it.

The aim of the HARPS-NOMADS program is to characterise planets in the Neptunian desert discovered by *TESS*, as the stars it observes are bright enough for effective radial velocity follow up. By precisely determining their masses and radii, we can constrain densities and thus the internal structures of these planets in order to understand their formation and evolution, leading to a better understanding of the origins of the desert itself.

We present here the detection and characterisation of TOI-332 b, an ultra-short period planet with an unusually high density located deep within the Neptunian desert. In Section 5.2, we present the observations of the TOI-332 system, including photometry, spectroscopy, and high-resolution imaging. The spectroscopic analysis and derivation of chemical abundances of the star is then described in Section 5.3. In Section 5.4, we describe the joint fit model to the data. In Section 5.5, we present the results of the joint fit, discuss the nature of TOI-332 b, theorise potential scenarios for its formation and evolution, and outline opportunities for further follow up of the system. Section 5.6 sets out our conclusions.

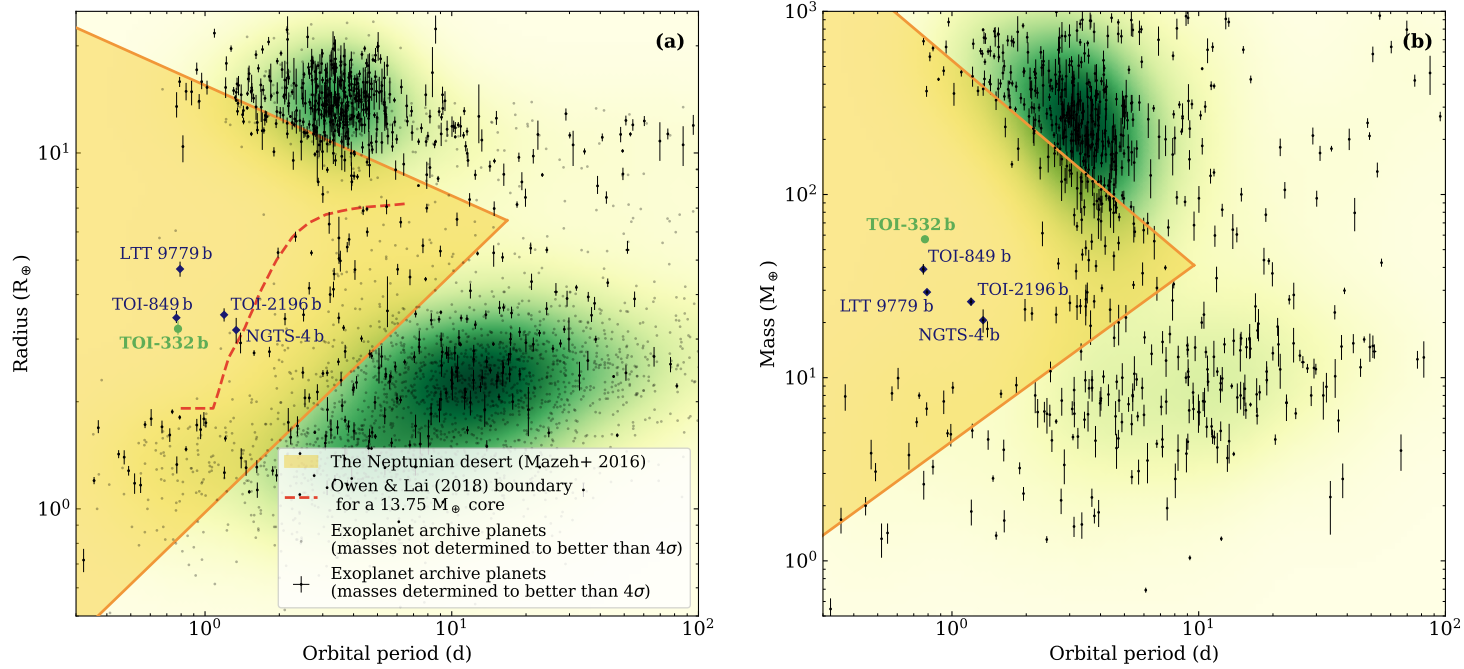


Figure 5.1: TOI-332 b (green circle) in the context of the Neptunian desert, with (a) showing period-radius space, and (b) showing period-mass. The Neptunian desert boundaries from Mazeh et al. (2016) are plotted as solid lines, with the enclosed Neptunian desert area shaded in yellow. In (a), the dashed line is a numerical solution for the lower boundary of the desert determined for a $13.75 M_{\oplus}$ core in Owen and Lai (2018). Known planets were sourced from the NASA exoplanet archive (<https://exoplanetarchive.ipac.caltech.edu/>) on 3 May 2023: those without mass determinations or mass determinations worse than 4σ are plotted as pale grey dots in (a) only; planets with mass determinations better than 4σ are plotted as black dots in both (a) and (b). Population density of known planets is shaded in green, where darker green denotes more planets discovered in that region of parameter space: in (a) this includes all planets; in (b) this includes only planets with mass determination better than 4σ . Particular planets with mass determination to better than 4σ that are considered to be in the “deep” Neptunian desert are labelled (dark blue diamonds).

5.2 Observations

In this section, we describe the instrumentation and observations used for the detection and characterisation of the TOI-332 system.

5.2.1 Photometry

TESS

The TOI-332 system (TIC 139285832, see Table 5.1) was observed in *TESS* Sectors 1 (25 July - 22 Aug 2018, hereafter S1, see Fig. 5.2) and 2 (22 August - 20 September 2018, hereafter S2) with a 30 min cadence in the full-frame images (FFIs). TOI-332.01 (now TOI-332 b) was detected in the FFIs by the MIT Quick-Look Pipeline (QLP, Huang et al., 2020) as part of the early *TESS* Data Alerts, and alerted on 20 December 2018 (Guerrero et al., 2021). It was then re-observed in Sector 28 (30 July - 26 August 2020, hereafter S28) on Camera 2 with a 2 min cadence. The Data Validation report (Twicken et al., 2018; Li et al., 2019) difference image centroid offsets determined from the S28 pixel data locate the transit source within 2.56 ± 2.76 arcsec of TOI-332, and exclude all other TICv8 objects as possible sources of the transit signal. The detection gave a period of 0.77685 ± 0.0003 d, a transit duration of 1.43 ± 0.442 h, and a depth of 830 ± 8 ppm. The data products, including calibrated full-frame images and light curves, are available on the Mikulski Archive for Space Telescopes (MAST; <https://archive.stsci.edu/missions-and-data/transiting-exoplanet-survey-satellite-tess>), and were produced by the *TESS* Science Processing Operations Center (SPOC, Jenkins et al., 2016; Caldwell et al., 2020) at NASA Ames Research Center.

We downloaded the publicly available photometry provided by the SPOC pipeline, and used the Presearch Data Conditioning Simple Aperture Photometry (PDCSAP), from which common trends and artefacts have been removed by the SPOC Presearch Data Conditioning (PDC) algorithm (Twicken et al., 2010; Smith et al., 2012; Stumpe et al., 2012, 2014). The median-normalised PDCSAP flux, after removal of data points flagged as being affected by excess noise, is shown in Fig. 5.3. No further detrending of the light curves was deemed necessary as they are relatively flat across the whole time series, showing little stellar activity. We also recover no periodicity from either the PDCSAP or SAP (Simple Aperture Photometry, where no trends and artifacts have been removed) flux that may be indicative of a stellar rotation period. The phase folded transits and best fit model are also shown in Fig. 5.3.

Table 5.1: Details for the TOI-332 system.

Property	(unit)	Value	Source
Identifiers			
TIC ID		139285832	TICv8
2MASS ID		J23121409-4452349	2MASS
Gaia ID		6529471108882243840	Gaia DR3
Astrometric properties			
R.A.	(J2000.0)	23:12:14.10	Gaia DR3
Dec	(J2000.0)	-44:52:34.77	Gaia DR3
Parallax	(mas)	4.54 ± 0.03	Gaia DR3
Distance	(pc)	222.85 ± 3.69	Gaia DR3
$\mu_{\text{R.A.}}$	(mas yr ⁻¹)	35.86 ± 0.01	Gaia DR3
μ_{Dec}	(mas yr ⁻¹)	-37.62 ± 0.02	Gaia DR3
Photometric properties			
<i>TESS</i>	(mag)	11.527 ± 0.006	TICv8
B	(mag)	13.10 ± 0.03	TICv8
V	(mag)	12.35 ± 0.05	TICv8
G	(mag)	12.0545 ± 0.0002	Gaia DR3
J	(mag)	10.78 ± 0.02	2MASS
H	(mag)	10.41 ± 0.02	2MASS
K	(mag)	10.32 ± 0.02	2MASS

Sources: TICv8 (Stassun et al., 2019), 2MASS (Skrutskie et al., 2006), Gaia Data Release 3 (Gaia Collaboration et al., 2023).

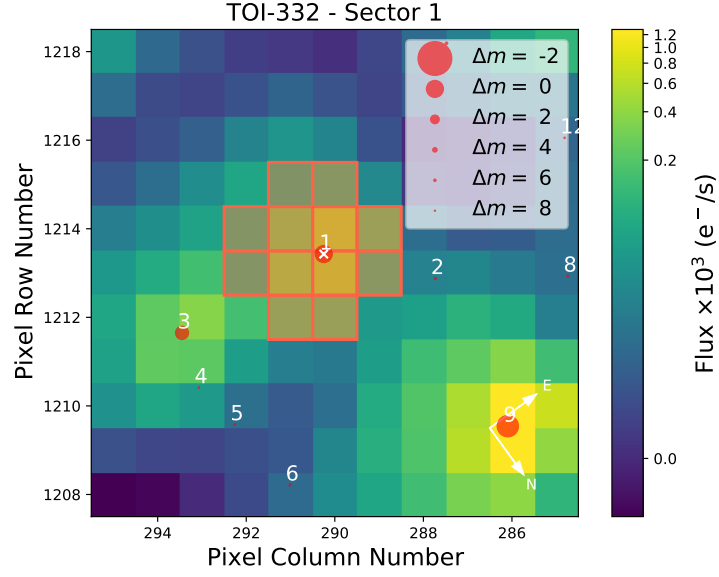


Figure 5.2: The Target Pixel File (TPF) for TOI-332 (marked as a white cross) from *TESS* S1. Other *Gaia* DR3 sources within a limit of 8 *Gaia* magnitudes difference from TOI-332 are marked as red circles, and are numbered in distance order from TOI-332. The aperture mask is outlined and shaded in red. This figure was created with *tpfplotter* (Aller et al., 2020).

LCOGT

The *TESS* pixel scale is ~ 21 arcsec per pixel and photometric apertures typically extend out to roughly 1 arcmin, generally causing multiple stars to blend in the *TESS* aperture (the aperture for the *TESS* S1 data for TOI-332 is shown in Fig. 5.2). To attempt to determine the true source of the TOI-332 detection in the *TESS* data and refine its ephemeris and transit shape, we conducted ground-based photometric follow-up observations of the field around TOI-332 as part of the *TESS* Follow-up Observing Program¹ Sub Group 1 (TFOP; Collins, 2019).

We observed six full predicted transit windows of TOI-332.01 using the Las Cumbres Observatory Global Telescope (LCOGT; Brown et al., 2013) 1.0m network nodes. The details of each observation are provided in the caption of Fig. 5.4. We used the *TESS* Transit Finder, which is a customized version of the *Tapir* software package (Jensen, 2013), to schedule our transit observations. The 1 m telescopes are equipped with 4096×4096 SINISTRO cameras having an image scale of 0.389 arcsec per pixel, resulting in a $26 \text{ arcsec} \times 26 \text{ arcsec}$ field of view. The images were calibrated by the standard LCOGT BANZAI pipeline (McCully et al., 2018). Differential photometric data were extracted using

¹<https://tess.mit.edu/followup>

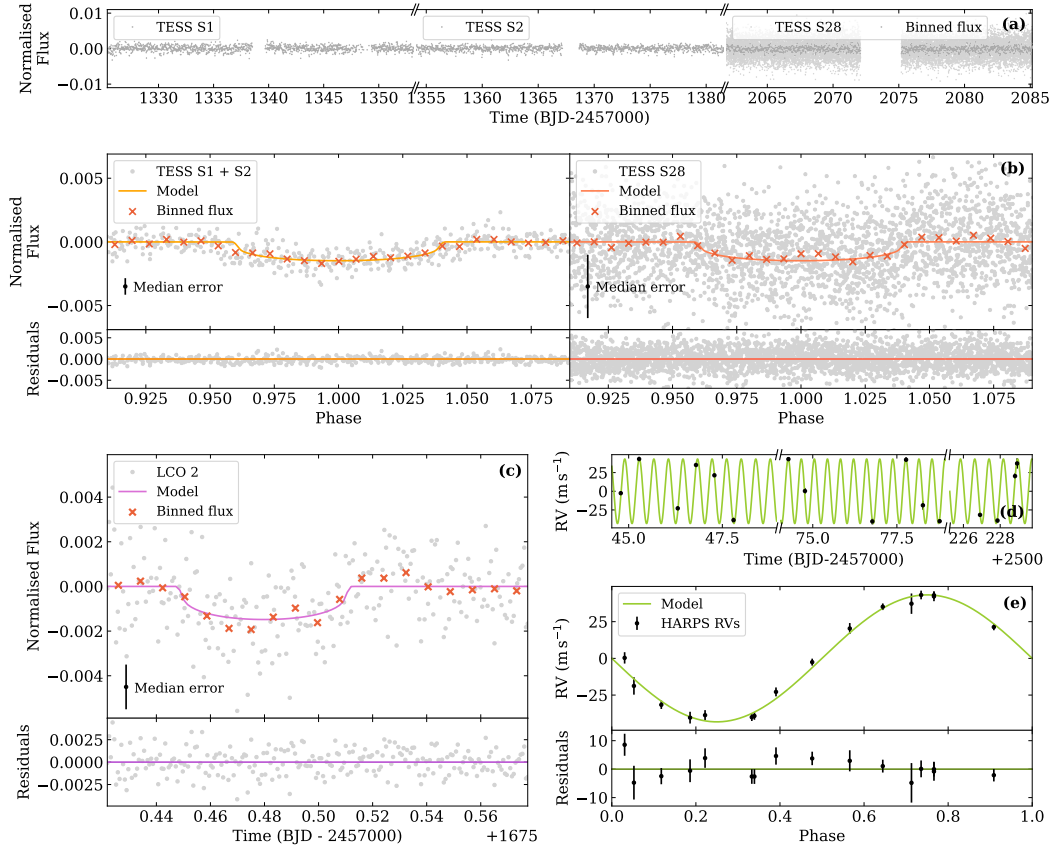


Figure 5.3: Joint fit model to the *TESS*, LCO, and HARPS data.

(a) *TESS* PDCSAP light curve for Sectors 1, 2 and 28 (circles), with time given as Barycentric Julian Date (BJD). Sectors 1 and 2 are 30 min cadence data, Sector 28 is 2 min cadence, hence the higher levels of noise in the latter. The Sector 28 data binned to 30 min has been overplotted in dark grey.

(b) Phase folded *TESS* 30 min cadence data (grey circles) from Sectors 1 and 2 (left), and 2 min cadence data from Sector 28 (right). Binned flux (red crosses) and the best fit model (solid line) are overplotted, and the median error on the flux is displayed (one standard deviation, black error bar, bottom left). Residuals when the model is subtracted are shown in the bottom panels.

(c) An example of the phase folded LCO data, using the second transit obtained (chronologically) by LCO (the model is fit to all of the LCO transits and the full data can be seen in Fig. 5.4). Symbols and model are as in **(b)**, with residuals in the bottom panel.

(d-e) The HARPS data (black circles), shown as a time series in **(d)**, and the phase folded data in **(e)**. The model is plotted as in **(b)**, with residuals in the bottom panel.

AstroImageJ (Collins et al., 2017). As shown in Fig. 5.4, we detected transit-like signals in all six TOI-332 follow-up light curves using photometric apertures with radii in the range of 3.1 arcsec to 7.8 arcsec, which exclude flux from the nearest neighbour of TOI-332 in the *Gaia* DR3 and TICv8 catalogs (which is 51 arcsec northeast, and is the target numbered 2 in Fig. 5.2). We therefore confirm that the TOI-332.01 signal in the *TESS* data occurs on-target relative to all known *Gaia* DR3 and TICv8 stars.

PEST

We observed TOI-332 in the V band from the Perth Exoplanet Survey Telescope (PEST) near Perth, Australia. At the time, the 0.3 m telescope was equipped with a 1530×1020 SBIG ST-8XME camera with an image scale of 1.2 arcsec per pixel resulting in a 31 arcmin × 21 arcmin field of view. A custom pipeline based on C-Munipack was used to calibrate the images and extract the differential photometry. Unfortunately, there is a gap during the transit egress due to cloud cover, and poor weather negatively affected the data quality; therefore, we do not include it in our joint fit model, but present the data with the model over-plotted in Fig. 5.5.

WASP

WASP-South, an array of 8 wide-field cameras, was the Southern station of the WASP transit-search project (Pollacco et al., 2006). It observed the field of TOI-332 in the years 2006, 2007, 2010 and 2011 when equipped with 200-mm, f/1.8 lenses, and then again in 2012, 2013 and 2014 when equipped with 85-mm, f/1.2 lenses. It observed on each clear night, with a typical 10 min cadence, and accumulated 88 000 photometric data points on TOI-332.

We searched the data for any rotational modulation using the methods from Maxted et al. (2011), shown in Fig. 5.6. We find a significant periodicity in data from one year, spanning 168 nights in 2007, with an estimated false-alarm probability of 0.15 per cent. The period is 20.9 ± 1.0 d and the amplitude 3 mmag. We note that there is also a peak near 36 days, though it is not significant in itself. In 2012, a possible modulation with a similar period (18.4 ± 1.5 d) has a lower significance (8 per cent false-alarm likelihood). No significant periodicity was detected in other years. We discuss the periodicity further in Section 5.3.

We also note that the standard WASP transit-detection algorithm (Collier Cameron et al., 2007), when run on the same 2007 dataset, detects the transit and reports a period of 0.77663 ± 0.00012 d with an epoch of TDB 2454343.4652 \pm 0.0079. The period matches the *TESS* period to 1 part in 2000 while the phase matches an extrapolated ephemeris to within 3 per cent. However, when run on the full dataset combined, the algorithm does not

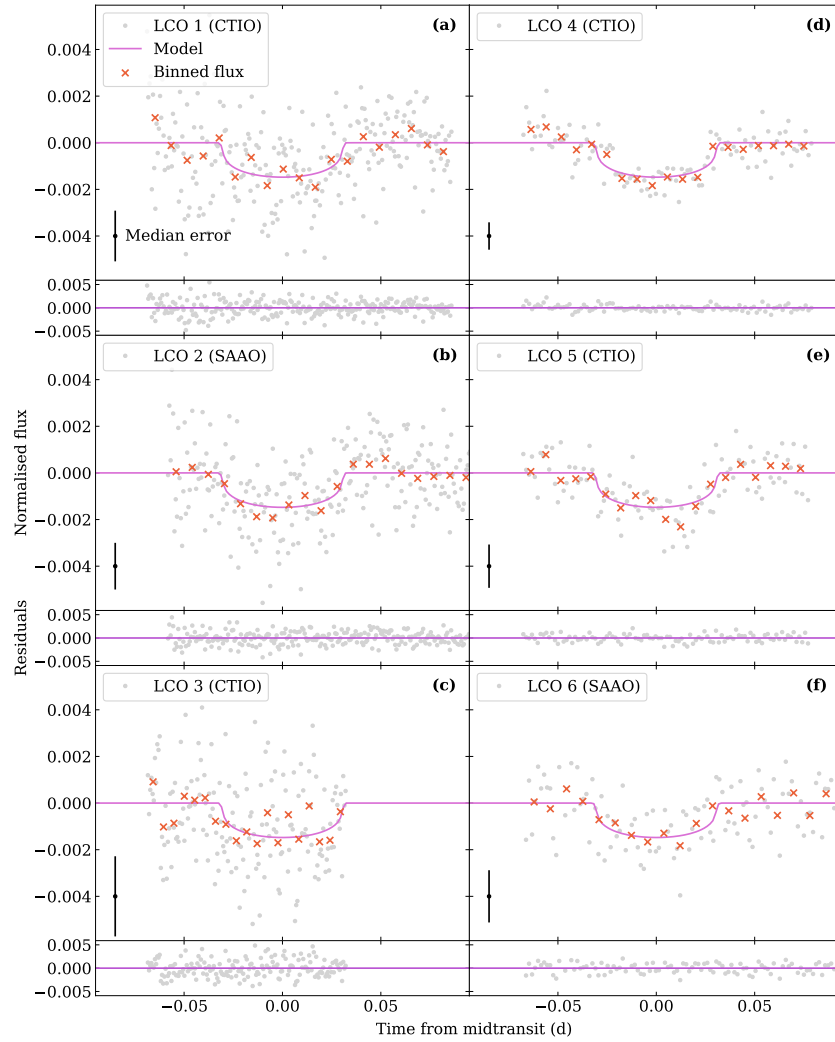


Figure 5.4: Photometric data from LCOGT. For each, the flux (grey circles), binned flux (red crosses), median error on the flux (one standard deviation, black error bar, bottom left) and best fit model (solid line) are shown. Residuals after the best fit model is subtracted are shown in the bottom panels. Data were captured at two different telescopes in the Global Network, the Cerro Tololo Inter-American Observatory (CTIO) and the South African Astronomical Observatory (SAAO), on the nights of:

- (a) 1 June 2019 at CTIO in Sloan i' band (“LCO 1”);
- (b) 10 July 2019 at SAAO in Sloan i' band (“LCO 2”);
- (c) 27 July 2019 at CTIO Pan-STARRS z -short band (“LCO 3”);
- (d) 10 Aug 2019 at CTIO in Sloan g' band (“LCO 4”);
- (e) 10 Aug 2019 at CTIO in Pan-STARRS z -short band (“LCO 5”);
- and (f), 24 Aug 2020 at SAAO in Pan-STARRS z -short band (“LCO 6”).

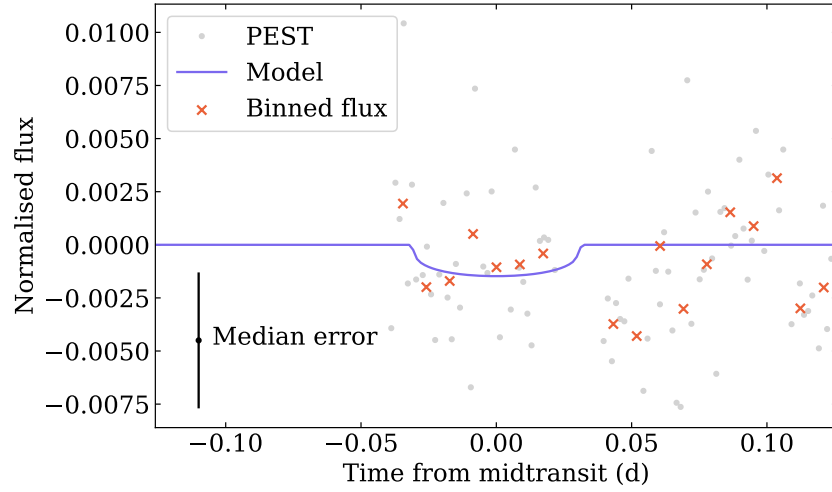


Figure 5.5: Photometric data from PEST (a single transit event), where the flux (grey circles), binned flux (red crosses), and median error on the flux (one standard deviation, black error bar, bottom left) are shown. This data was not included in our joint fit model (see Section 5.2.1), but we overplot the best fit model with a solid line.

find the transit, though this is explainable given that red noise in other years can destroy the detection. At a depth of 0.15 per cent, the transit would be the shallowest detected in WASP data, though it is comparable to the detection of the 0.17 per cent transit of HD 219666 b (Hellier et al., 2019b). We conclude that this detection is likely, but not securely, real, and thus we report it here as the earliest detection of the transit of TOI-332 b. Since planets in short-period orbits are expected to undergo tidal decay, timings over the longest possible time span are of interest, and we discuss this further in Section 5.5.5.

Due to the uncertainty in the detection, we do not include the WASP data in our joint fit model, but we present the 2007 dataset with the best fit model over-plotted in Fig. 5.7.

5.2.2 Spectroscopy

HARPS

We made radial velocity (RV) measurements of TOI-332 with the High Accuracy Radial velocity Planet Searcher (HARPS) spectrograph mounted on the ESO 3.6 m telescope at the La Silla Observatory in Chile (Pepe et al., 2002). A total of 16 spectra were obtained between 25 November 2021 and 29 May 2022 under the HARPS-NOMADS large programme (ID 1108.C-0697, PI: Armstrong). The instrument (with resolving power $R = 115\,000$) was used in high-accuracy mode (HAM) with an exposure time of 2400 s, and 1-2 observations of the star were made per night. The data were reduced using the standard offline HARPS data reduction pipeline, and a K5 template was used in a weighted cross-correlation function

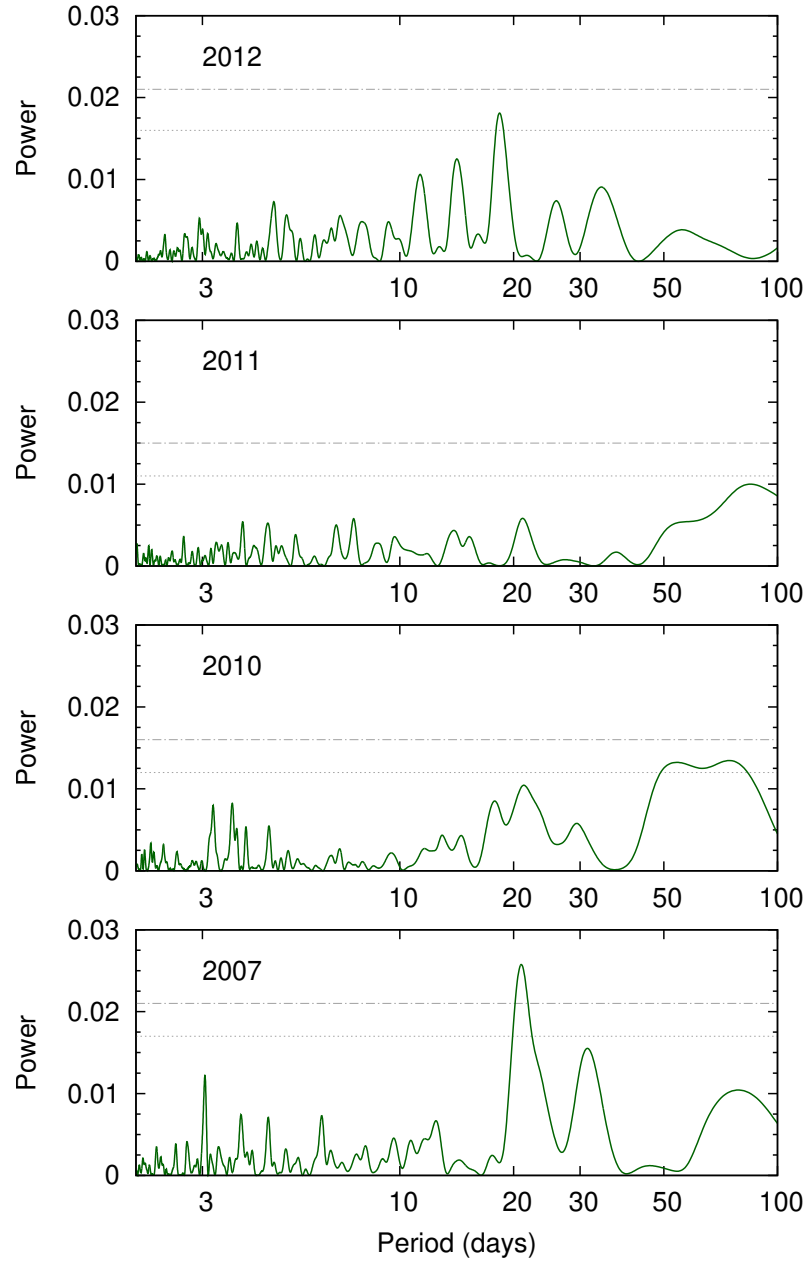


Figure 5.6: Periodograms of the WASP-South data for TOI-332 from four different years of data. The horizontal lines mark the estimated 10 per cent- and 1 per cent-likelihood false-alarm levels. A significant periodicity at 20.9 ± 1.0 d was seen in 2007.

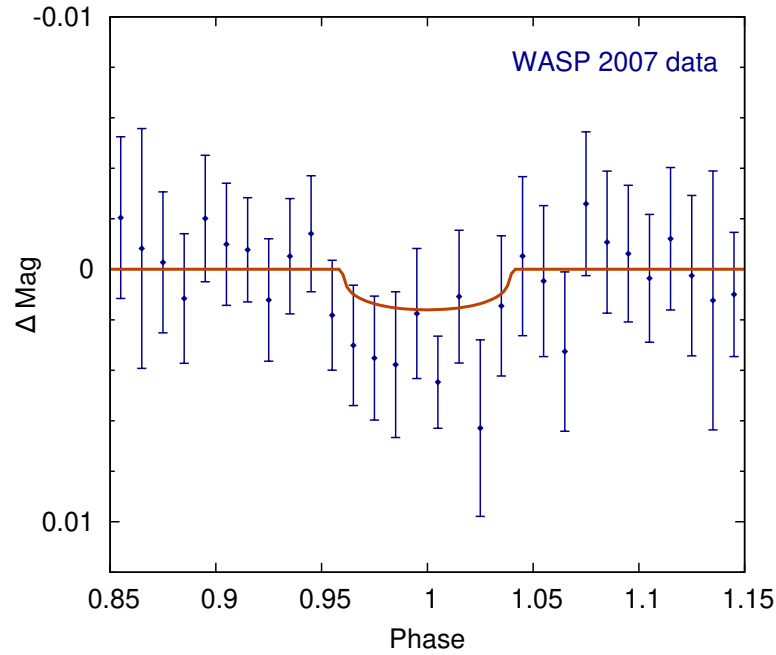


Figure 5.7: Photometric data from WASP (2007 data only), phase-folded on the transit ephemeris. This data was not included in our joint fit model (see Section 5.2.1), but we overplot the best fit model with a solid line.

(CCF) to determine the RV values (Baranne et al., 1996; Pepe et al., 2002). The line bisector (BIS) and full-width at half-maximum (FWHM) were measured using previously published methods (Boisse et al., 2011). The RV measurements can be found in Table 5.2, and the RV data and single-planet Keplerian best fit are shown in Fig. 5.3.

No correlation was detected between the RVs and the FWHM and bisector span of the CCF, or the S and $H\alpha$ activity indexes, shown in Fig. 5.8. After removing the contribution from TOI-332 b, we studied the RV residuals and found no evidence of further periodicity in those or in the activity indicators as shown in Fig. 5.9. Unfortunately, this means the RVs give no indication of a possible stellar rotation period that would corroborate that found by WASP in Section 5.2.1 or derived later in Section 5.3. We note that one of the $H\alpha$ points, from the night of 29 December 2021, is an outlier, shown in Fig. 5.8. Investigating the spectrum from this night, we find a narrow emission line in the centre of the $H\alpha$ line that may be indicative of a flare; however, the RV point corresponding to this night is not an outlier, nor does it have an anomalously large error, so we retain it.

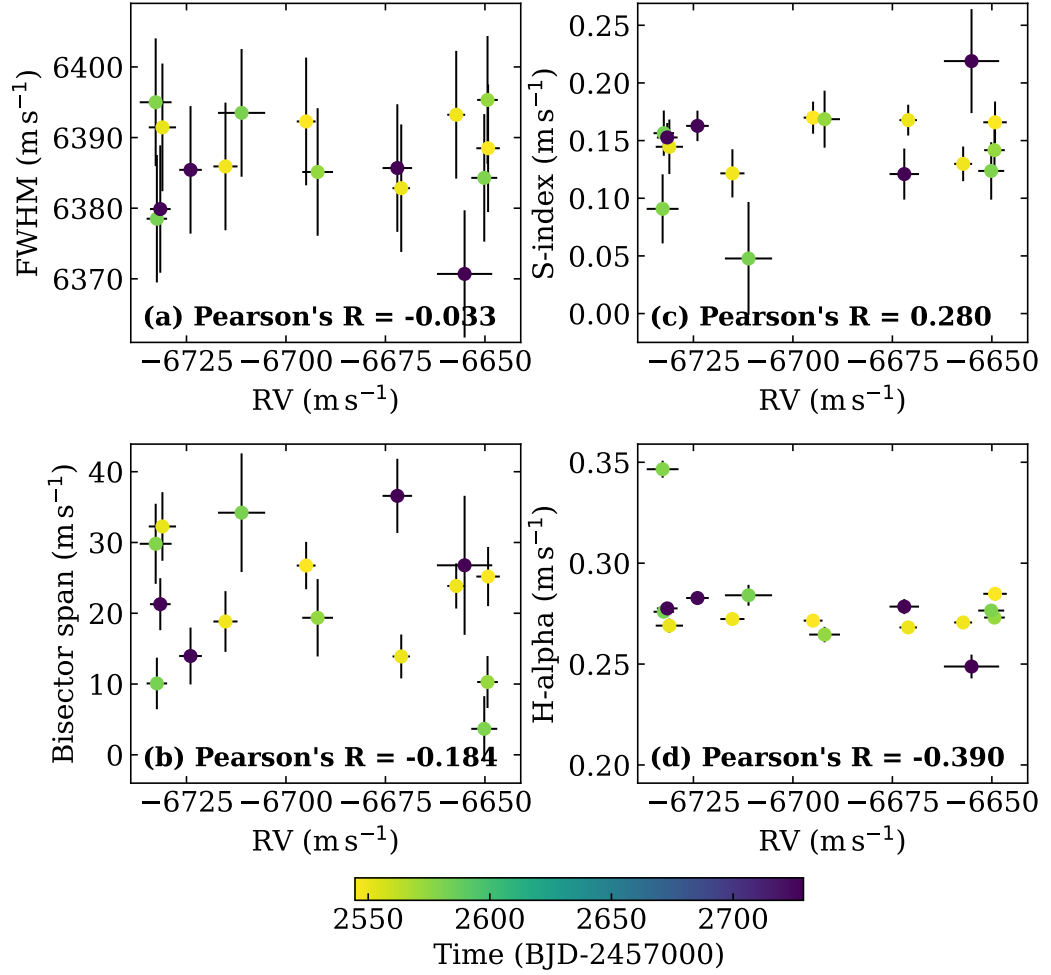


Figure 5.8: HARPS radial velocities plotted against stellar activity indicators: (a), the full-width at half-maximum (FWHM) of the cross-correlation function (CCF); (b), the bisector span of the CCF; (c), the s-index; and (d), h-alpha. The Pearson's R statistic, a measure of correlation strength, is given for each, and no significant correlation is seen. Colour represents the time of observation in Barycentric Julian Date (BJD). All error bars show one standard deviation.

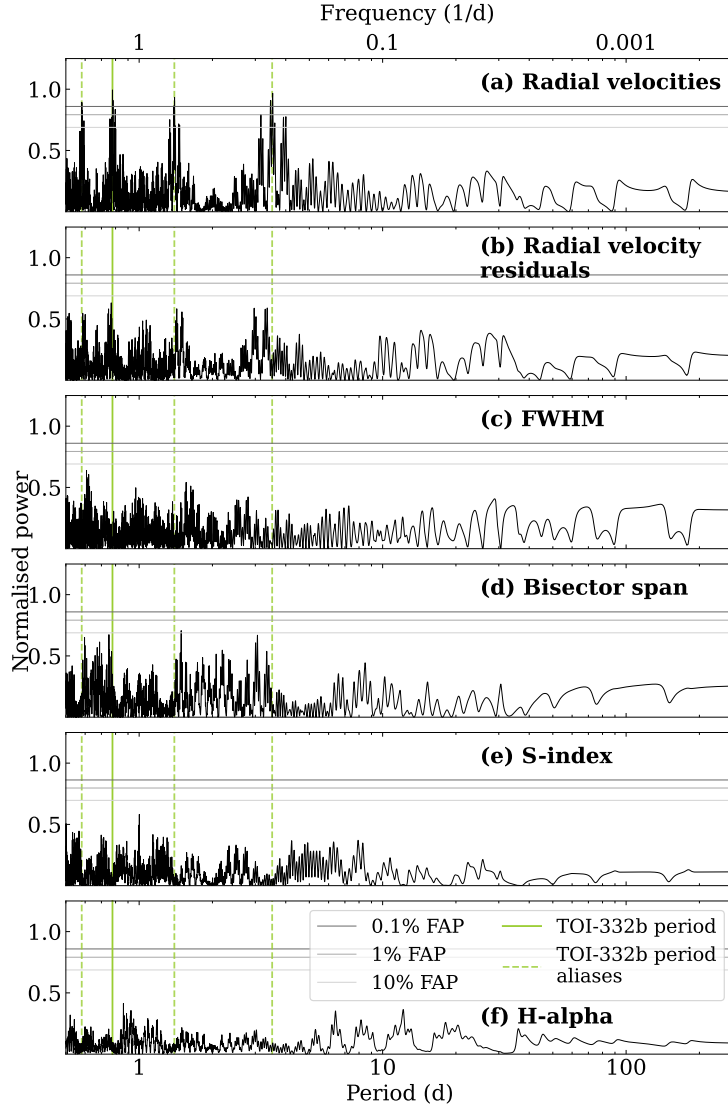


Figure 5.9: Periodograms for the HARPS data. The expected period of TOI-332 b is denoted by a solid vertical line, with the aliases of this period given as dashed lines. The 0.1, 1, and 10 per cent False Alarm Probabilities (FAPs) are shown as solid horizontal lines. The FAPs are calculated using the approximation from Baluev (2008). From top to bottom: (a) the periodogram for the raw radial velocities with a peak above the 0.1 per cent FAP at the expected planetary period; (b) the periodogram for the radial velocity residuals after the best fit model has been removed, showing no further significant peaks; (c)-(f) the periodograms for the stellar activity indicators full-width at half-maximum (FWHM, (c)), bisector span (d), s-index (e), and h-alpha (f), with no significant periodicity shown.

Table 5.2: HARPS radial velocities.

Time (RJD)	RV (ms^{-1})	σ_{RV} (ms^{-1})	FWHM (ms^{-1})	Bisector (ms^{-1})	Contrast
59544.57860093983	-6694.929826	2.367355	6392.280263	26.740961	46.535725
59545.55694768019	-6649.202052	2.955035	6388.488003	25.186455	46.576300
59547.619848280214	-6715.203341	3.033317	6385.913640	18.836766	46.524531
59548.59424903989	-6657.213721	2.254944	6393.227848	23.858339	46.540726
59549.57656604005	-6671.028545	2.193774	6382.831911	13.896637	46.601883
59550.5969126299	-6731.050809	3.426731	6391.442681	32.262988	46.546534
59573.55401741015	-6649.345311	2.595591	6395.329891	10.285248	46.332886
59574.5364167802	-6692.040443	3.867689	6385.126449	19.361261	45.940802
59578.54254523991	-6732.742673	4.000978	6394.998608	29.815993	46.371621
59580.54742026003	-6650.166676	3.263565	6384.294121	3.673344	46.562577
59581.54694268014	-6711.154291	5.927420	6393.494683	34.206176	46.592209
59582.541859869845	-6732.443878	2.576812	6378.507589	10.085069	46.563025
59726.90363577986	-6723.985683	2.841378	6385.428239	13.959675	46.604519
59727.8527730098	-6731.586132	2.598389	6379.880003	21.284249	46.570771
59728.80604236014	-6672.012893	3.704808	6385.685635	36.586262	46.578561
59728.92009065999	-6655.110375	6.944427	6370.694975	26.775309	46.895611

The full HARPS data products are publicly available from the ESO archive, at http://archive.eso.org/wdb/wdb/adp/phase3_main/form.

5.2.3 High resolution imaging

Gemini Zorro

High-angular resolution images of TOI-332 were obtained on 10 October 2019 using the Zorro² speckle instrument on the Gemini-South telescope (Scott et al., 2021). Zorro observes simultaneously in two bands (832 ± 40 nm and 562 ± 54 nm), obtaining diffraction limited images with inner working angles of 0.026 and 0.017 arcsec respectively. The TOI-332 data set consisted of 5 sets of 1000×0.06 s images, which were combined using Fourier analysis techniques, examined for stellar companions, and used to produce reconstructed speckle images (see Howell et al., 2011). The speckle imaging reveals TOI-332 to be a single star with no companions detected within 1.2 arcsec down to contrast limits of $\sim 5 - 7$ mag, shown in Fig. 5.10. At the distance of TOI-332 (220 pc), these angular limits correspond to spatial limits of 4 – 264 au.

VLT NaCo

We collected high-resolution adaptive optics imaging of TOI-332 with VLT/NaCo on 19 June 2019. These near-IR images complement the visible-band speckle data and provide greater sensitivity to late-type bound companions. We collected a sequence of nine images in the Ks filter, each with an integration time of 11 s; the telescope was dithered between

²<https://www.gemini.edu/sciops/instruments/alopeke-zorro/>

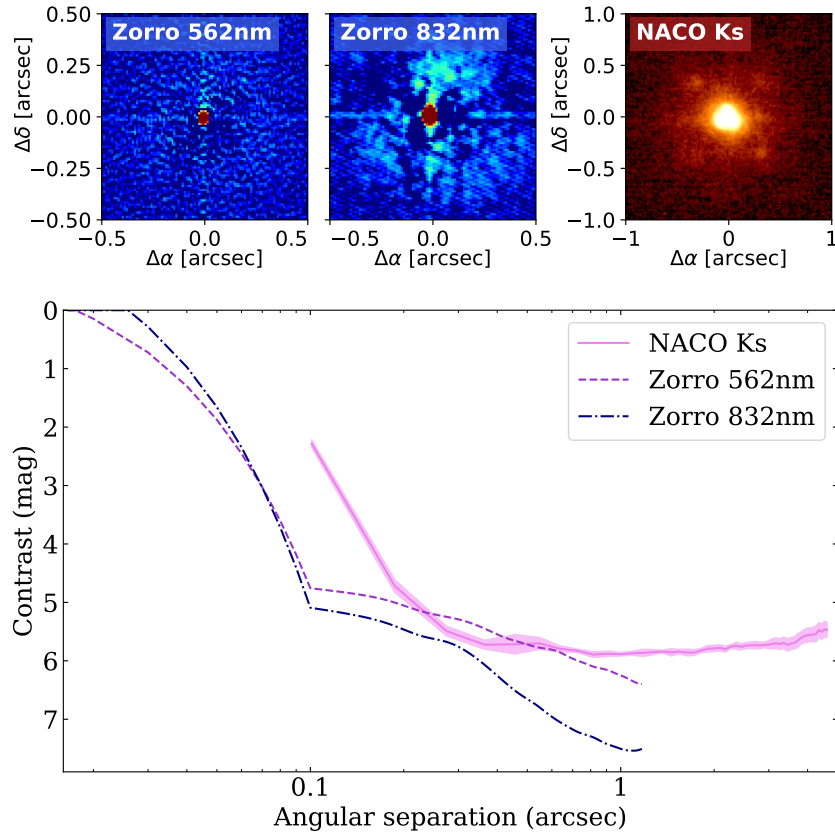


Figure 5.10: A compilation of reconstructed images for the sources of high-resolution imaging described in Sections 5.2.3 and 5.2.3 (top), with their corresponding 5σ contrast curves (bottom). Zorro observes simultaneously in the 562 and 832 nm bands; NACO observes in the near-IR with a Ks filter (labelled). No additional companions are detected.

each exposure and a sky background frame was created by median combining the science frames. We removed bad pixels, flat fielded, subtracted the sky background, and aligned the images on the stellar position before co-adding the sequence. We calculated the sensitivity of these images by injecting fake companions at several position angles and separations, and measuring the significance to which they could be recovered.

The AO image and the sensitivity limits are presented in Fig. 5.10. Some extended PSF structure is seen in the image, but the star is unresolved and no companions are identified. The data reveal that TOI-332 is a single star down to 5.5 mag of contrast, beyond 365 mas from the star, and exclude bright companions beyond 100 mas.

5.3 Spectroscopic analysis and chemical abundances

Here we perform several different methods to measure and derive a range of stellar parameters for TOI-332.

We first used ARES+MOOG to derive spectroscopic stellar parameters (T_{eff} , $\log g$, microturbulence v_{tur} , and $[\text{Fe}/\text{H}]$) following the same methodology as described in Santos et al. (2013); Sousa (2014); Sousa et al. (2021). The latest version of ARES ³ (Sousa et al., 2007, 2015) was used to consistently measure the equivalent widths (EW) of selected iron lines in the combined spectrum of TOI-332. For this, we used the iron line list presented in Sousa et al. (2008). The best spectroscopic parameters are found by converging into ionisation and excitation equilibrium. This process makes use of a grid of Kurucz model atmospheres (Kurucz, 1993) and the radiative transfer code MOOG (Snedden, 1973). We also derived a trigonometric surface gravity using *Gaia* DR3 data following the same procedure as described in Sousa et al. (2021). We find values of: $T_{\text{eff}} = 5251 \pm 71$ K; $\log g = 4.46 \pm 0.04$ c g s; $v_{\text{tur}} = 0.815 \pm 0.069$ km s⁻¹; and $[\text{Fe}/\text{H}] = 0.256 \pm 0.048$ dex.

To estimate the stellar mass and radius we used the calibrations in Torres et al. (2010); because the mass is between 0.7 and 1.3 M_{\odot} we used the correction in Santos et al. (2013). This gives $R_{\star} = 0.87 \pm 0.03 R_{\odot}$ and $M_{\star} = 0.88 \pm 0.02 M_{\odot}$.

Stellar abundances of the elements were then derived using the classical curve-of-growth analysis method assuming local thermodynamic equilibrium. The same codes and models were used for the abundance determinations. For the derivation of chemical abundances of refractory elements we closely followed the methods described in (e.g. Adibekyan et al., 2012, 2015; Delgado Mena et al., 2017). Abundances of the volatile elements, C and O, were derived following the method of Delgado Mena et al. (2021); Bertran de Lis et al. (2015). Since the two spectral lines of oxygen are usually weak and the 6300.3Å line can be contaminated by tellurics or an oxygen airglow, the EWs of these lines were manually measured with the task `splot` in IRAF. All the $[\text{X}/\text{H}]$ ratios are obtained by doing a differential analysis with respect to a high S/N solar (Vesta) spectrum from HARPS. The abundances of these elements are presented in Table 5.3.

Under the assumption that stellar composition serves as a reliable indicator of the disc composition during the planet formation phase, we can determine the mass fraction of the planet building blocks. Following the methodology outlined in Santos et al. (2015, 2017), which uses a simple stoichiometric model and chemical abundances of Fe, Mg, and Si, we computed that the anticipated iron-to-silicates mass fraction is 33.5 ± 3.1 per cent.

To estimate the activity level of TOI-332, we used the HARPS spectra to calculate the $\log R'_{\text{HK}}$ activity index. We co-added all spectra and used ACTIN2⁴ (Gomes da Silva

³The latest version, ARES v2, can be downloaded at <https://github.com/sousasag/ARES>

⁴Available at <https://github.com/gomesdasilva/ACTIN2>.

Table 5.3: Stellar abundances determined by the methods outlined in Section 5.3.

Chemical abundances	Value (dex)
[Ca/H]	0.22 ± 0.06
[Na/H]	0.37 ± 0.07
[Mg/H]	0.26 ± 0.06
[Al/H]	0.34 ± 0.06
[Si/H]	0.24 ± 0.04
[Ti/H]	0.33 ± 0.06
[Ni/H]	0.27 ± 0.04
[O/H]	0.23 ± 0.15
[C/H]	0.22 ± 0.04
[Cu/H]	0.41 ± 0.09
[Zn/H]	0.25 ± 0.05
[Sr/H]	0.33 ± 0.17
[Y/H]	0.18 ± 0.11
[Zr/H]	0.18 ± 0.11
[Ba/H]	0.10 ± 0.09
[Ce/H]	0.20 ± 0.15
[Nd/H]	0.27 ± 0.08

et al., 2018, 2021) to extract the S_{CaII} index. This index was calibrated to the Mt. Wilson scale using the calibration in Gomes da Silva et al. (2021) and converted to $\log R'_{\text{HK}}$ via Noyes et al. (1984), giving $\log R'_{\text{HK}} = -4.831 \pm 0.003$. This can then be used to derive a rotation period (P_{rot}) and age of the star (τ) via the relations in Mamajek and Hillenbrand (2008), giving a rotation period of 35.6 ± 4.6 d and an age of 5.0 ± 2.3 Gyr. This rotation period is approximately twice those obtained by WASP in Section 5.2.1, and so the WASP detection could be the first harmonic rather than the true rotational period.

Moreover, we used the chemical abundances of some elements to derive an alternative value for the age through the so-called chemical clocks (i.e. certain chemical abundance ratios which have a strong correlation for age). We applied the 3D formulas described in Table 10 of Delgado Mena et al. (2019), which also consider the variation in age produced by the effective temperature and iron abundance. The chemical clocks [Y/Mg], [Y/Zn], [Y/Ti], [Y/Si], [Y/Al], [Sr/Ti], [Sr/Mg] and [Sr/Si] were used from which we obtain a weighted average age of 6.3 ± 1.8 Gyr. This age is in agreement (within errors) with the age obtained from the stellar activity and rotation.

As an independent determination of the basic stellar parameters, we performed an analysis of the broadband spectral energy distribution (SED) of the star together with the *Gaia* EDR3 parallax (with no systematic offset applied; see, e.g., Stassun and Torres, 2021), in order to determine an empirical measurement of the stellar radius, following the procedures described in Stassun and Torres (2016b); Stassun et al. (2017, 2018). We pulled the JHK_S magnitudes from *2MASS*, the W1–W3 magnitudes from *WISE*, the $G_{\text{BP}}G_{\text{RP}}$

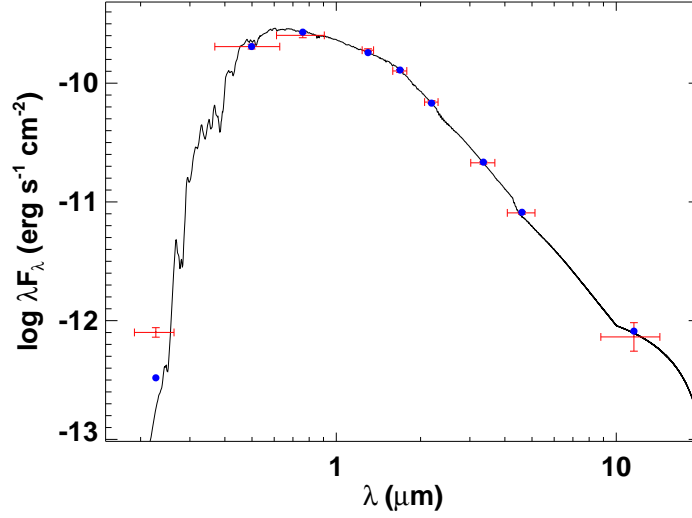


Figure 5.11: Spectral energy distribution of TOI-332. Red symbols represent the observed photometric measurements, where the horizontal bars represent the effective width of the passband. Blue symbols are the model fluxes from the best-fit Kurucz atmosphere model (black).

magnitudes from *Gaia*, and the NUV magnitude from *GALEX*. Together, the available photometry spans the full stellar SED over the wavelength range 0.2–10 μm (see Fig. 5.11).

We performed a fit using Kurucz stellar atmosphere models, with the effective temperature (T_{eff}), surface gravity ($\log g$), and metallicity ($[\text{Fe}/\text{H}]$) adopted from the spectroscopic analysis above. The remaining free parameter is the extinction A_V , which we limited to the maximum line-of-sight value from the Galactic dust maps of Schlegel et al. (1998). The resulting fit (Fig. 5.11) has a reduced χ^2 of 1.3, excluding the *GALEX* NUV flux which indicates a moderate level of activity (see below), and a best fit $A_V = 0.02 \pm 0.02$. Integrating the (unreddened) model SED gives the bolometric flux at Earth, $F_{\text{bol}} = 3.851 \pm 0.045 \times 10^{-10} \text{ erg s}^{-1} \text{ cm}^{-2}$. Taking the F_{bol} and T_{eff} together with the *Gaia* parallax gives the stellar radius, $R_{\star} = 0.923 \pm 0.016 R_{\odot}$. In addition, we can estimate the stellar mass from the empirical relations of Torres et al. (2010), giving $M_{\star} = 0.96 \pm 0.06 M_{\odot}$. These broadly agree with the previous values.

Finally, to obtain another independent check on the fundamental stellar parameters, and following Fridlund et al. (2020) and references therein, we analysed our spectrum with version 5.22 of the spectral analysis package SME (Spectroscopy Made Easy; Valenti and Piskunov, 1996; Piskunov and Valenti, 2017). This IDL based software is used to fit the observations to synthetic stellar spectra calculated with a given set of input parameters and a suitable atmospheric grid. Here, we used the Atlas12 (Kurucz, 2013) grids, together with atomic and molecular line data from VALD (Ryabchikova and Pakhomov, 2015) to calculate

the synthetic spectra. For T_{eff} , we modelled the line wings of the hydrogen alpha line, and derived the surface gravity, $\log g$, from the calcium triplet $\lambda 6102$, 6122 , and 6162 , and the $\lambda 6439$ line. For an independent check, we also modelled the Na I doublet at $5888/89 \text{ \AA}$. We find $T_{\text{eff}} = 5185 \pm 100 \text{ K}$ and $\log g = 4.4 \pm 0.1$, both in agreement with the values determined using ARES+MOOG.

We then fitted a large number of iron lines to obtain the abundances $[\text{Fe}/\text{H}] = 0.4 \pm 0.1 \text{ dex}$; $[\text{Ca}/\text{H}] = 0.47 \pm 0.1 \text{ dex}$; and $[\text{Na}/\text{H}] = 0.47 \pm 0.1 \text{ dex}$.

Following again schemes described in Fridlund et al. (2020) and keeping the macro-turbulent v_{mac} and microturbulent v_{mic} velocities fixed at the empirical values found in the literature (Bruntt et al., 2010; Doyle et al., 2014), we find $v \sin i_{\star} = 1.5 \pm 1.2 \text{ km s}^{-1}$. We can use the rotational period of 35.6 d derived earlier to estimate an equatorial velocity of $\approx 1.24 \text{ km s}^{-1}$ (assuming spin-orbit alignment) which is in agreement with this, supporting the hypothesis of the WASP period being half the true period. We therefore take forward the 35.6 d stellar rotation period into our later analysis.

There are uncertainties on the values for R_{\star} and M_{\star} due to the methods used to derive them: for example, in the calculation of the synthetic models used to fit the observed spectra, and in the Torres et al. (2010) calibration used. Errors on the primary derived stellar parameters (T_{eff} , $\log g$, $[\text{Fe}/\text{H}]$) are taken into account when applying the Torres et al. (2010) calibration, as explained in Santos et al. (2013).

We have used multiple methods to derive stellar parameters to account for unknown systematic effects. Our results from each method are consistent, implying that our stated errors are reasonable and the effect of unknown systematics is small. We note that any systematic errors remaining will propagate into the planetary parameters.

5.4 The joint fit

Using the `exoplanet` package (Foreman-Mackey et al., 2020), we fit the photometry from *TESS* and LCOGT simultaneously with the RVs from HARPS. `exoplanet` utilises the light curve modelling package `Starry` (Luger et al., 2019), `PyMC3` (Salvatier et al., 2016), and `celerite` (Foreman-Mackey et al., 2017). For consistency, all timestamps were converted to the same time system, that used by *TESS*, i.e. $\text{BJD} - 2457000$ (BJD-TDB). All prior distributions set on the parameters fit in this model are given in Table 5.4.

The photometric flux is normalised by dividing the full individual light curves by the median of their out-of-transit points and subtracting unity to produce a lightcurve with out-of-transit flux of zero. No further detrending is deemed to be necessary for either the LCOGT or *TESS* data, and so none is included in the joint fit.

To model the planetary transits, we use a limb-darkened transit model utilising the

Table 5.4: Prior distributions used in our joint fit model, fully described in Section 5.4, and the fit values resulting from the model. The priors are created using distributions in PyMC3 with the relevant inputs to each distribution described in the table footer. The fit values are given as the median values of the samples, and the uncertainties as the 16th and 84th percentiles. Further (derived) system parameters can be found in Table 5.5.

Parameter	(unit)	Prior Distribution	Fit value
Planet			
Period P	(days)	$\mathcal{U}(0.767, 0.787)$	0.777038 ± 0.000001
Reference time of midtransit t_c	(BJD-2457000)	$\mathcal{U}(2062.4439, 2062.4459)$	$2062.4447^{+0.0006}_{-0.0005}$
$\log(R_p)$	(R_\odot)	$\mathcal{U}(-4.6863, -2.6863)^*$	$-3.53^{+0.05}_{-0.04}$
Eccentricity e		0 (fixed)	-
Argument of periastron ω	($^\circ$)	0 (fixed)	-
Star			
Mass M_\star	(M_\odot)	$\mathcal{N}_\mathcal{B}(0.88, 0.02, 0.0, 3.0)$	0.88 ± 0.02
Radius R_\star	(R_\odot)	$\mathcal{N}_\mathcal{B}(0.87, 0.03, 0.0, 3.0)$	$0.87^{+0.03}_{-0.02}$
Photometry			
$TESS_{S1,2}$ mean		$\mathcal{N}(0.0, 1.0)$	0.00006 ± 0.00001
$TESS_{S28}$ mean		$\mathcal{N}(0.0, 1.0)$	0.00008 ± 0.00002
LCO_1 mean		$\mathcal{N}(0.0, 1.0)$	0.00050 ± 0.00007
LCO_2 mean		$\mathcal{N}(0.0, 1.0)$	0.00049 ± 0.00007
LCO_3 mean		$\mathcal{N}(0.0, 1.0)$	0.0008 ± 0.0001
LCO_4 mean		$\mathcal{N}(0.0, 1.0)$	0.00051 ± 0.00006
LCO_5 mean		$\mathcal{N}(0.0, 1.0)$	0.00053 ± 0.00009
LCO_6 mean		$\mathcal{N}(0.0, 1.0)$	0.0005 ± 0.0001
HARPS RVs			
$\log(K)$		$\mathcal{U}(0.0, 10.0)$	3.77 ± 0.02
Offset	(m s^{-1})	$\mathcal{U}(-6702, -6682)$	-6692.4 ± 0.8
$\log(\text{Jitter})$		$\mathcal{N}(2.193774^\dagger, 5.0)$	$-1.1^{+1.1}_{-1.8}$

Distributions:

$\mathcal{N}(\mu, \sigma)$: a normal distribution with a mean μ and a standard deviation σ ;

$\mathcal{N}_\mathcal{B}(\mu, \sigma, a, b)$: a bounded normal distribution with a mean μ , a standard deviation σ , a lower bound a , and an upper bound b (bounds optional);

$\mathcal{U}(a, b)$: a uniform distribution with a lower bound a , and an upper bound b .

Prior values:

* equivalent to $0.5(\log(D)) + \log(R_\star) \pm 1$ where D is the transit depth (ppm multiplied by 10^{-6}) and R_\star is the mean of the prior on the stellar radius (R_\odot), and -1 computes the lower bound while $+1$ computes the upper bound;

† equivalent to the log of the minimum error on the HARPS data (m s^{-1}).

quadratic limb-darkening parameterisation in Kipping (2013) and a Keplerian orbit model. We put Gaussian priors informed by the ARES+MOOG values on the stellar radius R_\star and the stellar mass M_\star .

The Keplerian orbit model is parameterised for the planet in terms of the orbital period P , the time of a reference midtransit t_c , the eccentricity e , and the argument of periastron ω . In an earlier iteration of this model, we found the eccentricity of TOI-332 b to be consistent with 0 (with the 95 per cent confidence interval for the eccentricity being 0 to 0.15), and so fix e and ω to 0 in the final model presented here. A close-to-zero eccentricity is also expected given the very short orbital period. These parameters are then input into light curve models created with *Starry*, alongside further parameters which are planetary radii R_p , the time series of the data t , and the exposure time t_{exp} of the instrument.

Individual light curve models are created for the LCOGT data, the combined *TESS* S1 and S2 data, and the *TESS* S28 data (S28 is kept separate to S1 and S2 due to differing cadence and exposure time of the S1 and S2 data compared to the S28 data). We use values from the *TESS* SPOC pipeline (Li et al., 2019) to estimate the placement of wide, uninformative uniform priors for the epoch, period, and radius of TOI-332 b. For each lightcurve, we put a Gaussian prior on the offset with a mean of zero and standard deviation of one.

To fit the HARPS RVs, we use DACE⁵ with a simple Keplerian model to estimate prior values for the systematic RV offset and the semi-amplitude of the RV signal K . We set wide, uninformative uniform priors on K and the offset. We also incorporate a separate jitter term with a wide Gaussian prior, the mean of which is the log of the minimum error on the HARPS data. This term encapsulates any uncharacterised signal or noise that is perceived as white noise in the RV data, for example instrumental effects and short-scale stellar activity. As the RV data does not show any significant stellar activity by visual inspection and in the stellar activity indicators (see Figs. 5.8 and 5.9), and does not exhibit any long-term trends, we do not perform any further detrending to it. This completes the joint fit model.

We use *exoplanet* to maximise the log probability of the model. The fit values that this optimisation obtains are then used as the starting point of the PyMC3 sampler, which draws samples from the posterior using a variant of Hamiltonian Monte Carlo, the No-U-Turn Sampler (NUTS). From examination of the chains from earlier test runs of the model, we use 5 chains of 50000 steps, 1000 steps of which are discarded as burn-in. To test for non-convergence, we calculate the rank-normalised split- \hat{R} statistic (Vehtari et al., 2021) for each parameter. $\hat{R} \approx 1.0$ for all parameters, implying convergence. We present our best fit parameters for the TOI-332 system from this joint fit in Table 5.5.

⁵The DACE platform is available at <https://dace.unige.ch>

Table 5.5: Stellar parameters of TOI-332, and transit, orbital, and physical parameters of TOI-332 b (further parameters from the joint fit model can be found in Table 5.4).

Parameter	(unit)	Value	Source
Host star			
Distance to Earth	(pc)	222.85 ± 3.69	Gaia DR3
Effective temperature T_{eff}	(K)	5251 ± 71	ARES+MOOG
Spectral type	-	K0V	Pecaut and Mamajek (2013)
Surface gravity $\log g$	(c g s)	4.46 ± 0.04	ARES+MOOG
Metallicity [Fe/H]	(dex)	0.256 ± 0.048	ARES+MOOG
Stellar radius R_{\star}	(R_{\odot})	$0.87^{+0.03}_{-0.02}$	Joint fit
Stellar mass M_{\star}	(M_{\odot})	0.88 ± 0.02	Joint fit
Rotational velocity $v \sin i_{\star}$	(km s^{-1})	$< 1.5 \pm 1.2$	SME
Chromospheric activity index $\log R'_{\text{HK}}$	-	-4.831 ± 0.003	ACTIN2
Rotation period P_{rot}	(days)	35.6 ± 4.6	$\log R'_{\text{HK}}$ + Mamajek and Hillenbrand (2008)
Age τ	(Gyr)	5.0 ± 2.3	$\log R'_{\text{HK}}$ + Mamajek and Hillenbrand (2008)
Planet			
Period P	(days)	0.777038 ± 0.000001	Joint fit
Full transit duration T_{dur}	(hours)	1.52 ± 0.03	Joint fit (derived)
Reference time of midtransit t_c	(BJD-2457000)	$2062.4447^{+0.0006}_{-0.0005}$	Joint fit
Radius R_p	(R_{\oplus})	$3.20^{+0.16}_{-0.11}$	Joint fit
Planet-to-star radius ratio R_p/R_{\star}	-	0.0341 ± 0.0009	Joint fit (derived)
Impact parameter b	-	$0.25^{+0.13}_{-0.15}$	Joint fit
Inclination i	($^{\circ}$)	$86.4^{+2.3}_{-2.0}$	Joint fit
Eccentricity e	-	0 (fixed)	Joint fit
The argument of periastron ω	($^{\circ}$)	0 (fixed)	Joint fit
Radial velocity semi-amplitude K	(ms^{-1})	43 ± 1	Joint fit
Mass M_p	(M_{\oplus})	57.2 ± 1.6	Joint fit (derived)
Bulk density ρ	(g cm^{-3})	$9.6^{+1.1}_{-1.3}$	Joint fit (derived)
Semi-major axis a	(AU)	0.0159 ± 0.0001	Joint fit (derived)
System scale a/R_{\star}	-	$3.94^{+0.11}_{-0.12}$	Joint fit (derived)
Equilibrium temperature* T_{eq}	(K)	1871^{+30}_{-25}	Joint fit (derived)

*Equilibrium temperature is calculated assuming an albedo of zero.

5.5 Results and discussion

The results of our joint fit model show that, with an orbital period of 0.777038 ± 0.000001 d, TOI-332 b is an “Ultra-Short Period” (USP) planet, defined as a planet with $P_{\text{orb}} < 1$ d (Winn et al., 2018). The host star, TOI-332, is a K0 dwarf with a mass of $0.88 \pm 0.02 M_{\odot}$ and a radius of $0.87^{+0.03}_{-0.02} R_{\odot}$. Assuming an albedo of zero, the proximity of the planet to the star gives the planet an equilibrium temperature of 1871^{+30}_{-26} K; it is highly irradiated, receiving approximately 2400 times the instellation of the Earth per unit area.

TOI-332 b has a mass of $57.2 \pm 1.6 M_{\oplus}$, more than half the mass of Saturn, yet a radius of $3.20^{+0.16}_{-0.11} R_{\oplus}$, smaller than that of Neptune. With a density of $9.6^{+1.1}_{-1.3} \text{ g cm}^{-3}$, it is one of the densest planets of those with the size of Neptune or greater found thus far (Fig. 5.12). These parameters place TOI-332 b deep in the Neptunian desert (Fig. 5.1).

Taking all of this into account, TOI-332 b is a very interesting addition to our current Neptunian desert discoveries and a case study to test planet formation theory.

5.5.1 Interior structure

As seen in Fig. 5.12, TOI-332 b occupies a unique and unpopulated spot in the mass-radius (M-R) diagram. Its mass and radius suggest a composition that is dominated by refractory materials, potentially more similar to that of terrestrial planets.

To put limits on the possible composition of TOI-332 b, we use a layered interior model similar to those used in Dorn et al. (2017) and Armstrong et al. (2020). This model consists of up to four layers including an iron core, a silicate mantle, a water layer, and a H-He atmosphere. For these layers, we solve the standard structure equations to estimate the possible ranges of H-He mass fractions. We note, however, that for such high mass planets, layers might not be as distinct as assumed here (e.g., Helled and Stevenson, 2017b; Bodenheimer et al., 2018). Overall, the planet is found to consist of 30 per cent iron core, 43 per cent rock mantle, 27 per cent water, and a negligible H-He envelope.

To constrain the H-He mass, we investigate the extreme situation of a planet without water and compare it with a planet where the water abundance is allowed to vary freely. For these cases, we construct structure models that reproduce the measured mass and radius of TOI-332 b. Moreover, we assume host star elemental abundances.

We find that even if TOI-332 b had no water, the H-He mass fraction would be only $1.8^{+0.6}_{-0.5} \%$. In the water-containing model, the H-He mass fraction is $\log(M_{\text{atm}}/M_p) = -6.7 \pm 3.2$, well below 0.1 per cent. We can therefore conclude that, unless the planet is devoid of water, the atmospheric mass of TOI-332 b is very small.

Typically, planets with comparable masses to TOI-332 b are expected to be H-He dominant in composition, and terrestrial planets are not expected to reach several tens of

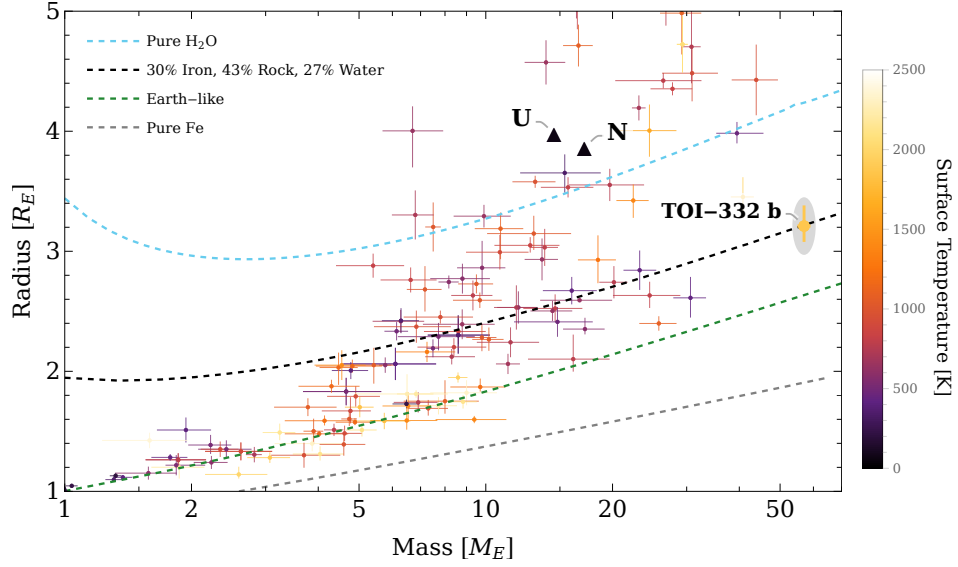


Figure 5.12: Mass-radius diagram of the exoplanets in the Otegi catalog (Otegi et al., 2020a). The color of the planets indicates their equilibrium temperature. The dashed blue, green, and gray line show the mass-radius relation for a pure water, an Earth-like, and a pure iron composition at TOI-332 b’s equilibrium temperature (1869.4 K), respectively. Uranus and Neptune are shown as black triangles.

Earth masses. Planetary embryos are expected to accrete only a few to ~ 20 Earth masses of heavy elements before the onset of rapid gas accretion (e.g., Pollack et al., 1996; Lambrechts et al., 2014; Piso et al., 2015), resulting in a large envelope. However, with such a large core mass and little envelope, the existence of TOI-332 b requires further explanation, perhaps having lost an initial envelope, or having managed to avoid core-accretion. We explore several scenarios below.

5.5.2 Co-orbital bodies

To try and explain the apparently excessive core mass of TOI-332 b, we first tested an alternative co-orbital configuration of two planets which may mimic the appearance of a single, more massive planet, and compare its evidence against the current one-planet scenario.

Co-orbital exoplanets (pairs of planets trapped in 1:1 resonances) are dynamically stable under very soft conditions (Laughlin and Chambers, 2002), and several formation mechanisms have already been proposed for these configurations (see e.g., Beaugé et al., 2007; Namouni and Morais, 2017; Leleu et al., 2019). However, no co-orbital exoplanets have yet been found despite several efforts (e.g., Janson, 2013; Hippke and Angerhausen, 2015; Ford and Holman, 2007; Madhusudhan and Winn, 2009; Lillo-Box et al., 2018b,a),

although different candidates have already been proposed (e.g., Lillo-Box et al., 2014b; Boyajian et al., 2016; Lillo-Box et al., 2020).

In the particular case of TOI-332 b, we explore the scenario where this planet is actually a pair of planets in 1:1 resonance where the lighter planet transits the host star while the more massive component does not. Assuming a low eccentricity scenario, we can approximate the sum of two Keplerians with the same periodicity as a single Keplerian. We might attribute the mass of the more massive (non-transiting) component to the only component that we see transiting the host star. This will imprint specific features in the radial velocity data that are testable through available techniques. In particular, in order to test this scenario, we apply the technique described in Leleu et al. (2017) (a generalisation of the technique proposed by Ford and Gaudi (2006)), which combines the transit and radial velocity information to infer time lags between the time of transit and the radial velocity phase. This technique is based on the modelling of the radial velocity data assuming the time of conjunction and period derived from the transit modelling. In this case, we assume the value $t_c = 2459062.444852864292$ and $P = 0.77703814$ days, obtained only by modelling the photometry from *TESS* and *LCOGT*. We apply Equation 18 in Leleu et al. (2017), which includes the RV semi-amplitude K , the orbital configuration parameters (c and d), and the additional parameter α , a measure of the mass imbalance between the Lagrangian regions L_4 and L_5 .

Here we test four different models: two assuming only one planet in the orbit (one assuming a circular orbit, “1p(c)”, the other including the possibility of a slightly eccentric orbit, “1p”), and two including the co-orbital scenario (again, one assuming a circular orbit, “1p(c)T”, and the other leaving the eccentricity as a free parameter, “1pT”). We use the implementation of Goodman and Weare (2010)’s affine invariant Markov chain Monte Carlo (MCMC) ensemble sampler *emcee* (Foreman-Mackey et al., 2013) to sample the posterior probability distribution of each of these parameters. The MCMC chains are subsequently used to estimate the Bayesian evidence ($\ln \mathcal{Z}_i$) of the models using the *perrakis* implementation (Díaz et al., 2016). The full set of priors and posteriors from each of the four co-orbital scenario models are presented in Table 5.6.

The results of our analysis show the model with just one planet in circular orbit as the most favourable model based on the current dataset ($\Delta \ln \mathcal{Z} > +7$ against the other more complex models). Consequently we can conclude that the current dataset does not support the presence of an additional co-orbital planet, hence confirming that all the mass at this periodicity is accumulated into the transiting body. For the simpler co-orbital model with circular orbit, we obtain $\alpha = -0.031^{+0.032}_{-0.031}$, hence compatible with zero (i.e., no mass imbalance between L_4 and L_5 and so potentially no co-orbitals) at the 1σ level.

Table 5.6: Co-orbital hypothesis results.

Parameter	Priors	Posteriors			
		1p(c)	1p	1p(c)T	1pT
P (days)	$\mathcal{G}(0.77703817, 10^{-8})$	$0.77703792 \pm 9.9 \times 10^{-7}$	$0.77703812 \pm 9.9 \times 10^{-7}$	Fixed	Fixed
t_0 (BJD-2457000)	$\mathcal{G}(2062.444852864292, 4 \times 10^{-7})$	$2062.44496903 \pm 4 \times 10^{-7}$	$2062.44496903 \pm 4 \times 10^{-7}$	Fixed	Fixed
K_b (m s $^{-1}$)	$\mathcal{U}(0.0, 100.0)$	$43.1^{+1.2}_{-1.2}$	$43.2^{+1.2}_{-1.2}$	$43.1^{+1.2}_{-1.2}$	$43.2^{+1.2}_{-1.2}$
e	$\mathcal{U}(0.0, 1.0)$	Circular	$0.025^{+0.021}_{-0.016}$	-	-
ω (deg.)	$\mathcal{U}(0.0, 359.99)$	Circular	215^{+38}_{-72}	-	-
c	$\mathcal{G}(0.0, 0.05)$	-	-	-	0.028 ± 0.025
d	$\mathcal{G}(0.0, 0.05)$	-	-	-	0.014 ± 0.026
δ_{HARPS} (km s $^{-1}$)	$\mathcal{U}(-15.0, -0.0)$	$-6.69198^{+0.00091}_{-0.00088}$	$-6.69194^{+0.00088}_{-0.00088}$	$-6.69209^{+0.00091}_{-0.00089}$	$-6.69184^{+0.00091}_{-0.00090}$
σ_{HARPS} (m s $^{-1}$)	$\mathcal{U}(0.0, 5.0)$	$1.6^{+1.3}_{-1.0}$	$1.33^{+1.3}_{-0.91}$	$1.5^{+1.4}_{-1.0}$	$1.4^{+1.4}_{-1.0}$
$\ln \mathcal{Z}$		55.5	52.7	48.9	48.4

Posteriors are given for the following models, as described in Section 5.5.2:

“1p(c)”, one planet on a circular orbit;

“1p”, one planet with the possibility of an eccentric orbit;

“1p(c)T”, a co-orbital scenario with a circular orbit;

“1pT”, a co-orbital scenario with the possibility of an eccentric orbit.

The period and time of conjunction posteriors coincide with the priors as they are much more constrained by the transit modelling. All posteriors for the systemic velocity agree within 1σ .

5.5.3 Evolution under XUV-driven escape

We then test whether a TOI-332 b-like planet could be reproduced by stripping an initial accreted envelope through X-ray and extreme-ultraviolet (EUV; together, XUV) driven escape.

We performed simulations on the evaporation history of TOI-332 b by taking into account the range of possible XUV emission histories of the star motivated by its stellar parameters. We fitted the star’s inferred spin period of 35.6 ± 4.6 days with the rotational evolution models of Johnstone et al. (2021) and estimated a gyrochronological age of $5.3^{+1.5}_{-1.4}$ Gyr.

By field age, the initial spread in stellar rotation periods have largely converged to a single track, leaving their histories degenerate, i.e., we cannot tell which history TOI-332 followed. We thus considered three spin histories in order to sample the diversity of possible X-ray activity pasts experienced by the planet: the *low*, *medium*, and *high* activity scenarios, which represent the model’s 5th, 50th, and 95th percentiles in the distribution of rotation periods at any given age. These rotational histories are shown in Fig. 5.13 (left hand panel), together with the star’s current place along these tracks. The corresponding XUV luminosity tracks for these scenarios are shown in Fig. 5.13 (right hand panel).

We then simulated the evaporation history of TOI-332 b using the *photoevolver* code⁶ (Fernández Fernández et al., 2023). For this analysis, we adopted the full hydrodynamic model of Kubyshkina et al. (2018) (and the interpolation routine of Kubyshkina and Fossati (2021)) to calculate the mass loss rates.

Taking the results of the interior structure characterisation, we assumed that the planet is currently a bare core with no gaseous envelope. We can thus estimate an upper limit on the initial envelope mass fraction assuming that it has just finished evaporating. Over this upper limit, the envelope would fail to evaporate in the lifetime of the planet, and this would be inconsistent with the planet’s current structure.

We achieved this by adding a tiny amount of gas to the planet, equivalent to 0.01 per cent of its total mass (such that it is completely evaporated within one simulation time step), and evolved this tenuous atmosphere backwards in time to the age of 10 Myr. We repeated this process using each of the three XUV emission scenarios, and plot the results in Fig. 5.14.

We find that the possible evaporation histories for TOI-332 b based on these scenarios lead to a narrow range of upper limits on the initial envelope mass fraction, between 3 and 6 per cent. We thus find that TOI-332 b starting out as a Jupiter-sized planet is inconsistent with photoevaporation as the only mechanism for mass loss.

⁶The evaporation evolution code is available on GitHub at <https://github.com/jorgefz/photoevolver>

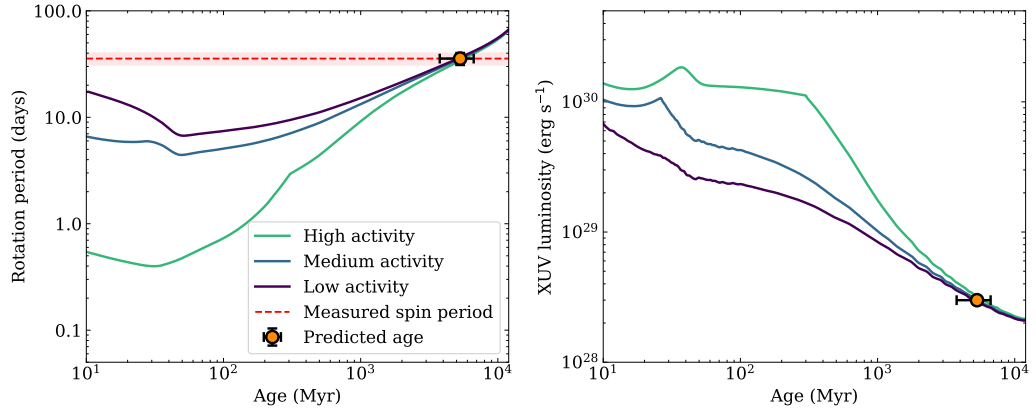


Figure 5.13: **Left panel:** plot of rotation period against age showing rotational evolution models by Johnstone et al. (2021), with high, medium, and low activity tracks for a $0.9 M_{\odot}$ star. Its measured rotation period is shown as a dashed red line, with the uncertainty as a shaded region. The age estimated with gyrochronology is plotted as an orange circle. **Right panel:** plot of XUV luminosity against age showing the corresponding XUV evolution tracks to the models on the left panel, as well as the predicted XUV luminosity based on its rotation period. The models were calculated using the methods described in Section 5.5.3.

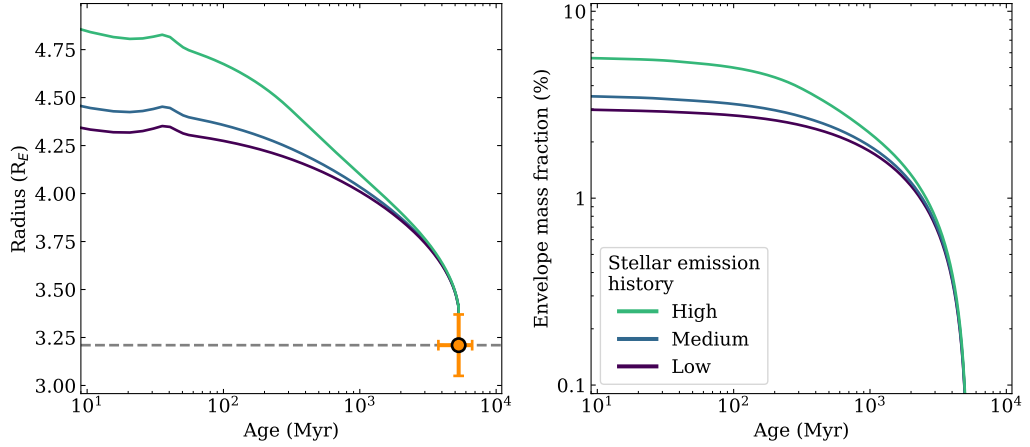


Figure 5.14: **Left panel:** plot of planet radius against age showing the evolution of the radius of TOI-332 b using the three stellar XUV emission histories described in Section 5.5.3. **Right panel:** plot of envelope mass fraction against age showing the evolution of the past envelope mass of TOI-332 b following the left panel.

5.5.4 Other formation scenarios

It is clear that if TOI-332 b originally had a Jupiter-like envelope as we would expect for a core of this size, photoevaporation could not have been the sole mechanism responsible for the removal of most of its atmosphere. So we can theorise other scenarios that could have caused this. An initially large envelope may have been removed by high-eccentricity migration and subsequent tidal thermalisation (e.g. Ivanov and Papaloizou, 2004; Vick and Lai, 2018; Wu, 2018; Vick et al., 2019). Alternatively, the atypical composition of TOI-332 b could be the result of a giant impact between two gas giants followed by efficient removal of the gaseous atmosphere (e.g. Liu et al., 2015; Emsenhuber et al., 2021; Ogihara et al., 2021). Finally, runaway accretion could have just been avoided entirely by, for example, gap opening in the protoplanetary disk (e.g. Crida et al., 2006; Duffell and MacFadyen, 2013; Lee, 2019).

However, we do not think it is currently possible to say which, if any, of these formation scenarios created TOI-332 b, though future observations may aid us in this.

5.5.5 Orbital decay rate

As TOI-332 b is both unusually massive and close to its host star, it may be one of the most well-placed non-gas giant planets for an orbital decay rate study.

We follow the method outlined in Jackson et al. (2023) to calculate the orbital decay rate, dP/dt . In short, we use the following equation (Goldreich and Soter, 1966; Ogilvie, 2014):

$$\frac{dP}{dt} = -\frac{27\pi}{2Q'_*} \left(\frac{M_p}{M_*} \right) \left(\frac{R_*}{a} \right)^5 \quad (5.1)$$

where $Q'_* = 3Q_*/2k_2$. Q_* is the tidal quality factor, Q'_* is the reduced tidal quality factor, and k_2 is the dimensionless quadrupolar Love number.

We do not consider the dynamical tide within the convective zone, only the equilibrium tide, as the orbital period of the planet is much less than twice the stellar rotation period. This allows us to calculate Q'_* as defined in Strugarek et al. (2017), which requires a value for the depth of the convective zone⁷.

We find $Q'_* = 8 \times 10^6$, resulting in $dP/dt = -1.05 \times 10^{-12}$. We use the decay rate together with the period to estimate the length of time it would take for the orbit to decay completely (i.e. reaching a period of zero) as 2.0 Gyr. We note that this is a likely upper-estimate of the decay timescale, as it assumes a constant rate of decay, ignoring any effects that may alter this (e.g. stellar wind and stellar evolution). Our method also does not

⁷We estimate the radius and mass of the stellar core using <http://www.astro.wisc.edu/~townsend/static.php?ref=e-z-web>.

take into account the structure of the planet, which some other more complex treatments of tidal effects do (e.g. Henning and Hurford, 2014; Clausen and Tilgner, 2015; Brasser et al., 2019), but this is beyond the scope of this paper.

We can also estimate boundaries on this decay timescale by assuming upper and lower limits on Q'_* of 10^5 and 10^8 , resulting in timescales of 0.25 and 250 Gyr respectively. Even the shortest decay timescale of 0.25 Gyr is several magnitudes longer than the current decay timescale estimate for WASP-12 b of 3.16 ± 0.10 Myr, to date the only planet we have confidently detected an orbital decay for (Maciejewski et al., 2016; Patra et al., 2017, 2020; Yee et al., 2020; Wong et al., 2022). Thus we conclude that measuring the orbital decay of TOI-332 b is not going to be possible over a realistic span of time.

5.5.6 Future observation prospects

TOI-332 b is undoubtedly an unusual and unique planet, and further observations will be needed to deduce more about its formation and evolutionary history and its current composition.

The Rossiter-McLaughlin (RM) effect allows us to measure the sky-projected obliquity of a system, and is important for constraining formation scenarios: disk-migration is expected to conserve alignment between the angular momentum of a disk and planetary orbits, but misalignment could imply, for example, planet-planet/planet-star scattering, high-eccentricity migration, or tidal disruption. If TOI-332 b lacks an atmosphere due to reduced gas accretion through gap opening, it should align to the stellar spin axis, but if it is misaligned, it might imply a more violent history has removed an initial envelope - though at such a short period, there is the possibility that tides might cause realignment even if the orbit and stellar spin axis began misaligned. We can predict a RM semi-amplitude of approximately 2.1 m s^{-1} (Triaud, 2018); though this signal is small, smaller RM amplitudes have been measured (e.g. Winn et al., 2010; Bourrier and Hébrard, 2014) and are obtainable with high-precision spectrographs like HARPS/HARPS-N and ESPRESSO, and methods such as the Rossiter-McLaughlin effect Revolutions (RMR) technique (Bourrier et al., 2021).

There is evidence from other USP planet discoveries that they often have companions with periods out to 50 days (Sanchis-Ojeda et al., 2014), and for a system like TOI-332 with $a/R_\star < 5$ they'd be expected to have a minimum mutual inclination of $5 - 10$ deg (Dai et al., 2018). We find no evidence of a companion in our current data, photometric or spectroscopic; further long-term monitoring of this system would be needed to discover or rule out a companion, and this may also help narrow down formation scenarios for this system.

Characterising potentially unusual atmospheres and surfaces of highly-irradiated rocky worlds is an exciting prospect. While the predicted atmospheric mass fraction of

TOI-332 b is small, the high equilibrium temperature of TOI-332 b may lead to evaporation of volatiles and formation of a secondary atmosphere that could contain core materials. The composition of such an atmosphere could be determined with JWST. Additionally, JWST could be used to obtain a phase curve of TOI-332 b, which would constrain its dayside and nightside temperatures and any phase offset, its Bond albedo, and heat recirculation efficiency. With little atmosphere we would expect high temperature contrast and poor recirculation, and may be able to distinguish between different surface composition scenarios.

5.6 Conclusion

We present here the discovery and characterisation of a new planet in the TOI-332 system. We use photometry from two *TESS* sectors at 30 min cadence and one sector at 2 min cadence, plus six LCOGT transit events. There is further photometry from PEST and WASP-South, but this was not included in the final fit due to the ambiguity of the transit detections. The photometric data were modelled jointly with 16 RV data points from the HARPS spectrograph. Multiple sources of high-resolution imaging confirm that the star is single with no unresolved companions.

The planet TOI-332 b is on an ultra-short period of 0.78 days, with a radius smaller than Neptune but an anomalously large mass of more than half that of Saturn, making it one of the densest known Neptune-sized planets discovered thus far. It is located deep within the Neptunian desert, and is one of only a handful of planets that have been found there, being one of even fewer to have a precise mass determination. Using a four layer model consisting of an iron core, silicate mantle, water, and a H-He envelope, interior structure characterisation determines that it likely possesses a negligible H-He envelope.

This unusual planet tests what we currently understand about planet formation; how such a giant core exists without a gaseous envelope remains an unanswered question. We determine that photoevaporation would be insufficient on its own in removing a Jupiter-like envelope, and we instead posit high-eccentricity migration or giant impacts as possible mechanisms for stripping the initial envelope from TOI-332 b. Alternatively, a mechanism like disc-gap opening could have led it to avoid gas accretion in the first instance. Further observations are needed to potentially disentangle TOI-332 b's formation history and current characteristics.

Chapter 6

Conclusion

In this thesis, I have presented the detection and characterisation of several exoplanetary systems, as well as the first steps towards creating, observing, and characterising a sample of planets within the Neptunian desert, all primarily using photometry from *TESS* and spectroscopy from HARPS. Here, I summarise the main results, and discuss the future outlook of this work and exoplanetary research going forward.

6.1 Thesis summary

In Chapter 3, I presented the discovery and characterisation of a new three planet system, with two of its planets released as *TESS* Objects of Interest. From a joint fit of *TESS*, *Spitzer*, LCOGT and NGTS photometry, and HARPS and HIRES RVs, I confirm the planetary nature of the TOIs and discover a third, non-transiting planet in the RV data. I obtain precise values for the periods, masses and radii of the transiting planets, and the period and minimum mass of the non-transiting planet. The system is explored in the context of its evolution under XUV-driven photoevaporation, as the two transiting planets straddle the radius gap. From a composition analysis, it is found the planet above the gap retains a H-He envelope, while the other is bare. TOI-431 is an exciting system for future observations from several perspectives: the atmospheric composition of TOI-431 d, the phase curve of TOI-431 b, and the comparative photoevaporation of them both.

In Chapter 4, I started work in creating and following up a sample of TOIs within the Neptunian desert, a first step towards a population study of the planets within the desert. From an initial list of TOIs within the desert boundaries that could be observed by HARPS to obtain their masses, I then hand-vetted and ranked the targets in this list to whittle it down to a sample that could have all necessary spectroscopic observations completed (i.e., a precise mass determined, a false positive confirmation, or a conclusion that a precise mass could not

be determined with HARPS) within the time span of a 2 year large HARPS program. I then directed these observations, continually monitoring the incoming data. First-look analysis was performed for each target, and a determination of the nature of any signal in the RV data was made. Out of a full sample of 73 TOIs, 47 had already been followed up, so I followed up the remaining 26. Out of these 26, 3 are already published as planets, a further 4 are in preparation to be published, and another 1 has its planetary nature confirmed. 3 targets are complicated by stellar activity and require dedicated detrending, though the planet signal is determined to likely exist in the data. 2 were confirmed false positives as they are double-lined spectroscopic binaries. The remaining 13 targets were determined to be null results, the RV data flatlining and in some cases, likely false positives due to stellar activity signals. I then examined the sample within the context of the Neptunian desert, noting where TOI release values can be unreliable and that mass estimations cannot be trusted. I finished by listing the extensive future work needed to finish this study of the Neptunian desert, and note that this work is going to be a positive contribution to that.

In Chapter 5, I performed the full analysis of one of the targets followed up by myself in the Nomads programme, TOI-332 b. Through a joint fit of *TESS* and LCOGT photometry and HARPS spectroscopy, I determined that it is an unusually massive planet for its size – with a radius smaller than Neptune but a mass half that of Saturn, it ranks as one of the densest Neptunes currently known. A composition analysis determined it has a likely negligible hydrogen-helium atmosphere, which is unexpected given its size. Either it accreted then lost an envelope, or managed to avoid accreting one in the first place. Thus, it required an explanation for its formation, migration and evolution beyond simple core-accretion, and I posited that it could have become what it is today through high-eccentricity migration, giant impacts, or gap-opening in the protoplanetary disc. Future observations of the Rossiter-McLaughlin effect could potentially help constrain these scenarios.

6.2 Future outlook

As stated at the end of Chapter 4, the work needed on Nomads is far from over. The first step is full analyses of each individual system in the sample, minus the spectroscopic double-lined binaries. This is in order to obtain precise masses where possible, or mass upper-limits where not. The null results need to be investigated to answer the question of why they are null results. Once every planet in the sample is fit, we can do an overall analysis and look for trends in e.g., planetary composition. Additionally, as noted several times throughout the Chapter, the current sample is far from perfect. A true statistical study required an unbiased and completeness-corrected search of the *TESS* FFIs, and this would be a very worthwhile avenue of work. Hopefully at that point, the contributions from the

Nomads sample and other teams following up *TESS* discoveries will mean that the number of planets left to follow up will be manageable, but that remains to be seen.

Taking this a step further, I think it would be a worthwhile endeavour to ensure all TOIs within the desert are followed up to some degree, at least to make a determination of whether they are planetary or not. This does not necessarily need to be via precise radial velocity measurements, and could instead be via statistical validation (e.g., Magliano et al., 2023). Looking at Fig. 4.1, there are a startling number of TOIs in the desert, enough to eradicate the boundaries we currently draw around it if they are all planetary. Why are there so many? What is the rate of false positives, and is this a higher false positive rate than for TOIs outside the desert? What false positive scenarios dominate in this regime? I ask these questions because I can see from the Nomads work that a good number of candidates within the desert are likely false positives. Greening the desert is an intriguing prospect, but from a very preliminary view, it still does not look like deep desert is very populated despite the initial excitement from the TOI releases that it may be.

As shown in my study of TOI-332 b, it is all well and good to obtain precise periods, radii, and masses of planets, but there is only so much you can infer from these properties when it comes to their formation, evolution, and migration histories. Further observations using different facilities and techniques can be leveraged to make further inferences. For example, as mentioned for TOI-332 b, a measurement of the Rossiter-McLaughlin effect can determine whether a planet's orbital plane is aligned with its host star's rotation plane. A misalignment could be important evidence towards high-eccentricity migration over disk migration. Taking it one step further, Rossiter-McLaughlin studies for many planets in and around the desert would allow us to perform a population-level study of how migration potentially sculpts the desert (or not). This is the goal of ATREIDES (Ancestry, Traits, and Relations of Exoplanets Inhabiting the Desert Edges and Savannah), a large programme on ESPRESSO that has just begun Rossiter-McLaughlin observations of planets at the desert edges and in the “savannah”, and for which I am in the consortium.

I think a valuable future avenue of study on the desert would be to redraw the original boundaries of Mazeh et al. (2016). In the years since this study, *TESS* launched, making thousands of new discoveries which are, importantly, amenable to follow-up observations due to their magnitudes. This has resulted in many more planetary mass measurements. We can see that many planet discoveries have been made along the edges of the desert, and that the Mazeh boundaries no longer accurately demarcate the density change in planet numbers as you move into the desert (see e.g. Fig. 5.1). It would be interesting to explore whether these empirically-defined boundaries change with, for example, different host star spectral types, as it has been noted that the properties of the star affects the nature of the planets that form around it and thus the desert boundaries (e.g., Szabó et al., 2023). It would also be

valuable work to compare new empirical boundaries to theoretically determined ones, for example from works like Owen and Lai (2018).

Zooming out from Neptunian desert studies, there is a lot to be excited about when it comes to exoplanet discovery prospects in the coming decades. *TESS* is far from finished, with several more years of observations already planned out and likely more to come. There are enough candidates to keep the follow-up sub groups busy for a while yet. Even *K2* still has candidates that are ripe for follow-up, as evidenced by the *K2* discovery that I followed up with Nomads.

The next facility focused on discovering transiting exoplanets is the ESA *PLATO* mission (Rauer et al., 2014; Nascimbeni et al., 2022). Set to launch in 2026, it will monitor two fields, one in each hemisphere, with a multiple year baseline. This is towards the goal of discovering longer period rocky planets around Solar-type stars, i.e. “true” Earth-analogues. HARPS3, a copy of the HARPS/HARPS-N spectrographs, is to be installed on the Isaac Newton Telescope for the Terra Hunting Experiment, a 10 year radial velocity programme also looking to find Earth-like exoplanets around our nearest and brightest Solar-type stars. (Thompson et al., 2016). The future data releases of *Gaia* will enable astrometric discovery of some tens of thousands of planets, sensitive to giant planets on wide orbits (Perryman et al., 2014). The Nancy Grace Roman Space Telescope, set to launch in 2027, will monitor millions of stars towards the centre of the galaxy in a microlensing survey in order to discover a large number of long period low-mass exoplanets (Penny et al., 2019). The Exoplanet *Euclid* Legacy Survey (ExELS) will also use microlensing to measure the abundance of exoplanets down to Earth masses for semi-major axes of 1 au out to the free-floating regime (Penny et al., 2013). The range of different detection techniques and upcoming missions will enable the discovery of exoplanets in regimes that have not yet been studied due to their inaccessibility.

It was only a mere thirty years ago when we had no confirmation that exoplanets even existed. In the few decades since, there have been over 5000 confirmed exoplanet discoveries, showcasing a breathtaking range of planetary sizes and orbital architectures, some of which do not exist in our own Solar System. Looking at population-level features has enabled us to make inferences about the formation, evolution, and migration processes they are subject to throughout their lifetimes. This thesis has presented work to both detect and characterise important planetary systems towards this goal, as well as steps towards a larger population study. This is one of many areas of exoplanet study that will ultimately lead us to a greater understanding of the Universe, and our place within it.

Bibliography

- Addison B. et al., 2019. *PASP*, 131(1005):115003.
- Addison B.C. et al., 2020. *arXiv e-prints*, arXiv:2001.07345.
- Adibekyan V., 2019. *Geosciences*, 9(3):105.
- Adibekyan V. et al., 2015. *A&A*, 583:A94.
- Adibekyan V.Z. et al., 2012. *A&A*, 545:A32.
- Affolter L. et al., 2023. *A&A*, 676:A119.
- Agol E. et al., 2021. *PSJ*, 2(1):1.
- Akana Murphy J.M. et al., 2021. *AJ*, 162(6):294.
- Aller A. et al., 2020. *A&A*, 635:A128.
- Almenara J.M. et al., 2022. *A&A*, 667:L11.
- Alonso R. et al., 2004. *ApJ*, 613(2):L153–L156.
- Armstrong D.J. et al., 2020. *Nature*, 583(7814):39–42.
- Astudillo-Defru N. et al., 2020. *A&A*, 636:A58.
- Baştürk Ö. et al., 2011. *A&A*, 535:A17.
- Baglin A. et al., 2006. In *36th COSPAR Scientific Assembly*, volume 36, 3749.
- Bailer-Jones C.A.L. et al., 2021. *AJ*, 161(3):147.
- Bakos G. et al., 2009. In F. Pont, D. Sasselov, and M.J. Holman, editors, *Transiting Planets*, volume 253, 354–357.
- Bakos G.Á. et al., 2002. *PASP*, 114(799):974–987.

- Baluev R.V., 2008. *MNRAS*, 385(3):1279–1285.
- Baranne A., Mayor M., and Poncet J.L., 1979. *Vistas in Astronomy*, 23(4):279–316.
- Baranne A. et al., 1996. *A&AS*, 119:373–390.
- Barbary K., 2016. *Journal of Open Source Software*, 1(6):58.
- Barnes S.I. et al., 2012. In *Ground-based and Airborne Instrumentation for Astronomy IV*, volume 8446 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 844688.
- Barragán O. et al., 2019. *MNRAS*, 490(1):698–708.
- Batalha N.M. et al., 2011. *ApJ*, 729(1):27.
- Batygin K., 2012. *Nature*, 491(7424):418–420.
- Baxter C. et al., 2020. *A&A*, 639:A36.
- Bean J.L. et al., 2023. *Nature*, 618(7963):43–46.
- Beaugé C. and Nesvorný D., 2013. *ApJ*, 763(1):12.
- Beaugé C. et al., 2007. *A&A*, 463(1):359–367.
- Becker J.C. et al., 2015. *ApJ*, 812(2):L18.
- Berger T.A. et al., 2018. *ApJ*, 866(2):99.
- Berger T.A. et al., 2020. *AJ*, 160(3):108.
- Bertin E. and Arnouts S., 1996. *A&AS*, 117:393–404.
- Bertran de Lis S. et al., 2015. *A&A*, 576:A89.
- Betancourt M., 2017. *arXiv e-prints*, arXiv:1701.02434.
- Beuzit J.L. et al., 2008. In I.S. McLean and M.M. Casali, editors, *Ground-based and Airborne Instrumentation for Astronomy II*, volume 7014 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 701418.
- Bhatti W. et al., 2020. waqasbhatti/astrobase: astrobase v0.5.0.
- Biddle L.I. et al., 2014. *MNRAS*, 443(2):1810–1820.
- Bodenheimer P., Tohline J.E., and Black D.C., 1980. *ApJ*, 242:209–218.

- Bodenheimer P. et al., 2018. *ApJ*, 868(2):138.
- Boisse I. et al., 2011. *A&A*, 528:A4.
- Bolton S.J. et al., 2017. *SSRv*, 213(1-4):5–37.
- Bond I.A. et al., 2001. *MNRAS*, 327(3):868–880.
- Bond I.A. et al., 2004. *ApJ*, 606(2):L155–L158.
- Bond J.C., O’Brien D.P., and Lauretta D.S., 2010. *ApJ*, 715(2):1050–1070.
- Bonomo A.S. et al., 2023. *A&A*, 677:A33.
- Borsato L. et al., 2014. *A&A*, 571:A38.
- Borsato L. et al., 2019. *MNRAS*, 484(3):3233–3243.
- Borucki W.J. et al., 2010. *Science*, 327(5968):977.
- Borucki W.J. et al., 2011a. *ApJ*, 728(2):117.
- Borucki W.J. et al., 2011b. *ApJ*, 736(1):19.
- Boss A.P., 1997. *Science*, 276:1836–1839.
- Bouchy F., Pepe F., and Queloz D., 2001. *A&A*, 374:733–739.
- Bouchy F. and Queloz D., 2003. HARPS DRS User Manual.
- Boué G. et al., 2012. *A&A*, 537:L3.
- Bourrier V. and Hébrard G., 2014. *A&A*, 569:A65.
- Bourrier V. et al., 2021. *A&A*, 654:A152.
- Bourrier V. et al., 2023. *A&A*, 669:A63.
- Boyajian T.S. et al., 2016. *MNRAS*, 457(4):3988–4004.
- Brahm R., Jordán A., and Espinoza N., 2017. *PASP*, 129(3):034002.
- Brahm R. et al., 2019. *AJ*, 158(1):45.
- Brasser R., Barr A.C., and Dobos V., 2019. *MNRAS*, 487(1):34–47.
- Brown T.M. et al., 2013. *PASP*, 125(931):1031.
- Bruno G., 1584. De l’infinito universo et mondi [on the infinite universe and worlds].

- Bruntt H. et al., 2010. *MNRAS*, 405(3):1907–1923.
- Bryant E.M., Bayliss D., and Van Eylen V., 2023. *MNRAS*, 521(3):3663–3681.
- Bryant E.M. et al., 2020. *arXiv e-prints*, arXiv:2004.07589.
- Buchhave L.A. et al., 2016. *AJ*, 152(6):160.
- Buchner J., 2014. *arXiv e-prints*, arXiv:1407.5459.
- Burdanov A. et al., 2018. In H.J. Deeg and J.A. Belmonte, editors, *Handbook of Exoplanets*, 130. Springer Nature.
- Burt J.A. et al., 2020. *AJ*, 160(4):153.
- Butler R.P. et al., 2004. *ApJ*, 600(1):L75–L78.
- Cadieux C. et al., 2022. *AJ*, 164(3):96.
- Caldwell D.A. et al., 2020. *Research Notes of the American Astronomical Society*, 4(11):201.
- Cale B. et al., 2019. *AJ*, 158(5):170.
- Campbell B., Walker G.A.H., and Yang S., 1988. *ApJ*, 331:902.
- Cegla H.M. et al., 2016. *A&A*, 588:A127.
- Cegla H.M. et al., 2019. *ApJ*, 879(1):55.
- Chabrier G., 2001. *ApJ*, 554:1274–1281.
- Chabrier G. et al., 1992. *ApJ*, 391:817.
- Chachan Y. et al., 2022. *ApJ*, 926(1):62.
- Chaplin W.J. and Miglio A., 2013. *ARA&A*, 51(1):353–392.
- Chaplin W.J. et al., 2019. *AJ*, 157(4):163.
- Charbonneau D. et al., 2000. *ApJ*, 529(1):L45–L48.
- Charbonneau D. et al., 2002. *ApJ*, 568(1):377–384.
- Charbonneau D. et al., 2009. *Nature*, 462(7275):891–894.
- Chauvin G. et al., 2004. *A&A*, 425:L29–L32.
- Chen J. and Kipping D., 2017. *ApJ*, 834(1):17.

- Christiansen J.L. et al., 2017. *AJ*, 154(3):122.
- Christiansen J.L. et al., 2022. *AJ*, 163(6):244.
- Ciardi D.R. et al., 2015. *ApJ*, 805(1):16.
- Clausen N. and Tilgner A., 2015. *A&A*, 584:A60.
- Cloutier R. and Menou K., 2020. *AJ*, 159(5):211.
- Cloutier R. et al., 2017. *A&A*, 608:A35.
- Cloutier R. et al., 2019. *A&A*, 629:A111.
- Cloutier R. et al., 2020a. *AJ*, 160(1):3.
- Cloutier R. et al., 2020b. *AJ*, 160(1):22.
- Collier Cameron A. et al., 2007. *MNRAS*, 380(3):1230–1244.
- Collier Cameron A. et al., 2010. *MNRAS*, 407(1):507–514.
- Collins K., 2019. In *American Astronomical Society Meeting Abstracts #233*, volume 233 of *American Astronomical Society Meeting Abstracts*, 140.05.
- Collins K. et al., 2018. In *American Astronomical Society Meeting Abstracts #231*, volume 231 of *American Astronomical Society Meeting Abstracts*, 439.08.
- Collins K.A. et al., 2017. *AJ*, 153(2):77.
- Connolly J.A.D., 2009. *Geochemistry, Geophysics, Geosystems*, 10(10):Q10014.
- Cowan N.B. and Abbot D.S., 2014. *ApJ*, 781(1):27.
- Crida A., Morbidelli A., and Masset F., 2006. *Icarus*, 181(2):587–604.
- Crossfield I. et al., 2018. Spitzer Transits of New TESS Planets. Spitzer Proposal.
- Crouzet N. et al., 2010. *A&A*, 511:A36.
- Csizmadia S., 2018. In H.J. Deeg and J.A. Belmonte, editors, *Handbook of Exoplanets*, 41. Springer Nature.
- Curiel S. et al., 2020. *AJ*, 160(3):97.
- Curiel S. et al., 2022. *AJ*, 164(3):93.
- Currie T. et al., 2023a. *Science*, 380(6641):198–203.

- Currie T. et al., 2023b. In S. Inutsuka, Y. Aikawa, T. Muto, K. Tomida, and M. Tamura, editors, *Protostars and Planets VII*, volume 534 of *Astronomical Society of the Pacific Conference Series*, 799.
- Dai F., Masuda K., and Winn J.N., 2018. *ApJ*, 864(2):L38.
- Dai F. et al., 2017. *AJ*, 154(6):226.
- D’Angelo G., Durisen R.H., and Lissauer J.J., 2010. In S. Seager, editor, *Exoplanets*, 319–346. University of Arizona Press.
- D’Angelo G. and Lissauer J.J., 2018. In H.J. Deeg and J.A. Belmonte, editors, *Handbook of Exoplanets*, 140. Springer Cham.
- D’Angelo G. and Lubow S.H., 2008. *ApJ*, 685(1):560–583.
- D’Angelo G., Lubow S.H., and Bate M.R., 2006. *ApJ*, 652(2):1698–1714.
- Davis A.B. et al., 2020. *AJ*, 160(5):229.
- Dawson R.I. and Fabrycky D.C., 2010. *ApJ*, 722(1):937–953.
- Dawson R.I. and Johnson J.A., 2018. *ARA&A*, 56:175–221.
- Debras F. and Chabrier G., 2019. *ApJ*, 872(1):100.
- Deeg H.J. and Belmonte J.A., 2018. *Handbook of Exoplanets*. Springer Cham.
- Delgado Mena E. et al., 2017. *A&A*, 606:A94.
- Delgado Mena E. et al., 2019. *A&A*, 624:A78.
- Delgado Mena E. et al., 2021. *A&A*, 655:A99.
- Demangeon O.D.S. et al., 2018. *A&A*, 610:A63.
- Deming D. et al., 2015. *ApJ*, 805(2):132.
- Desidera S. et al., 2023. *A&A*, 675:A158.
- Díaz M.R. et al., 2020. *MNRAS*, 493(1):973–985.
- Díaz R.F. et al., 2016. *A&A*, 585:A134.
- Dodson-Robinson S.E. et al., 2009. *ApJ*, 707(1):79–88.
- Dong J. et al., 2021. *ApJS*, 255(1):6.

- Dong J. et al., 2023. *ApJ*, 951(2):L29.
- Dorn C. and Heng K., 2018. *ApJ*, 853(1):64.
- Dorn C. et al., 2015. *A&A*, 577:A83.
- Dorn C. et al., 2017. *A&A*, 597:A37.
- Doyle A.P. et al., 2014. *MNRAS*, 444(4):3592–3602.
- Dragomir D. et al., 2019. *ApJ*, 875(2):L7.
- Dreizler S. et al., 2020. *A&A*, 644:A127.
- Duffell P.C. and MacFadyen A.I., 2013. *ApJ*, 769(1):41.
- Dumusque X. et al., 2019. *A&A*, 627:A43.
- Egeland R. et al., 2017. *ApJ*, 835(1):25.
- Elkins-Tanton L.T. and Seager S., 2008. *ApJ*, 688(1):628–635.
- Emsenhuber A. et al., 2021. *A&A*, 656:A70.
- Encrenaz T., Tinetti G., and Coustenis A., 2018. *Experimental Astronomy*, 46(1):31–44.
- Erkaev N.V. et al., 2007. *A&A*, 472(1):329–334.
- Espinoza N. and Jordán A., 2015. *MNRAS*, 450(2):1879–1899.
- Espinoza N., Kossakowski D., and Brahm R., 2019. *MNRAS*, 490(2):2262–2283.
- Espinoza N. et al., 2020. *MNRAS*, 491(2):2982–2999.
- Esposito M. et al., 2019. *A&A*, 623:A165.
- Essack Z. et al., 2023. *AJ*, 165(2):47.
- Evans T.M. et al., 2016. *ApJ*, 822(1):L4.
- Fabrycky D.C. and Winn J.N., 2009. *ApJ*, 696(2):1230–1240.
- Farihi J., 2016. *NewAR*, 71:9–34.
- Fazio G.G. et al., 2004. *ApJS*, 154:10–17.
- Fernández Fernández J., Wheatley P.J., and King G.W., 2023. *MNRAS*, 522(3):4251–4264.
- Fischer D.A. and Valenti J., 2005. *ApJ*, 622(2):1102–1117.

- Ford E.B. and Gaudi B.S., 2006. *ApJ*, 652(2):L137–L140.
- Ford E.B. and Holman M.J., 2007. *ApJ*, 664(1):L51–L54.
- Foreman-Mackey D., 2016. *The Journal of Open Source Software*, 1(2):24.
- Foreman-Mackey D. et al., 2013. *PASP*, 125(925):306.
- Foreman-Mackey D. et al., 2017. celerite: Scalable 1D Gaussian Processes in C++, Python, and Julia.
- Foreman-Mackey D. et al., 2020. exoplanet-dev/exoplanet v0.2.6.
- Foreman-Mackey D. et al., 2021. *arXiv e-prints*, arXiv:2105.01994.
- Fouesneau M. et al., 2023. *A&A*, 674:A28.
- Frame G. et al., 2023. *MNRAS*, 523(1):1163–1174.
- Fressin F. et al., 2013. *ApJ*, 766(2):81.
- Freudenthal J. et al., 2018. *A&A*, 618:A41.
- Fridlund M. et al., 2020. *MNRAS*, 498(3):4503–4517.
- Frustagli G. et al., 2020. *A&A*, 633:A133.
- Fu R., O’Connell R.J., and Sasselov D.D., 2010. *ApJ*, 708(2):1326–1334.
- Fulton B.J. and Petigura E.A., 2018. *AJ*, 156(6):264.
- Fulton B.J. et al., 2017. *AJ*, 154(3):109.
- Furlan E. et al., 2017. *AJ*, 153(2):71.
- Gaia Collaboration et al., 2016. *A&A*, 595:A2.
- Gaia Collaboration et al., 2018. *A&A*, 616:A1.
- Gaia Collaboration et al., 2021. *A&A*, 649:A1.
- Gaia Collaboration et al., 2023. *A&A*, 674:A1.
- Gan T. et al., 2023. *AJ*, 165(1):17.
- Gandolfi D. et al., 2017. *AJ*, 154(3):123.
- Gandolfi D. et al., 2018. *A&A*, 619:L10.

- Gaudi B.S., 2012. *ARA&A*, 50:411–453.
- Gaudi B.S. and Winn J.N., 2007. *ApJ*, 655(1):550–563.
- Gaudi B.S. et al., 2017. In *American Astronomical Society Meeting Abstracts #230*, volume 230 of *American Astronomical Society Meeting Abstracts*, 102.06.
- Giacalone S. et al., 2022. *AJ*, 163(2):99.
- Gill S. et al., 2020. *ApJ*, 898(1):L11.
- Ginzburg S., Schlichting H.E., and Sari R., 2018. *MNRAS*, 476(1):759–765.
- Girardi L. et al., 2012. *Astrophysics and Space Science Proceedings*, 26:165.
- Goldreich P. and Soter S., 1966. *Icarus*, 5(1):375–389.
- Gomes da Silva J. et al., 2018. *The Journal of Open Source Software*, 3(31):667.
- Gomes da Silva J. et al., 2021. *A&A*, 646:A77.
- Goodman J. and Weare J., 2010. *Communications in Applied Mathematics and Computational Science*, 5(1):65–80.
- Gorti U. and Hollenbach D., 2009. *ApJ*, 690(2):1539–1552.
- Gray D.F., 2005. *The Observation and Analysis of Stellar Photospheres*. Cambridge University Press.
- Grieves N. et al., 2021. *A&A*, 652:A127.
- Grunblatt S.K., Howard A.W., and Haywood R.D., 2015. *ApJ*, 808(2):127.
- Guerrero N.M. et al., 2021. *ApJS*, 254(2):39.
- Guillot T., 2005. *Annual Review of Earth and Planetary Sciences*, 33:493–530.
- Guilluy G. et al., 2023. *A&A*, 676:A130.
- Günther M.N. and Daylan T., 2019. Allesfitter: Flexible Star and Exoplanet Inference From Photometry and Radial Velocity. Astrophysics Source Code Library.
- Günther M.N. and Daylan T., 2021. *ApJS*, 254(1):13.
- Günther M.N. et al., 2019. *Nature Astronomy*, 3:1099–1108.
- Guo X. et al., 2020. *AJ*, 159(5):239.

- Gupta A. and Schlichting H.E., 2019. *MNRAS*, 487(1):24–33.
- Gustafsson B. et al., 2008. *A&A*, 486(3):951–970.
- Hadden S. and Lithwick Y., 2014. *ApJ*, 787(1):80.
- Haghighipour N., 2015. In M. Gargaud, W.M. Irvine, R. Amils, I. Cleaves Henderson James (Jim), D.L. Pinti, J.C. Quintanilla, D. Rouan, T. Spohn, S. Tirard, and M. Viso, editors, *Encyclopedia of Astrobiology*, 748–748. Springer Berlin Heidelberg.
- Hakim K. et al., 2018. *Icarus*, 313:61–78.
- Haldemann J. et al., 2020. *arXiv e-prints*, arXiv:2009.10098.
- Hallatt T. and Lee E.J., 2022. *ApJ*, 924(1):9.
- Halverson S. et al., 2015. *ApJ*, 806(1):61.
- Han C. et al., 2023. *A&A*, 674:A89.
- Hartman J.D. et al., 2011. *ApJ*, 728(2):138.
- Hastings W.K., 1970. *Biometrika*, 57(1):97–109.
- Haswell C.A., 2010. *Transiting Exoplanets*. Cambridge University Press, in association with The Open University.
- Hatzes A.P. et al., 2003. *ApJ*, 599(2):1383–1394.
- Hawthorn F. et al., 2023a. *arXiv e-prints*, arXiv:2310.17268.
- Hawthorn F. et al., 2023b. *MNRAS*, 524(3):3877–3893.
- Haywood R.D. et al., 2014. *MNRAS*, 443(3):2517–2531.
- Helled R. and Bodenheimer P., 2011. *Icarus*, 211(2):939–947.
- Helled R. and Fortney J.J., 2020. *Philosophical Transactions of the Royal Society of London Series A*, 378(2187):20190474.
- Helled R., Lozovsky M., and Zucker S., 2016. *MNRAS*, 455(1):L96–L98.
- Helled R., Podolak M., and Kovetz A., 2008. *Icarus*, 195(2):863–870.
- Helled R. and Schubert G., 2008. *Icarus*, 198(1):156–162.
- Helled R. and Schubert G., 2009. *ApJ*, 697(2):1256–1262.

Helled R. and Stevenson D., 2017a. *ApJ*, 840(1):L4.

Helled R. and Stevenson D., 2017b. *ApJ*, 840(1):L4.

Helled R. et al., 2011. *ApJ*, 726(1):15.

Helled R. et al., 2022. *Icarus*, 378:114937.

Hellier C. et al., 2010. *ApJ*, 723(1):L60–L63.

Hellier C. et al., 2019a. *MNRAS*, 488(3):3067–3075.

Hellier C. et al., 2019b. *Research Notes of the American Astronomical Society*, 3(10):156.

Henning W.G. and Hurford T., 2014. *ApJ*, 789(1):30.

Henry G.W. et al., 2000. *ApJ*, 529(1):L41–L44.

Herbig G.H., 1962. *Advances in Astronomy and Astrophysics*, 1:47–103.

Hillenbrand L.A., 2008. *Physica Scripta Volume T*, 130:014024.

Hippke M. and Angerhausen D., 2015. *ApJ*, 811(1):1.

Hippke M. and Heller R., 2019. TLS: Transit Least Squares.

Hirano T. et al., 2018. *AJ*, 155(3):127.

Ho C.S.K. and Van Eylen V., 2023. *MNRAS*, 519(3):4056–4073.

Hodapp K.W. et al., 2003. *PASP*, 115(814):1388–1406.

Hoffman M.D. and Gelman A., 2011. *arXiv e-prints*, arXiv:1111.4246.

Holman M.J. et al., 2010. *Science*, 330(6000):51.

Howard A.W. et al., 2010. *ApJ*, 721(2):1467–1481.

Howell S.B. et al., 2011. *AJ*, 142(1):19.

Howell S.B. et al., 2014. *PASP*, 126(938):398.

Hsu D.C. et al., 2019. *AJ*, 158(3):109.

Hu R. et al., 2021. *ApJ*, 921(1):L8.

Huang C., Wu Y., and Triaud A.H.M.J., 2016. *ApJ*, 825(2):98.

Huang C.X. et al., 2018. *ApJ*, 868(2):L39.

- Huang C.X. et al., 2020. *Research Notes of the American Astronomical Society*, 4(11):204.
- Ionov D.E., Pavlyuchenkov Y.N., and Shematovich V.I., 2018. *MNRAS*, 476(4):5639–5644.
- Irwin J. et al., 2009. In F. Pont, D. Sasselov, and M.J. Holman, editors, *Transiting Planets*, volume 253, 37–43.
- Isaacson H. and Fischer D., 2010. *ApJ*, 725(1):875–885.
- Ivanov P.B. and Papaloizou J.C.B., 2004. *MNRAS*, 347(2):437–453.
- Izidoro A. and Raymond S.N., 2018. In H.J. Deeg and J.A. Belmonte, editors, *Handbook of Exoplanets*, 142. Springer Cham.
- Jackson A.P., Davis T.A., and Wheatley P.J., 2012. *MNRAS*, 422(3):2024–2043.
- Jackson D.G. et al., 2023. *MNRAS*, 518(4):4845–4860.
- Jackson J.M. et al., 2021. *AJ*, 161(4):200.
- Janson M., 2013. *ApJ*, 774(2):156.
- Jehin E. et al., 2011. *The Messenger*, 145:2–6.
- Jenkins J.M. et al., 2016. In G. Chiozzi and J.C. Guzman, editors, *Software and Cyberinfrastructure for Astronomy IV*, volume 9913 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 99133E.
- Jenkins J.S. et al., 2020. *Nature Astronomy*, 4:1148–1157.
- Jensen E., 2013. Tapir: A web interface for transit/eclipse observability. Astrophysics Source Code Library, record ascl:1306.007.
- Jin S. and Mordasini C., 2018. *ApJ*, 853(2):163.
- Johnson J.A. et al., 2017. *AJ*, 154(3):108.
- Johnstone C.P., Bartel M., and Güdel M., 2021. *A&A*, 649:A96.
- Jontof-Hutter D. et al., 2014. *ApJ*, 785(1):15.
- Jordán A. et al., 2020. *AJ*, 159(4):145.
- Kálmán S. et al., 2023. *MNRAS*, 522(1):488–502.
- Kasting J.F., Whitmire D.P., and Reynolds R.T., 1993. *Icarus*, 101(1):108–128.

- Kaufer A. and Pasquini L., 1998. In *Proc. SPIE*, volume 3355 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 844–854.
- Kaufer A. et al., 1999. *The Messenger*, 95:8–12.
- Kempton E.M.R. et al., 2018. *PASP*, 130(993):114401.
- Kennedy G.M. and Kenyon S.J., 2008. *ApJ*, 673(1):502–512.
- Kenyon S.J. and Hartmann L., 1987. *ApJ*, 323:714.
- Kim S.L. et al., 2016. *Journal of Korean Astronomical Society*, 49(1):37–44.
- King G.W. et al., 2018. *MNRAS*, 478(1):1193–1208.
- Kipping D., 2023. *MNRAS*, 523(1):1182–1191.
- Kipping D.M., 2013. *mnras*, 435:2152–2160.
- Knutson H.A. et al., 2007. *ApJ*, 655(1):564–575.
- Konacki M. et al., 2003. *Nature*, 421(6922):507–509.
- Kopparapu R.K. et al., 2013. *ApJ*, 765(2):131.
- Korth J. et al., 2023. *A&A*, 675:A115.
- Kosiarek M.R. et al., 2021. *AJ*, 161(1):47.
- Kossakowski D. et al., 2021. *A&A*, 656:A124.
- Kostov V.B. et al., 2016. *ApJ*, 827(1):86.
- Kozai Y., 1962. *AJ*, 67:591–598.
- Kral Q., Davoult J., and Charnay B., 2020. *Nature Astronomy*, 4:769–775.
- Kral Q. et al., 2018. *MNRAS*, 479(2):2649–2672.
- Kreidberg L. et al., 2014. *ApJ*, 793(2):L27.
- Kubyshkina D. et al., 2018. *ApJ*, 866(2):L18.
- Kubyshkina D.I. and Fossati L., 2021. *Research Notes of the American Astronomical Society*, 5(4):74.
- Kuchner M.J., 2003. *ApJ*, 596(1):L105–L108.

- Kuchner M.J. and Seager S., 2005. *arXiv e-prints*, astro-ph/0504214.
- Kurokawa H. and Nakamoto T., 2014. *ApJ*, 783(1):54.
- Kurucz R.L., 1993. *SYNTH* spectrum synthesis programs and line data. Smithsonian Astrophysical Observatory.
- Kurucz R.L., 2013. ATLAS12: Opacity sampling model atmosphere program. Astrophysics Source Code Library, record ascl:1303.024.
- Kuzuhara M. et al., 2013. *ApJ*, 774(1):11.
- La Silla Observatory, 2019. HARPS User Manual.
- Lafarga M. et al., 2020. *A&A*, 636:A36.
- Lagrange A.M., Desort M., and Meunier N., 2010. *A&A*, 512:A38.
- Lai D., Foucart F., and Lin D.N.C., 2011. *MNRAS*, 412(4):2790–2798.
- Lam K.W.F. et al., 2021. *Science*, 374(6572):1271–1275.
- Lambrechts M. and Johansen A., 2012. *A&A*, 544:A32.
- Lambrechts M., Johansen A., and Morbidelli A., 2014. *A&A*, 572:A35.
- Laughlin G. and Chambers J.E., 2002. *AJ*, 124(1):592–600.
- Lecavelier des Etangs A. and Lissauer J.J., 2022. *NewAR*, 94:101641.
- Lee E.J., 2019. *ApJ*, 878(1):36.
- Lee E.J., Karalis A., and Thorngren D.P., 2022. *ApJ*, 941(2):186.
- Léger A. et al., 2004. *Icarus*, 169(2):499–504.
- Léger A. et al., 2009. *A&A*, 506(1):287–302.
- Léger A. et al., 2011. *Icarus*, 213(1):1–11.
- Lehmer O.R. and Catling D.C., 2017. *ApJ*, 845(2):130.
- Leleu A., Coleman G.A.L., and Ataiee S., 2019. *A&A*, 631:A6.
- Leleu A. et al., 2017. *A&A*, 599:L7.
- Lendl M. et al., 2020. *MNRAS*, 492(2):1761–1769.

- Li J. et al., 2019. *PASP*, 131(996):024506.
- Libby-Roberts J.E. et al., 2020. *AJ*, 159(2):57.
- Libby-Roberts J.E. et al., 2023. *AJ*, 165(6):249.
- Lidov M.L., 1962. *P&SS*, 9(10):719–759.
- Lillo-Box J., Barrado D., and Bouy H., 2014a. *A&A*, 566:A103.
- Lillo-Box J. et al., 2014b. *A&A*, 562:A109.
- Lillo-Box J. et al., 2018a. *A&A*, 618:A42.
- Lillo-Box J. et al., 2018b. *A&A*, 609:A96.
- Lillo-Box J. et al., 2020. *A&A*, 642:A121.
- Lillo-Box J. et al., 2023. *A&A*, 669:A109.
- Lin D.N.C. and Papaloizou J., 1986. *ApJ*, 309:846.
- Lissauer J.J., 1987. *Icarus*, 69(2):249–265.
- Lissauer J.J. and Stevenson D.J., 2007. In B. Reipurth, D. Jewitt, and K. Keil, editors, *Protostars and Planets V*, 591.
- Lissauer J.J. et al., 2009. *Icarus*, 199(2):338–350.
- Lithwick Y., Xie J., and Wu Y., 2012. *ApJ*, 761(2):122.
- Liu S.F. et al., 2015. *MNRAS*, 446(2):1685–1702.
- Livingston J.H. et al., 2018. *AJ*, 156(6):277.
- Lomb N.R., 1976. *Ap&SS*, 39(2):447–462.
- Lopez E.D. and Fortney J.J., 2014. *ApJ*, 792(1):1.
- Lopez E.D. and Rice K., 2018. *MNRAS*, 479(4):5303–5311.
- Lothringer J.D. et al., 2022. *Nature*, 604(7904):49–52.
- Lovis C. et al., 2011. *arXiv e-prints*, arXiv:1107.5325.
- Lubow S.H. and D’Angelo G., 2006. *ApJ*, 641(1):526–533.
- Luger R. et al., 2019. *aj*, 157:64.

- Lundkvist M.S. et al., 2016. *Nature Communications*, 7:11201.
- Luque R. et al., 2019. *A&A*, 628:A39.
- Luque R. et al., 2022. *A&A*, 664:A199.
- Maciejewski G. et al., 2016. *A&A*, 588:L6.
- Macintosh B.A. et al., 2008. In N. Hubin, C.E. Max, and P.L. Wizinowich, editors, *Adaptive Optics Systems*, volume 7015 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 701518.
- Madhusudhan N., Lee K.K.M., and Mousis O., 2012. *ApJ*, 759(2):L40.
- Madhusudhan N., Piette A.A.A., and Constantinou S., 2021. *ApJ*, 918(1):1.
- Madhusudhan N. and Winn J.N., 2009. *ApJ*, 693(1):784–793.
- Madhusudhan N. et al., 2023. Carbon-bearing molecules in a possible hycean atmosphere.
- Magliano C. et al., 2023. *MNRAS*, 519(1):1562–1577.
- Mamajek E.E., 2009. In T. Usuda, M. Tamura, and M. Ishii, editors, *Exoplanets and Disks: Their Formation and Diversity*, volume 1158 of *American Institute of Physics Conference Series*, 3–10.
- Mamajek E.E. and Hillenbrand L.A., 2008. *ApJ*, 687(2):1264–1293.
- Marois C. et al., 2008. *Science*, 322(5906):1348.
- Marois C. et al., 2010. *Nature*, 468(7327):1080–1083.
- Martinez C.F. et al., 2019. *ApJ*, 875(1):29.
- Masuda K., 2014. *ApJ*, 783(1):53.
- Matsakos T. and Königl A., 2016. *ApJ*, 820(1):L8.
- Maxted P.F.L. et al., 2011. *PASP*, 123(903):547.
- Mayor M. and Queloz D., 1995. *Nature*, 378(6555):355–359.
- Mayor M. et al., 2009. *A&A*, 507(1):487–494.
- Mayor M. et al., 2011. *arXiv e-prints*, arXiv:1109.2497.
- Mazeh T., Holczer T., and Faigler S., 2016. *A&A*, 589:A75.

- McCormac J. et al., 2013. *PASP*, 125(927):548.
- McCullough P.R. et al., 2005. *PASP*, 117(834):783–795.
- McCully C. et al., 2018. In J.C. Guzman and J. Ibsen, editors, *Software and Cyberinfrastructure for Astronomy V*, volume 10707 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 107070K.
- McLaughlin D.B., 1924. *ApJ*, 60:22–31.
- Metropolis N. et al., 1953. *JChPh*, 21(6):1087–1092.
- Meunier N., Desort M., and Lagrange A.M., 2010. *A&A*, 512:A39.
- Miguel Y., Guillot T., and Fayon L., 2016. *A&A*, 596:A114.
- Mireles I. et al., 2023. *ApJ*, 954(1):L15.
- Mistry P. et al., 2023. *MNRAS*, 521(1):1066–1078.
- Morbidelli A. and Raymond S.N., 2016. *Journal of Geophysical Research (Planets)*, 121(10):1962–1980.
- Morbidelli A. et al., 2012. *Annual Review of Earth and Planetary Sciences*, 40(1):251–275.
- Mortier A. et al., 2014. *A&A*, 572:A95.
- Morton T.D. et al., 2016. *ApJ*, 822(2):86.
- Murgas F. et al., 2022. *A&A*, 668:A158.
- Namouni F. and Morais H., 2017. *arXiv e-prints*, arXiv:1704.00550.
- Naoz S. et al., 2011. *Nature*, 473(7346):187–189.
- Naponiello L. et al., 2023. *arXiv e-prints*, arXiv:2309.01464.
- Nascimbeni V. et al., 2022. *A&A*, 658:A31.
- Nesvorný D., 2009. *ApJ*, 701(2):1116–1122.
- Nesvorný D. and Vokrouhlický D., 2016. *ApJ*, 823(2):72.
- Nettelmann N. et al., 2012. *ApJ*, 750(1):52.
- Newton E.R. et al., 2021. *AJ*, 161(2):65.

- Newton I., 1726. The General Scholium to Isaac Newton’s *Principia Mathematica*, translated by Andrew Motte.
- Nielsen L.D. et al., 2019. *MNRAS*, 489(2):2478–2487.
- Nielsen L.D. et al., 2020a. *MNRAS*, 492(4):5399–5412.
- Nielsen L.D. et al., 2020b. *A&A*, 639:A76.
- Noyes R.W., Weiss N.O., and Vaughan A.H., 1984. *ApJ*, 287:769–773.
- Ogihara M. et al., 2021. *A&A*, 648:L1.
- Ogilvie G.I., 2014. *ARA&A*, 52:171–210.
- Osborn A. et al., 2021a. *MNRAS*, 507(2):2782–2803.
- Osborn A. et al., 2023. *MNRAS*.
- Osborn H.P., 2022. MonoTools: Planets of uncertain periods detector and modeler. *Astrophysics Source Code Library*, record ascl:2204.020.
- Osborn H.P. et al., 2021b. *MNRAS*, 502(4):4842–4857.
- Osborn H.P. et al., 2022. *A&A*, 664:A156.
- Otegi J.F., Bouchy F., and Helled R., 2020a. *A&A*, 634:A43.
- Otegi J.F. et al., 2020b. *arXiv e-prints*, arXiv:2006.12353.
- Owen J.E. and Lai D., 2018. *MNRAS*, 479(4):5012–5021.
- Owen J.E. and Murray-Clay R., 2018. *MNRAS*, 480(2):2206–2216.
- Owen J.E. and Schlichting H.E., 2023. *arXiv e-prints*, arXiv:2308.00020.
- Owen J.E. and Wu Y., 2013. *ApJ*, 775(2):105.
- Owen J.E. and Wu Y., 2017. *ApJ*, 847(1):29.
- Palle E. et al., 2019. *A&A*, 623:A41.
- Papaloizou J.C.B. and Lin D.N.C., 1995. *ARA&A*, 33:505–540.
- Pascucci I. et al., 2013. *ApJ*, 779(2):178.
- Patra K.C. et al., 2017. *AJ*, 154(1):4.

- Patra K.C. et al., 2020. *AJ*, 159(4):150.
- Pearson K., 1895. *Proceedings of the Royal Society of London Series I*, 58:240–242.
- Pecaut M.J. and Mamajek E.E., 2013. *ApJS*, 208(1):9.
- Pelletier S. et al., 2023. *Nature*, 619(7970):491–494.
- Penny M.T. et al., 2013. *MNRAS*, 434(1):2–22.
- Penny M.T. et al., 2019. *ApJS*, 241(1):3.
- Pepe F. et al., 2002. *The Messenger*, 110:9–14.
- Pepe F. et al., 2014. *Astronomische Nachrichten*, 335(1):8.
- Pepe F. et al., 2018. In H.J. Deeg and J.A. Belmonte, editors, *Handbook of Exoplanets*, 190. Springer Nature.
- Pepe F.A. and Lovis C., 2008. *Physica Scripta Volume T*, 130:014007.
- Pepper J. et al., 2007. *PASP*, 119(858):923–935.
- Perruchot S. et al., 2008. In I.S. McLean and M.M. Casali, editors, *Ground-based and Airborne Instrumentation for Astronomy II*, volume 7014 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 70140J.
- Perryman M., 2018. *The Exoplanet Handbook*. Cambridge University Press.
- Perryman M. et al., 2014. *ApJ*, 797(1):14.
- Persson C.M. et al., 2018. *A&A*, 618:A33.
- Persson C.M. et al., 2022. *A&A*, 666:A184.
- Petigura E.A., Howard A.W., and Marcy G.W., 2013. *Proceedings of the National Academy of Science*, 110(48):19273–19278.
- Petigura E.A. et al., 2017. *AJ*, 154(3):107.
- Petigura E.A. et al., 2018. *AJ*, 155(2):89.
- Piaulet C. et al., 2023. *Nature Astronomy*, 7:206–222.
- Piskunov N. and Valenti J.A., 2017. *A&A*, 597:A16.
- Piskunov N.E. et al., 1995. *A&AS*, 112:525.

- Piso A.M.A., Youdin A.N., and Murray-Clay R.A., 2015. *ApJ*, 800(2):82.
- Podolak M. and Helled R., 2012. *ApJ*, 759(2):L32.
- Podolak M., Weizman A., and Marley M., 1995. *P&SS*, 43(12):1517–1522.
- Poleski R. et al., 2014. *ApJ*, 795(1):42.
- Poleski R. et al., 2021. *AcA*, 71(1):1–23.
- Pollacco D.L. et al., 2006. *PASP*, 118(848):1407–1418.
- Pollack J.B. et al., 1996. *Icarus*, 124(1):62–85.
- Psaridi A. et al., 2023. *A&A*, 675:A39.
- Queloz D. et al., 2000. *A&A*, 354:99–102.
- Queloz D. et al., 2001. *A&A*, 379:279–287.
- Rajpaul V., Aigrain S., and Roberts S., 2016. *MNRAS*, 456(1):L6–L10.
- Rasmussen C.E. and Williams C.K.I., 2006. *Gaussian Processes for Machine Learning*. MIT Press.
- Rauer H. et al., 2014. *Experimental Astronomy*, 38(1-2):249–330.
- Rayner J. et al., 2016. In *Proc. SPIE*, volume 9908 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 990884.
- Reefe M.A. et al., 2022. *AJ*, 163(6):269.
- Ricker G.R. et al., 2015. *Journal of Astronomical Telescopes, Instruments, and Systems*, 1:014003.
- Robertson P. et al., 2013. *ApJ*, 764(1):3.
- Rodrigues T.S. et al., 2017. *MNRAS*, 467(2):1433–1448.
- Rogers J.G. et al., 2021. *MNRAS*, 508(4):5886–5902.
- Rogers L.A., 2015. *ApJ*, 801(1):41.
- Rogers L.A. and Seager S., 2010. *ApJ*, 712(2):974–991.
- Rossiter R.A., 1924. *ApJ*, 60:15–21.
- Rouan D. et al., 2011. *ApJ*, 741(2):L30.

- Rowe J.F. et al., 2014. *ApJ*, 784(1):45.
- Ryabchikova T. and Pakhomov Y., 2015. *Baltic Astronomy*, 24:453–461.
- Safronov V.S., 1972. In *Evolution of the protoplanetary cloud and formation of the earth and planets*.
- Sahlmann J. et al., 2013. *A&A*, 556:A133.
- Salvatier J., Wiecki T.V., and Fonnesbeck C., 2016. *PeerJ Computer Science*, 2:e55.
- Sanchis-Ojeda R. et al., 2013. *ApJ*, 774(1):54.
- Sanchis-Ojeda R. et al., 2014. *ApJ*, 787(1):47.
- Santerne A. et al., 2016. *A&A*, 587:A64.
- Santos N.C. et al., 2013. *A&A*, 556:A150.
- Santos N.C. et al., 2015. *A&A*, 580:L13.
- Santos N.C. et al., 2017. *A&A*, 608:A94.
- Saumon D., Chabrier G., and van Horn H.M., 1995. *ApJS*, 99:713.
- Scargle J.D., 1982. *ApJ*, 263:835–853.
- Schlegel D.J., Finkbeiner D.P., and Davis M., 1998. *ApJ*, 500(2):525–553.
- Schlichenmaier R. et al., 2010. *A&A*, 512:L1.
- Scott N.J. et al., 2021. *Frontiers in Astronomy and Space Sciences*, 8:138.
- Seager S. et al., 2007. *ApJ*, 669(2):1279–1297.
- Serrano L.M. et al., 2022. *Nature Astronomy*, 6:736–750.
- Sha L. et al., 2023. *MNRAS*, 524(1):1113–1138.
- Shporer A. et al., 2020. *ApJ*, 890(1):L7.
- Shu F.H., Adams F.C., and Lizano S., 1987. *ARA&A*, 25:23–81.
- Shu F.H., Johnstone D., and Hollenbach D., 1993. *Icarus*, 106(1):92–101.
- Silverstein M.L. et al., 2022. *AJ*, 163(4):151.
- Skrutskie M.F. et al., 2006. *AJ*, 131(2):1163–1183.

- Smith A.M.S. et al., 2020. *arXiv e-prints*, arXiv:2002.05591.
- Smith B.A. and Terrile R.J., 1984. *Science*, 226(4681):1421–1424.
- Smith J.C. et al., 2012. *PASP*, 124(919):1000.
- Snedden C.A., 1973. *Carbon and Nitrogen Abundances in Metal-Poor Stars*. Ph.D. thesis, University of Texas, Austin.
- Sotin C., Grasset O., and Mocquet A., 2007. *Icarus*, 191(1):337–351.
- Sousa S.G., 2014. In *Determination of Atmospheric Parameters of B*, 297–310. Springer.
- Sousa S.G. et al., 2007. *A&A*, 469(2):783–791.
- Sousa S.G. et al., 2008. *A&A*, 487(1):373–381.
- Sousa S.G. et al., 2015. *A&A*, 577:A67.
- Sousa S.G. et al., 2021. *A&A*, 656:A53.
- Sowmya K. et al., 2023. *ApJ*, 956(1):L10.
- Stassun K.G., Collins K.A., and Gaudi B.S., 2017. *AJ*, 153(3):136.
- Stassun K.G. and Torres G., 2016a. *AJ*, 152(6):180.
- Stassun K.G. and Torres G., 2016b. *AJ*, 152(6):180.
- Stassun K.G. and Torres G., 2021. *ApJ*, 907(2):L33.
- Stassun K.G. et al., 2018. *AJ*, 155(1):22.
- Stassun K.G. et al., 2019. *AJ*, 158(4):138.
- Stefánsson G. et al., 2022. *ApJ*, 931(2):L15.
- Stevens D.J. and Gaudi B.S., 2013. *PASP*, 125(930):933.
- Stevenson D.J., 2020. *Annual Review of Earth and Planetary Sciences*, 48:465–489.
- Strugarek A. et al., 2017. *ApJ*, 847(2):L16.
- Struve O., 1952. *The Observatory*, 72:199–200.
- Stumpe M.C. et al., 2012. *PASP*, 124(919):985.
- Stumpe M.C. et al., 2014. *PASP*, 126(935):100.

- Suárez Mascareño A. et al., 2020. *A&A*, 639:A77.
- Swain M.R. et al., 2009. *ApJ*, 704(2):1616–1621.
- Szabó G.M. and Kálmán S., 2019. *MNRAS*, 485(1):L116–L120.
- Szabó G.M. and Kiss L.L., 2011. *ApJ*, 727(2):L44.
- Szabó G.M. et al., 2023. *A&A*, 671:A132.
- Teske J.K. et al., 2013. *ApJ*, 778(2):132.
- Thompson M.A. et al., 2021. *Nature Astronomy*, 5:575–585.
- Thompson S.J. et al., 2016. In C.J. Evans, L. Simard, and H. Takami, editors, *Ground-based and Airborne Instrumentation for Astronomy VI*, volume 9908 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 99086F.
- Tokovinin A., 2018. *PASP*, 130(985):035002.
- Torres G., Andersen J., and Giménez A., 2010. *A&ARv*, 18(1-2):67–126.
- Torres G. et al., 2011. *ApJ*, 727(1):24.
- Triaud A.H.M.J., 2018. In H.J. Deeg and J.A. Belmonte, editors, *Handbook of Exoplanets*, 2. Springer.
- Twicken J.D. et al., 2010. In N.M. Radziwill and A. Bridger, editors, *Software and Cyber-infrastructure for Astronomy*, volume 7740 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 77401U.
- Twicken J.D. et al., 2018. *PASP*, 130(988):064502.
- Udalski A., 2003. *AcA*, 53:291–305.
- Udry S. et al., 2007. *A&A*, 469(3):L43–L47.
- Ueta S. and Sasaki T., 2013. *ApJ*, 775(2):96.
- Valenti J.A. and Piskunov N., 1996. *A&AS*, 118:595–603.
- Van Eylen V. et al., 2016. *ApJ*, 820(1):56.
- Van Eylen V. et al., 2018. *MNRAS*, 479:4786–4795.
- van Maanen A., 1917. *PASP*, 29(172):258.

- Vanderburg A. et al., 2016. *ApJ*, 829(1):L9.
- VanderPlas J.T., 2018. *ApJS*, 236(1):16.
- Vazan A. et al., 2013. *MNRAS*, 434(4):3283–3292.
- Vazan A. et al., 2015. *ApJ*, 803(1):32.
- Vehtari A. et al., 2021. *Bayesian Analysis*, 16(2):667–718.
- Venturini J. et al., 2020. *A&A*, 643:L1.
- Vick M. and Lai D., 2018. *MNRAS*, 476(1):482–495.
- Vick M., Lai D., and Anderson K.R., 2019. *MNRAS*, 484(4):5645–5668.
- Vines J.I. et al., 2023. *MNRAS*, 518(2):2627–2639.
- Vissapragada S. et al., 2022. *AJ*, 164(6):234.
- Vogt S.S. et al., 1994. In D.L. Crawford and E.R. Craine, editors, *Instrumentation in Astronomy VIII*, volume 2198 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 362.
- Walker G.A.H. et al., 1992. *ApJ*, 396:L91.
- Ward W.R., 1986. *Icarus*, 67(1):164–180.
- Watson A.J., Donahue T.M., and Walker J.C.G., 1981. *Icarus*, 48(2):150–166.
- Weidenschilling S.J., 1980. *Icarus*, 44(1):172–189.
- Weiss L.M. et al., 2021. *AJ*, 161(2):56.
- West R.G. et al., 2019. *MNRAS*, 486(4):5094–5103.
- Wheatley P.J. et al., 2018. *MNRAS*, 475(4):4476–4493.
- Wigner E. and Huntington H.B., 1935. *JChPh*, 3(12):764–770.
- Williams J.P. and Cieza L.A., 2011. *ARA&A*, 49(1):67–117.
- Wilson H.F. and Militzer B., 2014. *ApJ*, 793(1):34.
- Winn J.N., 2010. *arXiv e-prints*, arXiv:1001.2010.
- Winn J.N. and Fabrycky D.C., 2015. *ARA&A*, 53:409–447.

- Winn J.N., Sanchis-Ojeda R., and Rappaport S., 2018. *NewAR*, 83:37–48.
- Winn J.N. et al., 2009. *ApJ*, 703(2):L99–L103.
- Winn J.N. et al., 2010. *ApJ*, 723(2):L223–L227.
- Wolszczan A., 1994. *Science*, 264(5158):538–542.
- Wolszczan A. and Frail D.A., 1992. *Nature*, 355(6356):145–147.
- Wong I. et al., 2022. *AJ*, 163(4):175.
- Wright J.T. and Gaudi B.S., 2013. In T.D. Oswalt, L.M. French, and P. Kalas, editors, *Planets, Stars and Stellar Systems. Volume 3: Solar and Stellar Planetary Systems*, 489. Springer Dordrecht.
- Wright N.J. et al., 2011. *ApJ*, 743(1):48.
- Wright N.J. et al., 2018. *MNRAS*, 479(2):2351–2360.
- Wu D.H. et al., 2018. *AJ*, 156(3):96.
- Wu Y., 2018. *AJ*, 155(3):118.
- Wu Y. and Murray N., 2003. *ApJ*, 589(1):605–614.
- Yee S.W., Petigura E.A., and von Braun K., 2017. *ApJ*, 836(1):77.
- Yee S.W. et al., 2020. *ApJ*, 888(1):L5.
- Yoffe G., Ofir A., and Aharonson O., 2021. *ApJ*, 908(1):114.
- Zeng L. et al., 2019. *Proceedings of the National Academy of Science*, 116(20):9723–9728.
- Zhou G. et al., 2019. *AJ*, 158(4):141.
- Zhou G. et al., 2021. *AJ*, 161(1):2.
- Ziegler C. et al., 2020. *AJ*, 159(1):19.