Simulation, detection and characterisation of monotransits from the TESS mission

by

Benjamin F. Cooke

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

April 2021
Contents

List of Tables vi
List of Figures vii
Acknowledgments x
Declarations xi
Abstract xii
Abbreviations xiii

Chapter 1 Introduction 1
  1.1 Exoplanet history ............................................. 1
  1.2 Discovery methods ........................................... 2
    1.2.1 Transit method ...................................... 2
    1.2.2 Radial velocity ..................................... 8
    1.2.3 Other methods ....................................... 10
  1.3 Exoplanet knowledge ........................................ 12
    1.3.1 Formation and evolution ................................. 12
    1.3.2 Occurrence rates .................................... 16
  1.4 Exoplanet characterisation ................................ 17
    1.4.1 Physical parameters ................................ 17
    1.4.2 Types of exoplanet .................................. 22
  1.5 Thesis outline .............................................. 25

Chapter 2 Methods 26
  2.1 Transit detection methods ................................ 26
    2.1.1 Multitransit detection ............................... 26
    2.1.2 Monotransit detection ............................... 28
Chapter 3  WASP-150b & WASP-176b

3.1 Introduction ................................................. 50
3.2 SuperWASP discovery photometry .......................... 50
3.3 Spectroscopic follow-up ................................. 53
   3.3.1 WASP-150b ........................................ 53
   3.3.2 WASP-176b ........................................ 55
3.4 Photometric follow-up ..................................... 56
   3.4.1 WASP-150b ........................................ 56
   3.4.2 WASP-176b ........................................ 60
3.5 High-spatial-resolution follow-up ......................... 63
3.6 Results ....................................................... 65
   3.6.1 WASP-150b ........................................ 65
   3.6.2 WASP-176b ........................................ 69
3.7 Discussion and conclusions ............................... 71
   3.7.1 WASP-150b ........................................ 71
   3.7.2 WASP-176b ........................................ 74

Chapter 4  TESS primary mission monotransit simulation

4.1 Introduction ................................................ 76
4.2 Stellar population ........................................ 78
4.3 Planet population ........................................ 81
4.4 Detectability .............................................. 82
4.5 Single site follow-up .................................... 85
4.6 Analysis and results ..................................... 87
   4.6.1 TESS single transit demographics .................. 87
   4.6.2 Single site observations ............................ 93
   4.6.3 Realistic follow-up ................................ 96
   4.6.4 False positives ..................................... 98
4.7 Discussion and conclusions ................................ 98
Chapter 8  Period aliases

8.1 Introduction ........................................... 150
8.2 Observations .......................................... 151
  8.2.1 Simulation population ......................... 151
  8.2.2 TESS observations ......................... 152
  8.2.3 Simulated photometry ..................... 153
  8.2.4 Simulated spectroscopy .................... 153
8.3 Alias determination ................................. 154
  8.3.1 Period aliases from TESS ................. 154
  8.3.2 Period aliases using additional data .... 155
8.4 Results and Discussion ............................ 156
  8.4.1 TESS extended mission monotransits ...... 156
  8.4.2 Number of aliases, \( N_{\text{aliases}} \) .......... 156
  8.4.3 True period index, \( I \) ....................... 161
  8.4.4 Improvements in number of aliases, \( \Delta N_{\text{aliases}} \) 162
  8.4.5 Time taken vs nights observed ............. 164
  8.4.6 Effect of increasing follow-up time ........ 166
8.5 Conclusions ........................................ 168
8.A Number of aliases ................................. 169

Chapter 9  TESS monotransit search and NGTS follow-up

9.1 NGTS monotransit working group .................. 171
9.2 Difference Imaging Analysis light curves ........ 172
  9.2.1 DIA light curve generation ............... 172
  9.2.2 Monotransit identification ............... 174
  9.2.3 Candidate eyeballing ....................... 174
9.3 NGTS light curves ................................ 176
  9.3.1 NGTS light curve generation ............... 176
  9.3.2 Monofind ..................................... 176
List of Tables

3.1 WASP-150 & WASP-176 photometric properties .......................... 53
3.2 Photometric follow-up of WASP-150 ........................................ 57
3.3 Photometric follow-up of WASP-176 ................................. 60
3.4 WASP-150 & WASP-176 stellar analysis parameters ............... 66
3.5 WASP-150 & WASP-176 Gaia DR2 parameters ....................... 67
3.6 WASP-150 & WASP-176 bagemass parameters ..................... 67
3.7 WASP-150b & WASP-176b MCMC parameters ....................... 72

4.1 Spectral type classification as a function of $T_{\text{eff}}$ ............... 80
4.2 TESS simulated and detectable planets ............................... 85
4.3 Detectable TESS planets as a function of period .................. 91
4.4 Detectable TESS planets as a function of period - 2min targets only .... 94

5.1 TESS simulated and detectable planets - updated .................. 113
5.2 Detectable TESS planets as a function of period - 2min targets only - updated ... 114

7.1 Parameter space amenable to follow-up as a function of host magnitude .... 143

8.1 Period aliases as a function of follow-up campaign ................ 170
## List of Figures

1.1 Exoplanet discoveries .............................................. 3  
1.2 Transit geometry ................................................. 4  
1.3 Ideal transit light curve ........................................... 5  
1.4 Limb darkening comparison ....................................... 7  
1.5 Phase folded RV curves .......................................... 9  
1.6 PSD70b detection ................................................. 14  
1.7 Planetary occurrence rates from RV surveys ................. 17  
1.8 Planetary occurrence rates from transit surveys ........... 18  
1.9 Atmospheric processes ......................................... 21  
1.10 Planetary regimes ................................................. 22  

2.1 Light curve and BLS periodogram ............................... 27  
2.2 TLS Vs BLS template ............................................. 28  
2.3 Photometric reduction steps .................................... 32  
2.4 Differential photometry steps .................................. 34  
2.5 TESS sky coverage .............................................. 39  
2.6 Basic spectrograph schematic .................................. 41  
2.7 Spectrograph drift removal .................................... 44  
2.8 Solar spectrum .................................................. 45  

3.1 WASP-150b WASP photometry .................................. 51  
3.2 WASP-176b WASP photometry .................................. 52  
3.3 WASP-150 RV data ............................................. 54  
3.4 WASP-150 RV bisector data ................................... 55  
3.5 WASP-176 RV data ............................................. 56  
3.6 WASP-176 RV bisector data ................................... 57  
3.7 WASP-150b follow-up photometry ............................. 61  
3.8 WASP-150b combined photometry ............................. 62  
3.9 WASP-176b follow-up photometry ............................. 64
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.10</td>
<td>WASP-150b combined photometry</td>
<td>65</td>
</tr>
<tr>
<td>3.11</td>
<td>WASP-150 contrast light curve</td>
<td>66</td>
</tr>
<tr>
<td>3.12</td>
<td>WASP-150 BAGEMASS fit</td>
<td>68</td>
</tr>
<tr>
<td>3.13</td>
<td>WASP-176 BAGEMASS fit</td>
<td>70</td>
</tr>
<tr>
<td>3.14</td>
<td>Exoplanet density plot</td>
<td>73</td>
</tr>
<tr>
<td>4.1</td>
<td>North ecliptic hemisphere TIC CTL stars</td>
<td>79</td>
</tr>
<tr>
<td>4.2</td>
<td>TESS sectors sky coverage</td>
<td>79</td>
</tr>
<tr>
<td>4.3</td>
<td>TIC CTL stellar population</td>
<td>80</td>
</tr>
<tr>
<td>4.4</td>
<td>Simulated planet input distributions</td>
<td>82</td>
</tr>
<tr>
<td>4.5</td>
<td>TESS noise budget in 1hr</td>
<td>84</td>
</tr>
<tr>
<td>4.6</td>
<td>Dilution of simulated TESS planet hosts</td>
<td>86</td>
</tr>
<tr>
<td>4.7</td>
<td>TESS simulated planets sky distribution</td>
<td>86</td>
</tr>
<tr>
<td>4.8</td>
<td>Simulated TESS planet demographics</td>
<td>88</td>
</tr>
<tr>
<td>4.9</td>
<td>Simulated TESS planets period distribution</td>
<td>90</td>
</tr>
<tr>
<td>4.10</td>
<td>Number of in transit observations for simulated TESS detections</td>
<td>92</td>
</tr>
<tr>
<td>4.11</td>
<td>TESS simulated single transits visible from Paranal over 1yr</td>
<td>95</td>
</tr>
<tr>
<td>4.12</td>
<td>Distribution of simulated TESS detections with at least 5% transit depth</td>
<td>95</td>
</tr>
<tr>
<td>4.13</td>
<td>Visibility of TESS single transit targets from Paranal</td>
<td>97</td>
</tr>
<tr>
<td>5.1</td>
<td>Period distribution of TESS year 1 monotransits</td>
<td>107</td>
</tr>
<tr>
<td>5.2</td>
<td>Period distribution of TESS monotransits after the extended mission</td>
<td>108</td>
</tr>
<tr>
<td>5.3</td>
<td>Radius distribution of TESS year 1 monotransits</td>
<td>109</td>
</tr>
<tr>
<td>5.4</td>
<td>Magnitude distribution of TESS year 1 monotransits</td>
<td>109</td>
</tr>
<tr>
<td>5.5</td>
<td>Effective temperature distribution of TESS year 1 monotransits</td>
<td>110</td>
</tr>
<tr>
<td>5.6</td>
<td>Distribution of true period index for repeated monotransits</td>
<td>111</td>
</tr>
<tr>
<td>6.1</td>
<td>CHEOPS sky coverage</td>
<td>119</td>
</tr>
<tr>
<td>6.2</td>
<td>TICv8 stars sky distribution</td>
<td>120</td>
</tr>
<tr>
<td>6.3</td>
<td>TESS monotransit hosts overlaid on CHEOPS sky coverage</td>
<td>121</td>
</tr>
<tr>
<td>6.4</td>
<td>CHEOPS coverage of observable TESS monotransit targets</td>
<td>122</td>
</tr>
<tr>
<td>6.5</td>
<td>Period distribution of TESS primary mission monotransits</td>
<td>123</td>
</tr>
<tr>
<td>6.6</td>
<td>Radius distribution of TESS primary mission monotransits</td>
<td>124</td>
</tr>
<tr>
<td>6.7</td>
<td>Probability of CHEOPS observing additional transits</td>
<td>125</td>
</tr>
<tr>
<td>6.8</td>
<td>Cumulative solved systems as a function of follow-up time</td>
<td>127</td>
</tr>
<tr>
<td>7.1</td>
<td>Stellar and planetary mass-radius relations</td>
<td>134</td>
</tr>
<tr>
<td>7.2</td>
<td>Instrumental precision as a function of target magnitude</td>
<td>135</td>
</tr>
<tr>
<td>7.3</td>
<td>SNR ratio in NGTS/HARPS parameter space</td>
<td>139</td>
</tr>
</tbody>
</table>
Acknowledgments

Firstly, thank you to my supervisor, Don Pollacco, for providing invaluable advice, guidance and help over the last 4 years. The chance to work with some of the leaders in this field that I enjoy so much has been incredible. Additionally, I’d like to thank the rest of the Warwick exoplanets group, every member of which was always happy and willing to give support whenever I asked for it, regardless of how basic the question.

Thank you, also, to my parents for always being interested in hearing about what I was working on, even if they didn’t understand it, and letting me know I could always count on them for support, even while teasing me for not having a ‘real job’.

I’d like to thank the friends I’ve made here, including, but not limited to, Alex, Ben, Darion, George and Hollie. Thank you for helping Warwick become somewhere I really felt at home, and providing food, films, quizzes, games and conversation about physics and everything else under the sun.

Finally, thank you to Kat. These years would have been much more stressful and difficult without your constant love and support. I could fill the rest of this thesis with how much it meant to me.
Declarations

I declare that this thesis is my own work except where stated otherwise and has not been submitted to this or any other academic institution for any other degree or qualification. Science chapters 3 through 7 are reproductions of my peer reviewed and published papers Cooke et al. [2020a], Cooke et al. [2018], Cooke et al. [2019], Cooke et al. [2020b], Cooke and Pollacco [2020] and Cooke et al. [2021]. Exact contributions from collaborators are stated at the beginning of each science chapter. Chapters 4 and 9 contain work presented in a poster at the UK EXOplanet Meeting on 15th April 2019 (UKEXOM19) and in a talk given at the PLATO ExtraSolar Planets meeting on 4th September 2019 (PLATO ESP 2019) respectively.
Abstract

The discovery of exoplanets is a key astrophysics goal. I begin in Chapter 3 with details of two new planets from the WASP survey; WASP-150b, a 1.1R_J hot Jupiter on a 5.6 day orbit, and WASP-176b, a 1.5R_J hot Jupiter on a 3.9 day orbit.

The exploration of longer period systems, such as those more similar to Earth, requires more effort. One method, explored extensively in this thesis is the use of monotransits. In Chapters 4 and 5 I describe simulations to predict the yield of monotransiing systems from TESS. The results indicate a significant number of detections, on the order of several hundred per hemisphere, with a key result being that monotransiits outnumber multitransiits for periods $\geq$ 30 days.

The use of multiple instruments to characterise exoplanet systems is common practice and in Chapter 6 I explore using CHEOPS to follow-up monotransiits identified in TESS. I show that CHEOPS could follow-up 90% of the monotransiits and would observe additional transits for the majority.

Monotransiits, and exoplanets in general, are usually followed-up with a combination of photometry and spectroscopy. In Chapter 7 I examine the idea of using these two methods to follow-up poorly constrained exoplanet systems in a general manner. I show that photometry and spectroscopy are each preferred for different regions of exoplanet parameter space.

Chapter 8 extends the TESS monotransiit simulations and explores using follow-up on the recurring systems. I show that $\sim$400 planets will transit once in each of the primary and extended TESS missions but each has an average of 38 period aliases. I show that photometric and spectroscopic observations are both capable of resolving these aliases.

Finally, Chapter 9, details the efforts of the NGTS monotraniit working group to identify monotransiits in the TESS data and characterise them using follow-up techniques. I describe the methods used and some of the key results including NGTS-11b.
Abbreviations

BLS  Box Least Squares
CBV  Co-trending Basis Vector
CCD  Charge Coupled Device
CHEOPS  CHaracterising ExOPlanet Satellite
CTL  Candidate Target List
CTOI  Community-identified TESS Object of Interest
DIA  Difference Imaging Analysis
EBLM  Low Mass Eclipsing Binary
FFI  Full Frame Image
HARPS  High Accuracy Radial velocity Planet Searcher
MCMC  Markov Chain Monte Carlo
NGTS  Next Generation Transit Survey
PSF  Point-Spread Function
RV  Radial Velocity
SNR  Signal-to-Noise Ratio
TESS  Transiting Exoplanet Survey Satellite
TFOP  TESS Follow-up Observing Program
TIC  TESS Input Catalogue
TOI  TESS Object of Interest
TTV  Transit Timing Variation
WASP  Wide Angle Search for Planets
R\(_\oplus\)  Earth Radius
M\(_\oplus\)  Earth Mass
R\(_J\)  Jupiter Radius
M\(_J\)  Jupiter Mass
R\(_\odot\)  Solar Radius
M\(_\odot\)  Solar Mass
AU  Astronomical Unit
Chapter 1

Introduction

So, then, mono means “one”,
and transit means “transit”
and that concludes our intensive four-year course.

- Lyle Lanley (paraphrased), The Simpsons, S04E12

1.1 Exoplanet history

Planets other than the Earth have been known about for thousands of years. Excluding Earth, 5 of the solar system planets, Mercury, Venus, Mars, Jupiter and Saturn are visible with the naked eye and the earliest known measurements of them come from Ancient Babylonia around 17th century BC [Reiner and Pingree, 1975]. The same planets were first thought of in a heliocentric model by Aristarchus of Samos around 300 BC [Heath and Aristarchus of Samos, 1913]. The non-naked eye planets had to wait for the invention of the telescope in 1608 which led to the discovery of Uranus in 1781 by Herschel [Pasachoff and Olson, 2011] (though technically a naked eye planet Uranus was sufficiently dim to have been thought to be a star before the use of a telescope) and Neptune in 1846 by Galle [Galle, 1846] (based on calculations performed by Le Verrier). Exoplanet discovery is a fairly recent development, although their existence had been theorised centuries before by Italian philosopher Giordano Bruno in the 16th century [Bryan, 1904] and Isaac Newton in his 1713 General Scholium [Newton, 1999].

The first officially discovered exoplanet was found by Wolszczan and Frail [1992] who discovered two exoplanets orbiting the pulsar PSR 1257+12. Subsequent follow-up identified a third planet in 1994 [Bisnovatyi-Kogan, 1993]. Though this was the first confirmed detection earlier candidates were suggested with some having been confirmed in recent years. For example, in 1988 astronomers Cambell, Walker and Yang suggested the
presence of a planet around Gamma Cephei based on radial velocity measurements [Campbell et al., 1988]. At the time this announcement was met with scepticism but was eventually confirmed to be a planet in 2003 [Hatzes et al., 2003]. The first evidence of exoplanet systems dates back even further, to a spectrographic image of the star Van Maanen 2 from 1917 which reveals calcium absorption indicating planetary material, however this realisation was not made until 2016 [van Maanen, 1917; Farihi, 2016].

The first detection of a planet around a main-sequence star was made in 1995 by [Mayor and Queloz, 1995]. The planet, a hot Jupiter orbiting the G-type star 51 Pegasi, was discovered using radial velocity measurements and its discovery led to a share of the 2019 physics Nobel prize for Mayor and Queloz. Finally, the first transiting exoplanet was discovered in 1999 by two independent teams who detected transits of a radial velocity planet around the star HD 209458 [Charbonneau et al., 2000; Henry et al., 1999]. The utilisation of these two methods, radial velocity and transits, have led to an explosion in exoplanet discovery since the late 90’s with over 4000 confirmed exoplanets currently known.

1.2 Discovery methods

Since the first detection of an exoplanet in 1992 [Wolszczan and Frail, 1992] and the first detection of a exoplanet around a sun-like star in 1995 [Mayor and Queloz, 1995] the field has exploded with many thousands of detections in the last 25 years. Figure 1.1 shows the year of detection and detection method of over 4000 currently known exoplanets.

Although dominated by the transit and radial velocity methods there are a whole host of methods for discovering exoplanets. A selection of these are described below.

1.2.1 Transit method

The transit method is currently the most prolific method of exoplanet discovery, responsible for over three quarters of all confirmed exoplanets (3266/4292) at last count. This is mainly due to large, photometric surveys such as WASP and Kepler which are discussed in sections 2.2.2 and 2.2.3 respectively. These surveys are capable of monitoring thousands of stars simultaneously, increasing the chances of detecting a transiting planet.

At the most basic level the transit method is designed to measure the change in brightness of a star caused by the passage of a planet between said star and the observer. This measurement requires high precision photometry since the drop in flux (and therefore brightness) caused by a planet transiting its host star is small, being proportional to the ratio of the areas of the planet and the star, given by

1https://exoplanetarchive.ipac.caltech.edu/
\[ \frac{\Delta F}{F} = \left( \frac{R_p}{R_\star} \right)^2 \]  

(1.1)

where \( R_p \) and \( R_\star \) are the radii of the planet and star respectively, \( F \) is the flux measured out of transit and \( \Delta F \) is the change of flux measured during transit. To put this into perspective, Jupiter transiting the sun produces a 1% drop in flux whereas the Earth produces only a 0.008% change.

On top of the difficulty caused by trying to detect such a small effect, transits are geometrically rare. For a transit to be observed the observer must lie in, or very close to, the orbital plane of the planet else the planet will not block any light from its host star. If we define a transit as occurring when any amount of the planetary disc occults any amount of the stellar disc then we require our inclination to satisfy \( i \geq i_{\text{crit}} \) where \( i_{\text{crit}} \) is given by

\[ \cos (i_{\text{crit}}) = \frac{R_p + R_\star}{a} \]  

(1.2)

where \( a \) is the orbital separation of star and planet. This equation necessitates an inclination of very close to 90°. From this equation we see that, geometrically, planets with smaller orbital separations, and therefore shorter periods, are more likely to be observed using the transit method. This geometry gives rise to an additional quantity, the impact factor, \( b \). The impact factor, measured as a fraction of the stellar radius, is the height above the stellar equator of the planet centre during transit and is given by the equation
Figure 1.2: Transiting planet geometry. Taken from https://www.paulanthonywilson.com/

\[ b = \frac{a \cos i}{R_*}. \]  

\(|b|\) ranges from 1 + \frac{R_p}{R_*} for a truly grazing transit to 0 for a perfectly central planet. Figure 1.2(a) shows this geometry.

Measuring the flux from a star as a function of time produces a light curve and witnessing a transit in the light curve allows us to make inferences about the planet. The depth of the feature is a proxy for planetary radius and, if observed for long enough, the separation between events gives the orbital period. The transit duration can also be related to properties of the planet-star system using geometrical arguments. If we assume a circular orbit then the transit duration is related to the length of the section of the orbit for which the planet is between the star and the observer. The geometry of this is shown in Figure 1.2(b) and gives the relation

\[ T_{\text{dur}} = P \frac{\alpha}{2\pi} = P \frac{\arcsin\left(\frac{l}{a}\right)}{\pi} = P \frac{\arcsin\left(\sqrt{(R_p + R_*)^2 - (bR_*)^2} / a\right)}{\pi} \]  

(1.4)

where \(P\) is the orbital period and \(\alpha\) and \(l\) are defined as in Figure 1.2(b).

An idealised transit light curve and corresponding geometry is shown in Figure 1.3.

In this figure the four numbered circles represent the four transit contact points, points at which the edge of the planet and star coincide. The solid line corresponds to a transit with impact parameter \(b = 0\) whereas the dotted line corresponds to one with \(0 < b < 1\). \(t_T\) in this figure is analogous to \(T_{\text{dur}}\) from equation 1.4 whereas \(t_F\) is the full disc transit time, the time for which the full planetary disc is eclipsing the stellar disc. This figure displays the idealised transit shape; a linear decrease in flux as the planet moves on to the stellar disc, a constant lower flux while the planet is fully on the stellar disc and a linear increase in flux as the planet leaves the stellar disc. In reality there is one more effect
Figure 1.3: Ideal transit light curve and geometry, reproduction of Figure 1 from Seager and Mallén-Ornelas [2003].
that must be accounted for to produce a realistic transit shape, limb darkening.

Limb darkening is the name given to the phenomenon that causes stars to appear brighter at the centre than towards the limb. This is actually the result of a visual effect by which light from the centre of the disc travels radially towards an observer whereas the closer to the limb one observes the more obliquely light must travel. The light we see from a star comes from the same optical depth which corresponds to a deeper radius near the disc centre. Since the temperature of a star increase radially inwards so too does the brightness of light from that depth. This results in a decreasing brightness as one moves outwards from the stellar disc centre. This phenomenon then causes a visible effect on a transit light curve. Since the centre of a stellar disc is brighter a planet will block a greater proportion of the total flux as it transits the central part of the disc compared to the edges. This results in a more ‘u’ shaped transit bottom than the perfectly flat line seen in Figure 1.2. Limb darkening is a function of wavelength and is usually modelled using one a range of laws, including linear, logarithmic, quadratic and cubic [Csizmadia, 2018]. Figure 1.4 shows the change in transit shape as a result of limb darkening. It is key to note that the transit depth increases by approximately 0.15%.

As a discovery tool the transit method is most sensitive to planets with large radii relative to their host star, this leads to the largest signals. This is then tempered by signal to noise constraints; the shallower signals are more easily detected around brighter hosts. Finally, as mentioned above, the transit method has a geometric bias towards close in, short period planets.

**Multitransits and monotransits**

With the fundamentals behind transits understood, transit discoveries can be separated into two categories, multitransits and monotransits. Multitransiting planets are those that exhibit multiple transit events per light curve, this generally requires the host star flux to be monitored for longer than 1 orbital period (although it is possible to luckily catch two or more transits without the intervening data). Monotransiting planets however are those for which only a single transit is caught during an observing campaign. There are some key differences between what we can learn from the light curve of a multitransiting vs monotransiting system.

The most important difference is that multitransits give the orbital period immediately. We simply fold the light curve until the transits match up in phase and the folding step gives us our period (assuming there are no missing transits between those observed). Multitransits tend to favour shorter periods as the shorter the period is, the greater the chances that multiple transit events are caught within an observing baseline. Because of the increased geometric probability that a short period system will transit (see section 1.2.1), multitransits
Figure 1.4: Limb darkening comparison, reproduction of Figure 1.3 from Coughlin [2013]. The solid and dotted lines show the transit with and without limb darkening.
make up the vast majority of confirmed transit discoveries.

In the cases of single transiting systems, or monotransits, measuring the period is much harder. One way of doing this is to rearrange equation 1.4 for period. This gives an equation involving $a$, a parameter not obviously measurable from a transit light curve. However, using Kepler’s third law, and under the assumption that a transit duration is a small fraction of an orbital period, the orbital period can be found from the following equation,

$$
P = \frac{M \cdot G \pi (t_T^2 - t_F^2)}{32 \Delta F^2}. \quad (1.5)
$$

This equation uses the nomenclature described in Figure 1.3. $M$ is the stellar mass and $G$ is the gravitational constant. The full derivation of the equation is shown in Seager and Mallén-Ornelas [2003]. From the equation we see that to measure an accurate period from a single transit we require a light curve of sufficient precision to measure $t_T$ and $t_F$ separately as well as accurate knowledge of the stellar density, $\frac{M \cdot R}{G \pi}$. This is generally only possible for the clearest transits but equation 1.5 gives us an upper limit on period, assuming the light curve is not precise enough to measure $t_F$ accurately,

$$
P \leq \frac{M \cdot G \pi t_T^3}{R^3 \Delta F^2}. \quad (1.6)
$$

1.2.2 Radial velocity

The radial velocity method of exoplanet discovery is the second most effective source of new discoveries and has only been surpassed by the transit method in the last 5-10 years. This method relies on the fact that, in practice, both elements of a planet-star system orbit the system barycentre. This means that there is a movement of the stellar host that is induced by the presence of a planet. Measuring this motion takes advantage of the Doppler effect. When a light source is moving either towards or away from an observer emitted light is either red- or blue-shifted to a different wavelength. Measuring this shift allows the determination of the velocity of the motion. The equation governing this velocity is given in Perryman [2018] and reproduced here,

$$
K_\star = \left(\frac{2 \pi G}{P}\right)^{1/2} \frac{M_p \sin i}{(M_\star + M_p)^{1/2}} \frac{1}{\sqrt{1 - e^2}}. \quad (1.7)
$$

Here $K_\star$ is the semi-amplitude of the velocity of the stellar host and $e$ is the eccentricity. Depending on the size of the planet and star this measurement can be extremely hard to measure. For example, Jupiter induces a $K_\star$ value of 12.7 ms$^{-1}$ on the sun whereas the Earth produces only 0.09 ms$^{-1}$. Figure 1.5(a) shows a typical radial velocity signal, in this
Creating an RV curve requires taking individual spectra of a host star at various orbital phases. Each spectrum is cross-correlated against a template spectra at rest and the wavelength offset is converted to a host velocity via the relation

\[ \nu_r = \frac{\lambda_o - \lambda_e}{\lambda_e} c. \]  \hspace{1cm} (1.8)

\( \lambda_o \) and \( \lambda_e \) are the observed and emitted wavelengths respectively with \( \nu_r \) being the line-of-sight velocity and \( c \) the speed of light. This equation makes it clear that we can only measure stellar motion along the line-of-sight. This means that the measured \( K_\star \) is at maximum for \( i = 90^\circ \) and falls to zero for \( i = 0^\circ \).

Since the reflex motion of the star is dependent on the mass of the planet we can use radial velocity data to measure this mass. Using equation 1.7 we see that only the minimum mass, \( M_p \sin i \), can be measured directly, we require an external measure of inclination to get the true planetary mass. Another factor of note is the impact of eccentricity \( e \) on radial velocity. Figure 1.5(a) shows a negligible eccentricity, resulting in a sine wave shape but RV signals for significantly eccentric planets are very different. Figure 1.5(b) shows the RV curve for HAT-P-2, a planet with \( e = 0.5 \).

The radial velocity method obviously favours high mass planets, and, in particular, systems where the planetary mass is large relative to the host mass. Additionally, as with the transit method, the radial velocity method favours short period, close in planets. A benefit
compared to the transit method however is that the radial velocity approach is sensitive to a wider range of inclination values, although the signal is maximum at $i = 90^\circ$.

1.2.3 Other methods

Though the transit and radial velocity methods are the two most productive in terms of number of planets discovered there are various other methods employed as well, some of which are mentioned briefly here.

Astrometry

Astrometry is the oldest method of searching for exoplanets and was originally used to identify binary star systems. Astrometry relies on a similar idea to radial velocity. The presence of a planet causes a star to move about the common centre of mass of the system but instead of measuring the velocity of the motion, astrometry tries to measure the actual movement of the star on the sky. For many star-planet systems this motion is very small and often the centre of mass is within the star itself. Because of the high precision required astrometry is currently responsible for only 1 planetary discovery, DENIS-P J082303.1-491201 b [Sahlmann et al., 2013]. However, ESA’s Gaia mission [Gaia Collaboration et al., 2016] is expected to discover thousands of planets via astrometry over its lifetime. Astrometry is different to both the transit and radial velocity methods in that it favours planets on long periods, since these create the largest reflex motion of the host, and perpendicularly aligned systems, since these create the largest component of the signal perpendicular to the line-of-sight, the only component that can be measured. A disadvantage of the method is that the further away a star is, the smaller its angular displacement due to a companion. Therefore this method is most sensitive to planetary systems close to our own.

Transit timing variations

The method of transit timing variation (TTV) for planetary discovery is a unique method as it can only be used to find planets in a system where at least one other planet (or stellar companion) is already known. For a single planet orbiting a star there will be one fixed period (ignoring orbital decay effects which are usually negligible) and the transit will reoccur at an expected rate. In cases where the system contains another planet however, this is not necessarily true. The second planet exerts a gravitational force on the first, causing it to accelerate and decelerate in its orbit which may cause its transit to occur earlier or later than predicted. Precise measurements of these TTVs can reveal the presence of an additional planet and modelling of the size of the signals can be used to infer planet masses [Lithwick et al., 2012]. This method was first used successfully on the Kepler-19c system in which
the first planet was detected via transits and the second was inferred through the transit vari-
ations of the first [Ballard et al., 2011]. There are currently 37 confirmed planet discoveries
via this method. The sensitivities of using TTVs as a detection method are not as clear as
other methods since multiple aspects of the system architecture interact. However, it has
been seen that the TTV method favours larger mass planets, which create larger gravita-
tional effects, and particularly, planets in resonant orbits. Additionally, the TTV signal can
be amplified in cases of large inclinations or eccentricities [Nesvorný, 2009].

Direct imaging

Directly imaging a planet, separately from its host star, is a huge task as the star is usually
many orders of magnitude brighter than the planet. For example, the sun is around $10^9$
times brighter than Jupiter and $10^{10}$ times brighter than Earth (at visual wavelengths; the
effect is lessened, but still large, in the infrared). For direct imaging to be successful specific
conditions are usually required. First, a coronagraph is used to block as much stellar flux
as possible, secondly the planet is required to have a large semi-major axis to avoid being
overwhelmed and thirdly the planet is required to be hot enough to produce significant
infrared radiation that can be detected independently. The first planet discovery using this
method was 2M1207b [Chauvin et al., 2004] and there are currently 51 confirmed planetary
detections using this technique. This method is particularly sensitive to massive, young
planets on large orbits. Young planets have yet to cool after formation so are relatively
bright compared to their hosts (especially in the infrared).

Gravitational microlensing

Gravitational microlensing as an exoplanet discovery method is an extension of the general
relativistic result that light is bent by massive objects. In gravitational lensing experiments
the light from a distant source is bent around a massive foreground lens, such as a galaxy,
producing multiple distorted, magnified and brightened images of itself. For microlensing
the lens is much less massive, for example, a star or planet, meaning the distorted images are
not visible but the brightening of the source can still be measured. The degree of brightening
is related to the mass of the lens. A stellar lens produces a single peak in brightness of the
source but if the lens hosts a sufficiently massive planet then the source is brightened in an
additional peak. Analysis of this double peak structure can reveal a planetary companion
and allow measurements of its mass. The first planetary discovery using this method was
OGLE 2003-BLG-235b [Bond et al., 2004]. There are now 101 confirmed exoplanets that
have been discovered via this method. Gravitational microlensing is most sensitive to high
mass planets sufficiently separated from their hosts but can be used for planets down to an
Earth mass on orbits up to a few AU. This method is one of the few capable of reaching this region of parameter space.

1.3 Exoplanet knowledge

Using the techniques above, the discovery of over 4000 exoplanets over the last 25 years has shaped our understanding of the field. Analyses of the populations, alongside sophisticated physical models, have given us more information than ever about exactly how these myriad systems came to be.

1.3.1 Formation and evolution

Protoplanetary discs

Planets are thought to form in protoplanetary discs surrounding young stars [Williams and Cieza, 2011]. These discs are a natural consequence of angular momentum conservation during gravitational collapse. The start of this formation begins when a region of a molecular cloud satisfies the Jean criterion and becomes unstable to gravitational collapse [Tsiklauri, 1998]. As the cloud contracts the conservation of angular momentum leads to an increasing rotation rate. This rate is naturally greatest at the equator of the forming protostar and thus the material there becomes smeared out in to a protostellar disc structure. As the disc forms and evolves it continues to accrete material on to the central star. At this point the star is undergoing its T-Tauri phase, characterised by excess infrared emission [Herbig, 1962; Chen et al., 2020]. This emission is a direct result of the protoplanetary disc which absorbs radiation from the star and re-emits it as infrared.

Throughout this process the molecular cloud is still collapsing, transferring more mass to the protostellar disc. However, it has been seen that the disc mass is insensitive to its evolutionary state [Looney et al., 2003; Jørgensen et al., 2009]. This therefore implies that the accretion of material on to the star is efficient. Accretion is caused by a combination of viscous and instability effects. Viscosity in the disc allows the transport of angular momentum to external material, allowing close in material to accrete on to the star. Additionally, gravitational instabilities within the disc have been shown to create episodic accretion events which prevent the disc growing faster than the star [Zhu et al., 2009; Vorobyov and Basu, 2010]. This accretion creates the separation between class I and class II discs and the transition from protostellar to protoplanetary.

The final stage of disc evolution is the dispersal, the time span of which constrains planetary formation which cannot occur without a disc. The oldest observed discs are found to be $10^6$ – $10^7$ years old implying that the majority of disc dispersal occurs within this time-
Dispersal is thought to be the result of a combination of effects including accretion to the central star, planet formation and photoevaporation where intense radiation from the star can evaporate gas from the protoplanetary disc which in turn drags small dust particles along with it [Hollenbach and Adams, 2004; Gorti et al., 2009].

Planetesimals and planets

The formation of planets must occur sometime between the formation and dispersal of the protoplanetary disc. The most commonly accepted theory for the formation of planets is via a core-accretion process [Matsuo et al., 2007], although other avenues have been suggested (e.g. gravitational collapse [Boss, 1997]). In the core accretion model tiny grains that condense from the molecular cloud accrete via multiple collisions and build their way up to planet sized bodies. For the very smallest grains, sub-micron in size, particles are held together by weak, short range, Van der Waals forces [Weidenschilling, 1980]. The weakness of these forces means that particles are prone to fragmentation, especially in more turbulent zones.

There have been multiple laboratory based experiments to quantify the way in which these tiny particles can begin to coalesce [Blum and Wurm, 2008; Gütlinger et al., 2010]. In any grain-grain collision there are three possibilities. The grains can: coalesce (the preferred outcome for core-accretion); rebound; or fragment (the most detrimental to the core-accretion). The probability of coalescence is related to the grain properties as well as the speed of collisions, with slower collisions being preferable for accretion [Blum et al., 2006]. Unfortunately, according to simulations, as grain size increases so does the average collision speed leading to more fragmentation. Additionally, as the grains reach sizes on the order of 1 metre they are susceptible to accretion on to the star since it is at this size that grains feel the largest gas-drag effects, increasing their in-spiral rate [Weidenschilling, 1977a,b]. A solution to both of these issues is that simulations suggest that grains can become trapped at disc radii which correspond to pressure maxima. The pressure maxima can be caused by disc transitions, such as the ice line (beyond which the disc is sufficiently cool for volatiles to condense) or between small eddies in turbulent discs [Brauer et al., 2008; Cuzzi et al., 2008]. Grains trapped in these maxima no longer fall inwards towards the star and also have much lower relative velocities, allowing for more accretion collisions.

These collisions can allow the grains to grow in size to become planetesimals, ~km sized planetary building blocks. The larger the planetesimal the stronger its gravitational cross section and therefore the more easily it can continue to accrete. This leads to runaway accretion until these budding protoplanets have cleared their respective orbits up to a Hill sphere in radius [Tanigawa and Ohtsuki, 2010]. Figure 1.6 shows the first robust detection of planetary-mass companion in a protoplanetary disc, observed by Müller et al. [2018].
Further out planets are thought to grow more massive in size due to the availability of condensed volatiles not available in the inner disc [Boss, 1988]. The largest bodies are then even capable of accreting significant amount of gas leading to rich atmospheric envelopes.

**Evolution and migration**

There is plenty of evidence that shows that the formation of planets is not the end of the story. Based on the theory of planetary formation discussed above there is no viable scenario for a Jupiter mass planet to form at a period of just a few days but many such systems are observed. This means that at some point between formation and the present day these planets must have migrated inwards from beyond the snow line where they presumably formed. There are additional systems which imply the necessity of planetary migration such as those planets with extreme eccentricity, or multi-planet systems which don’t share a common inclination [Rein, 2012; Dittkrist et al., 2014]. Much theoretical work on this problem suggests that the migration can happen either prior to the dispersal of the gas disc.
or afterwards.

Migration prior to the dispersal of the planetary disc is separated into two regimes based on the mass of the planet [Bate et al., 2003]. For low mass planets ($< 0.1 \text{ M}_J$) regime 1 migration occurs where the interaction between the planet and the dust disc at particular resonant radial locations can set up spiral density waves in the disc. The planet can then exchange angular momentum with the disc and move in or out depending on whether momentum is lost or gained [Armitage and Rice, 2005]. The rate of migration is a function of the torques created by inner and outer Linblad resonances [Goodman and Rafikov, 2001]. For more massive planets the immediate area has been cleared meaning that regime 1 disc migration is halted [Lubow et al., 1999] and regime 2 migration can occur. The edges of the cleared area are set by a balance between angular momentum transport and viscosity. Thus as the viscosity of the disc evolves with time, the gap, and planet along with it, can migrate inwards [Lin and Papaloizou, 1986].

Post-disc dispersal planetary migration is caused by interactions between multiple giant bodies. There are many scenarios that can take place here. One such interaction is that an external body scatters an inward planet closer to the host star while it itself is ejected from the system. The now close-in planet can then circularise much closer to the star than otherwise possible [Petrovich et al., 2014]. An alternative method is that an external massive body, perhaps a giant planet or unbound star, may set up Kozai-Lidov oscillations in a closer planet [Kozai, 1962; Lidov, 1962]. If the perihelion of the inner planet is sufficiently close to its host it will circularise and result in a giant planet on a very short period [Valsecchi and Rasio, 2014]. Combining these methods can produce simulated populations similar to those observed.

The evolution of such systems does not end here however. Due to continuing tidal interactions between the planet and star, enhanced due to the small distances between them, hot Jupiters are unstable to orbital decay. These tidal interactions lead to energy dissipation and the transfer of angular momentum from the orbit of the planet to the rotation of the star. Eventually, when the angular momentum in the orbit becomes sufficiently low, the planet will be engulfed by the star [Rasio et al., 1996; Levrard et al., 2009]. This is a very long-term process and thus it is hard to measure directly in all but a few cases. The most extreme example known is that of WASP-12b which shows an orbital decay rate of $\frac{dP}{dt} \sim 10^{-9}$ [Patra et al., 2020]. Though not much direct evidence exists this theory is further supported by population analysis results including a dearth of giant planets with periods less than 1 day [Ogilvie, 2014] and the anomalously fast rotation of some hot Jupiter host stars [Penev et al., 2018].
1.3.2 Occurrence rates

One of the key goals for large-scale exoplanet discovery surveys is to aid in the determination of the occurrence rate of planets. That is, to determine on average how many planets of a particular type a given star will host. Within this search there is great interest in eta Earth ($\eta_\oplus$), the mean number of rocky, Earth sized planets within their host stars habitable zones [Haghighipour, 2015]. Calculating occurrence rates requires the observation of a large number of stars with a well understood and robust method (be that photometric or spectroscopic). The observations are then searched in a consistent manner for planet signals. The number of found planets must then be corrected for bias and incompleteness in the detection and search methods [Bryson et al., 2020]. For example, transit surveys will only catch systems with $i \sim 90^\circ$ therefore, for every observed planet there must be many more on un-observable orbits. Different methods will also require different levels of precision to confidently detect planets [Burke and McCullough, 2014]. These corrected rates can then be separated into specific rates for different types of planets, e.g. giants or terrestrials.

Occurrence rate calculations can only be realistically attempted when a significant number of stars have been searched and planets found, thus the first calculations were based on radial velocity searches, the original workhorse of exoplanet discovery. One of the first studies to attempt this was Marcy et al. [2005] who found that $\sim 12\%$ of FGK stars host a gas giant within 20 AU and $\sim 7\%$ host a gas giant within 5 AU. This study also showed the strong relation between planet occurrence rate and increasing stellar metallicity. Figure 1.7 shows their planets as functions of $m \sin i$ and semi-major axes. Other radial velocity surveys have shown the details of multi-planet systems [Wright et al., 2009], the increased occurrence of lower mass planets (at periods of less than 50 days) [Howard et al., 2010] and the preference of giant planets to orbit lower mass hosts [Reffert et al., 2015] (excluding M dwarfs, which are shown to host fewer giant planets than solar-mass stars [Johnson et al., 2010]).

In recent years, with the huge numbers of planet discoveries via photometry, transit surveys have been used to calculate occurrence rates. The first sufficiently large and sensitive transit survey was Kepler, launched in 2009 [Borucki et al., 1997, 2010]. Some of the first occurrence rates coming from the Kepler survey were produced by Fressin et al. [2013] who analysed the first 6 quarters of Kepler data and found 16.5% of FGK stars host a planet with a radii $0.8 - 1.25 \, R_\oplus$ and a period less than 85 days. A similar analysis was carried out by Dressing and Charbonneau [2013] looking exclusively at M dwarfs which found an average of 0.51 Earth-sized planet per star with 0.15 Earth-sized planets in the habitable zone per star. Burke et al. [2015] used 16 quarters of Kepler data to find 0.77 planets per GK star with radii $0.75 - 2.5 \, R_\oplus$ and periods 50-300 days. Limiting the results to radii and periods within 20% of Earth’s they find 0.01 planets per star. More recently Hsu
et al. [2019] use Kepler in conjunction with Gaia to infer an $\eta_\oplus$ of $< 0.27$, and suggest that future surveys should plan for a value between 0.03-0.4. Figure 1.8 shows the occurrence rate around FGK stars as a function of radius and period produced by this work. Using a similar data set Bryson et al. [2021] examine main sequence dwarf stars and define $\eta_\oplus$ as the rate of rocky, habitable zone planets with radii $0.5 - 1.5 \, R_\oplus$ orbiting stars with effective temperatures between 4800 and 6300 K. This work finds a value of $\eta_\oplus$ between 0.37 and 0.60 assuming a conservative habitable zone.

### 1.4 Exoplanet characterisation

The detection and confirmation of exoplanets is comprised of both discovery, covered in section 1.2, and characterisation. Characterisation is the method of determining the exoplanets physical characteristics use a range of techniques as discussed below. Once characterised, some natural categories of exoplanets arise which are also mentioned.

#### 1.4.1 Physical parameters

Listed below are some of the key physical parameters used in the characterisation of exoplanets and the most common methods for measuring them.

**Radius**

Exoplanet radius is most commonly measured via the transit method. As discussed in section 1.2.1 the drop in flux during a transit is proportional to the ratio of the square of the radii of the planet and star (equation 1.1). Therefore, assuming the stellar radius is known,
Figure 1.8: FGK star planet occurrence rates. The colour of cells shows an occurrence rate that is normalised by the size of the cells in period and radius space. Taken from Figure 2 of Hsu et al. [2019].

The planetary radius is easily inferred from a transit light curve. However, it should be noted that there are ambiguities to this method. For example, due to limb-darkening effects, planets that transit the centre of their hosts block more light. If not properly modelled this may lead to an over- or under-estimation of radius depending on whether the planet transits the brighter centre or dimmer edge of its host. Additionally, the impact parameter $b$ has an effect: for $b$ values close to 1 only part of the planetary disc may transit the star. This results in less flux being blocked and an underestimated radius.

**Mass**

Planetary mass can be measured in a variety of ways, mainly based on how it affects the movement of its host. The most common method is through radial velocities where, using equation 1.7, planet mass can be determined from the semi-amplitude of the radial velocity signal. The particular subtlety in the method is that it actually only allows the measurement of the minimum mass $m \sin i$. To determine the true mass we require an independent measurement of the inclination $i$.

In a similar way, the stellar motion caused by a planet of a particular mass is the cause of exoplanet discoveries via astrometry therefore this method can also lead to a planetary mass measurement. Planet mass is also the cause of the gravitational microlensing effect thus can be inferred from the size of the effect. Finally, in planetary systems with multiple bodies the size of any TTV signals can be modelled and used to infer the masses of non-transiting bodies.
Period

Orbital period is one of the few planetary parameters that is simply measurable from both transit and radial velocity measurements. Both the transit signal and the radial velocity signal are periodic with the period of the planets orbit. Folding the measured data from either method until the signals overlap is usually all that is required to measure period. This can be complicated in the case of long period systems or systems with multiple planet signals. In the case of the transit method period is hard to measure from a single transit (but not impossible, as shown in section 1.2.1) and if the coverage is non-continuous it is possible to miss transits and infer the wrong period. Partial phase coverage also affects radial velocity measurements. Modelling the period is made more challenging in cases where less than a full cycle is observed or phase coverage is spotty.

Orbital period can be further complicated by additional effects such as orbital decay or TTVs. In the case of TTVs the measured period of a planet will actually fluctuate around a central value due to gravitational effects from other planets in the system. For orbital decay the period will decrease as a function of time. For descriptions of the magnitudes of these effects see Nesvorný and Vokrouhlický [2016] and Rasio et al. [1996] respectively.

Orbital separation

Measuring the orbital separation of a planet around its star is a direct result of Kepler’s third law;

\[ P^2 = \frac{4\pi^2}{G(M_\star + M_p)}a^3. \] (1.9)

Using this equation, along with the stellar and planetary masses and the orbital period (1.4.1) allows us to measure orbital separation, \( a \), directly.

In the special case of directly imaged exoplanets the orbital separation can be determined from the angular separation between star and planet and the distance from the observer to the system.

Density and composition

Measuring the bulk density of an exoplanet is trivial if the radius and mass are known since \( \rho_p \propto M_p/R_p^3 \). Using these measurements to infer an internal composition however, is more complex. It is generally assumed that one can make inferences of planetary composition based on the composition of its stellar host which can be inferred from spectra [Wang et al., 2019a; Spaargaren et al., 2019]. This method, however, just produces bulk compositions whereas in reality the radial stratification of elements can be vital for understanding the
evolution and behaviour of a planet [Foley and Driscoll, 2016]. In an attempt to understand the specific nature of exoplanet compositions, work usually focuses on our in-depth knowledge of the planets of our own solar systems and then attempts to extrapolate to exoplanet systems [Taylor, 1992; Newsom, 1995].

Gas giants are generally modelled as spheres of hydrogen and helium with a small fraction of heavier elements and complex molecules [Spiegel et al., 2014]. One key unknown for the composition of gas giants is whether or not they contain solid cores in their centre, in fact we are not even sure whether Jupiter has a solid core, though it is assumed that Saturn does [Wilson and Militzer, 2012]. Recent measurements from the Juno satellite [Bolton et al., 2017] suggest that Jupiter may contain a dilute core, occupying a significant fraction of the planets radius, though these models are strongly affected by the chosen equation of state [Wahl et al., 2017]. For more solid or terrestrial exoplanets the composition and internal structure is informed by 3 key equations. The mass of a spherical shell,

$$\frac{dm(r)}{dr} = 4\pi r^2 \rho(r),$$  \hspace{1cm} (1.10)

where $m$, $r$ and $\rho$ are the mass, radius and density of the exoplanet; the equation of hydrostatic equilibrium,

$$\frac{dP(r)}{dr} = -\frac{Gm(r)\rho(r)}{r^2},$$  \hspace{1cm} (1.11)

where $P$ is pressure; and the equation of state, which varies for each specific material. Using these equations, planets are often modelled as radially stratified spheres comprised of iron, silicates and water-ice [Zeng and Seager, 2008]. Different combinations of these three materials produce different mass and radius measurements. There is obvious degeneracy between the three fractions which means it is impossible to fully constrain compositions [Rogers and Seager, 2010]. Current work focuses in finding possible combinations of the three fractions that lead to the observed mass and radius which are usually presented via a ternary diagram [Valencia et al., 2007].

Atmospheres

The detection and characterisation of exoplanet atmospheres is a relatively young field within exoplanetology with the first real atmospheric detection occurring for HD 209458b in 2002 [Charbonneau et al., 2002]. Observations of exoplanet atmospheres can be carried out via transit photometry, Doppler spectroscopy and direct imaging [Madhusudhan, 2019]. Transit photometry is the most successful technique so far [Kreidberg, 2018] and can act in two ways. Firstly, during transit the light from the star passes through the planet atmosphere leading to absorption features. Subtracting the out of transit spectrum reveals the
planetary features. Similarly, just before secondary transit (of the planet behind the star) the spectrum contains a combination of stellar and planetary features, removing the stellar spectrum again reveals the planetary day-side features [Knutson et al., 2007; Stevenson et al., 2014]. Doppler spectroscopy is less commonly used, partly because of the extreme precision requirements. This technique measures the combined spectrum of the star and planet [Birkby, 2018]. Since the planet is moving faster its spectral lines shift by a greater degree than those of the star (on the order of $10^3$ time larger). This allows the stellar lines to be identified and subtracted leaving just the planetary spectra [Brogi et al., 2012; Cabot et al., 2019]. Finally, direct imaging measures the spectra of the planet directly by blocking the stellar spectra through use of a coronagraph [Fischer et al., 2014]. However, this is challenging based on the high contrast ratios required (see section 1.2.3).

Using different spectral ranges for observations means that, due to opacity effects, different regions of atmospheres can be probed. Figure 1.9 shows some key processes and at what heights/pressures in exoplanet atmospheres they may occur.

This results in different physical atmospheric processes being identifiable. A few key properties that are now detectable are mentioned here. Firstly, chemical composition can be inferred from the absorption and emission lines present in planetary spectra, these include elemental species (H, He, Na etc.) and molecular species (H$_2$O, CO, HCN etc.) [Crossfield, 2015]. The degree of suppression of spectral features at some wavelengths allows for the measurement of clouds and hazes in atmospheres [Deming et al., 2013; Madhusudhan et al., 2016]. Finally, we are able to measure the vertical temperature profiles of atmospheres by measuring spectral emission lines generated at different wavelengths, and
thus depths in the atmosphere. These profiles can lead to inferences of radiative processes and chemical and dynamical interactions [Madhusudhan, 2018].

1.4.2 Types of exoplanet

Based on the parameters described above a number of natural classifications of exoplanets arise in the data. Some of these are described here. Figure 1.10 shows the distribution in radius and period of all NASA exoplanet archive\(^2\) confirmed planets with radius and period constrained to 10% or better. The radii are separated into terrestrial, ice giants and gas giants as discussed below.

**Terrestrial**

Terrestrial planets, sometimes called telluric or rocky, are those comprised mainly of silicate rocks or metals. Within our solar system the terrestrial planets are Mercury, Venus, Earth and Mars, but many extra-solar terrestrial planets are also known. Depending on the specific definition of a planet used (geophysical vs IAU definition) some rocky moons may

\(^2\)[https://exoplanetarchive.ipac.caltech.edu/]
be referred to as terrestrial planets [Russell, 2013]. Terrestrial planets are generally thought to have formed via pebble accretion interior to the snow line (see section 1.3.1), explaining their lack of volatiles. Additionally, it is believed that most terrestrial planets have a secondary atmosphere. A secondary atmosphere is one formed by volcanic out-gassing or cometary impacts, having lost the primary atmosphere that was accreted from the protoplanetary disc [Dorn and Heng, 2018; Kral et al., 2020]. At least one terrestrial planet (Earth) has had its atmosphere further changed by biological life resulting in a tertiary atmosphere.

Terrestrial planets do not contain large quantities of ice or gas and are therefore limited in potential radius compared to ice or gas giants as discussed below. Various radii limits have been proposed [e.g. Fressin et al., 2013; Batalha et al., 2013] but here (in Figure 1.10) we assume an upper limit of 2 \( R_\oplus \). This boundary is not definitive however, there have been planets in this radius regime that have been shown to have significant ice or gas fractions [Kepler-138d, Kipping et al., 2014] but it is a good approximation. Other definitions focus on density instead, requiring \( \rho_p \geq 5 \text{ g/cm}^3 \) to be classed as terrestrial. The first confirmed terrestrial extra-solar planet was Kepler-10b [Batalha et al., 2011].

**Ice giants**

Ice giants are a type of giant planet comprised mainly of elements heavier than hydrogen and helium such as nitrogen, oxygen and carbon. They do however retain significant hydrogen/helium envelopes. Our solar system contains two ice giants, Uranus and Neptune. In this context ice can be any volatile with a freezing point \( \gtrsim 100\text{K} \) such as water, ammonia or methane [Podolak and Hubbard, 1998]. The formation of ice giants, especially in our own solar system is a matter of debate [Lambrechts et al., 2014]. To accrete sufficient ice and gas these planets must form beyond the snow line but if formation occurs too far from the star they are liable to be ejected through interactions with the interior, more massive, gas giants. One theory suggests a formation location between Jupiter and Saturn before gravitational interactions scattered the ice giants out to more distant orbits whereas other theories favour a gravitational instability based approach [Boss et al., 2002; Boss, 2003].

Ice giants contain small rocky cores surrounded by large reservoirs of ices before a final atmospheric layer of hydrogen and helium. In ice giants this atmospheric layer comprises \( \sim 20\% \) in mass in contrast to gas giants which are \( \sim 90\% \) hydrogen and helium [Stevenson, 2013; Helled and Guillot, 2018]. As before, the radii distribution of ice giants is not a definitive boundary and multiple values exist [e.g. Fressin et al., 2013; Batalha et al., 2013]. Here we take the ice giant regime to be between 2 and 6 \( R_\oplus \). Other studies, however, argue for an intermediate regime between terrestrial and ice giant planets called gas dwarfs, or transitional planets, at \( 1.7 - 3.9 R_\oplus \) [Buchhave et al., 2014].

Of particular note when discussing Neptune-like planets is the so-called Neptunian
desert. This is a region in parameter space where short period, Neptune-radius planets are very sparse (as can be seen in the middle left region of Figure 1.10). The exact nature and cause of this desert is not fully understood but current theories ascribe it to strong irradiation from the stellar host [Mazeh et al., 2016]. Ice giant planets at short periods are subject to evaporation of their atmospheres caused by the host star. The planets are not sufficiently massive to be able to maintain this atmosphere meaning planets in this range shrink in radius, becoming stripped, rocky cores, falling into the terrestrial radius regime. In recent years some planets have been found in this desert however [e.g. West et al., 2019; Armstrong et al., 2020].

Gas giants

A gas giant is a type of giant planet that contains \( \sim 90\% \) hydrogen and helium by mass [D’Angelo and Lissauer, 2018]. Gas giants were initially thought to comprise all giant planets until a separate regime was defined for ice giants as described above. Gas giant composition is generally thought to contain a small, rocky (possibly molten) core surrounded by hydrogen and helium [Lambrechts and Johansen, 2012]. The hydrogen is differentiated into molecular hydrogen, around the surface of the planet, and metallic hydrogen, comprising the bulk of the mass. It is the presence of metallic hydrogen that sets these planets apart from ice giants which do not contain this form of hydrogen [Chabrier et al., 1992]. Though conventionally thought of as a gas, the majority of hydrogen in a gas giant planet is actually above the critical point at which there is no distinction between liquid and gas.

As with the previous regimes the radius boundaries for gas giants are not definitive [e.g. Fressin et al., 2013; Batalha et al., 2013] but here we use a lower limit of 6 \( R_\oplus \). At the upper end of the scale gas giants are bounded by the transition to brown dwarfs. Brown dwarfs (sometimes referred to as failed stars) are sub-stellar objects dense enough to exhibit deuterium fusion but not enough to trigger sustained fusion of hydrogen. The boundary between gas giants and brown dwarfs is thought to be around 13 \( M_J \) [Burrows et al., 2001]. Gas giant formation is fairly well understood with the general belief being the run-away accretion of gas on to rocky cores, out beyond the snow line of a protoplanetary disc [Pollack et al., 1996] (although other theories exist; see section 1.3.1 and Boley [2009]). This mechanism implies the existence of gas giants only at long periods, there is, however, a significant population of close in gas giants, called hot Jupiters. Hot Jupiters are Jupiter mass gas giants at periods of <10 days [Masset and Papaloizou, 2003]. These planets are usually though to be the result of inwards migration of exterior giant planets, rather than forming in situ, which would have been difficult, due to the lack of volatiles that close to a host star [Michael et al., 2011; Bitsch et al., 2019]. Due to the ease with which they can be detected (through both transit and radial velocity signals) there are many such hot Jupiters
known [Wright et al., 2012] including the first extra-solar planet around a sun-like star, 51 Pegasi b [Mayor and Queloz, 1995]. Planets of this type can clearly be seen in the upper left portion Figure 1.10.

1.5 Thesis outline

This thesis continues in Chapter 2 with a description and discussion of the methods employed within the science chapters. The remainder of the thesis is split into 7 science chapters. Chapter 3 is a reproduction of Cooke et al. [2020a] and covers the discovery, confirmation and characterisation of two new hot Jupiter planets from the WASP survey, WASP-150b and WASP-176b. Chapter 4 is a reproduction of Cooke et al. [2018] and details the simulation of monotransit detections from the primary TESS mission and examines the feasibility of follow-up from a single ground-based location. Chapter 5 is a reproduction of Cooke et al. [2019] and discusses an improved simulation of TESS monotransits and its extension to the TESS extended mission. Chapter 6 is a reproduction of Cooke et al. [2020b] and explores the use of CHEOPS as a potential follow-up tool to monotransit targets detected in the TESS mission. Chapter 7 is a reproduction of Cooke and Pollacco [2020] and looks at exoplanet follow-up in a generalised way, comparing the benefits of photometric (NGTS) and spectroscopic (HARPS/CORALIE) follow-up across a range of exoplanet system parameter space. Chapter 8 is a reproduction of Cooke et al. [2021] and considers the problem of follow-up of TESS monotransits using both photometric and spectroscopic instruments, specifically looking at the problem of period aliases which is key to monotransits. The last science chapter, Chapter 9, outlines the work and current results of the NGTS monotransits working group. The thesis ends with conclusions in Chapter 10 and a list of all first- and co-author papers produced during this project in Chapter 11.
Chapter 2

Methods

2.1 Transit detection methods

Transiting exoplanets create periodic dips in flux, measurable in photometric light curves (as described in section 1.2.1). However, when it comes to exoplanet discovery it is unknown which of usually thousands of light curves may have a transit signal present. Therefore, methods are required to identify the transit signals and to attempt to limit the amount of data that must be searched by eye. The methods used may be different depending on the type of transit signals one is looking for (e.g. multitransits vs monotransits).

2.1.1 Multitransit detection

The most common transit detection method is the Box Least Squared (BLS) technique, pioneered by Kovács et al. [2002]. This method assumes a strictly periodic transit signal and attempts to fit 5 parameters (described here using Kovács et al. [2002] nomenclature and symbols) to the data: period, $P_0$; flux level in and out of transit, $H$ and $L$; the epoch of transit, $t_0$; and the fraction of period in transit $q$ (where $q$ is assumed to be small, typically 1-5%). For each chosen period the temporal data is phase folded on the period. The start and end phases of the transit are then iterated over allowing the transit to occur at any phase and at a range of $q$ values. $H$ and $L$ are calculated as the level interior and exterior to the transit start and end phases. For each iteration a chi-squared value is found between the data and the model with the minimum value being chosen for each period. This produces a BLS periodogram for the data. An example periodogram is shown in Figure 2.1.

In Figure 2.1 the BLS periodogram signal peaks at the true period but also at period aliases, that is, integer or inverse integer multiples of $P_0$ ($P_{0}/2, 2P_{0}, 3P_{0}$ etc.). These must be carefully accounted for when determining the true period. For an ideal periodogram the strongest peak is that of the true period (see Figure 2.1 whose true frequency is 0.2 d$^{-1}$).
practice however, some amount of human eyeballing is generally required to rule out aliases. This method can then be applied to thousands of light curves automatically allowing further study to focus on only the strongest detections.

In recent years attempts have been made to improve on the simplicity of BLS without sacrificing its computational speed. One of these, Circumbinary Box Least Squares (CB-BLS) is optimised to search for circumbinary planets [Ofir, 2008]. This method is required since circumbinary planets may not have periodic signals and may produce variable depth transits. Another attempt to improve upon BLS is the Transit Least Squared (TLS) method [Hippke and Heller, 2019]. TLS is broadly similar to BLS except that it uses a more realistic transit shape rather than a box. The default transit fit is shown in Figure 2.2 and is determined from the normalised average of thousands of real and verified transit shapes.

TLS accounts for ingress and egress times as well as limb-darkening effects. TLS is comparable in computational time to BLS while potentially being more reliable, especially in the case of small radius planets and has already been used for the discovery and statistical validation of multiple exoplanets [Heller et al., 2019a,b, 2020].

A separate transit detection method which should also be mentioned here is that of wavelet analysis. In brief, wavelet analysis is based around the fact that time-series photometry may be contaminated with white noise as well as temporally correlated red noise.

---

Figure 2.1: Example light curve with system parameters in the figure header. Also shown are the phase folded light curve and corresponding BLS periodogram. Taken from Kovács et al. [2002].

2http://exoplanets.org/
Both components are assumed to be time-invariant Gaussian processes, with the white noise being uncorrelated and the red noise having a $1/f^\gamma$ power spectral density. The technique relies on minimising likelihood functions, as more traditional methods do, but instead of using the raw data, the data is first transformed into a wavelet basis. The benefit of this method is that many realistic correlated noise models produce nearly diagonal covariance matrices in the wavelet basis (meaning individual variables are not correlated with each other). Additionally, there exist computationally fast methods for wavelet transforms (for details of wavelet transforms see e.g. Wornell [1996]; Press et al. [2007]). The basic method is to subtract a transit model from a data set giving the residuals, transform these residuals into a wavelet basis and then calculate the likelihood function in this basis. Repeating for a range of transit parameters and maximising this likelihood function will then determine the best fitting values. When compared to assuming a simple white noise contamination the wavelet analysis method performs better as the white noise only approach tends to underestimate fitted uncertainties. This method is not further utilised in this thesis so we do not go into detail here but for a more in-depth description see Carter and Winn [2009].

2.1.2 Monotransit detection

Detecting monotransit events requires a different approach to multitransit events. Since only a single transit is seen there is no periodicity to the light curve, thus testing periods
and folding the data is not possible. Instead, we require a method to locate individual transit-like dips below the out of transit flux levels.

**Difference Imaging Analysis monotransit search**

One such method, originally optimised for K2 data, is laid out in Osborn et al. [2016] and Osborn [2017]. This method is used in this thesis on the Difference Imaging Analysis (DIA) light curves produced by Oelkers and Stassun [2018] as discussed in section 9.2.

First, individual light curves are detrended for systematics as described in Armstrong et al. [2015]. Next, additional long-term variability is removed using third-order polynomials around a central window, iterated over the full light curve as in Armstrong et al. [2014]. The transit search is then carried out as follows. A box of fixed transit time and duration is placed on a light curve and the depth compared to the out of transit level is found. Events with depths above a given threshold are recorded. The box is then iterated though the light curve and the duration is increased from 1.95 hrs to 1.2 days in an attempt to cover the reasonable range of realistic durations. Detections above a chosen Signal-to-Noise Ratio (SNR) limit are then modelled twice. First, with a transit model, including depth, transit centre, duration and shape using chi-squared minimisation. The shape parameter is found using an array of transit models create by the *transit* Python package. The second model is a step model used to identify non-transiting signals such as end of observation ramps and flux drops with no ingress/egress etc. Each model is optimised using the relevant *scipy* package and a minimum reduced chi-squared value is found. The ratio between the values suggests which model is more likely for each event. The high SNR events which favour transits over non-transits are then the most likely monotransit events. The reliance on SNR, vital here to remove the vast number of false positives and increase confidence in the detections, hampers the detection of small planets which are low SNR. The technique used in the NGTS monotransit working group, described below, attempts to remove this bias.

**NGTS working group monotransit search**

Another variant on this method is that used by the NGTS monotransit working group, optimised for TESS Full Frame Image (FFI) light curves. This method is used by this project on the primary and extended TESS mission light curves produced by the NGTS monotransit working group as discussed in section 9.3. This method first detrends the light curves using individual Co-trending Basis Vectors (CBVs). CBVs are measurements of the most common trends found across all stars in a particular region of an image. Since these trends

---

3https://github.com/dfm/transit
are common to all stars they are assumed to be systematic effects and are therefore removed from the light curves [Kinemuchi et al., 2012]. Additionally a Savitsov-Goly filter is used to smooth the data without removing transit events [Savitzky and Golay, 1964]. Next the light curves are searched for data that are below the baseline flux by a certain threshold. The method requires at least 3 consecutive data points to be below this threshold so as to avoid triggering on individual drop out events that can appear in TESS data [Vanderspek et al., 2018]. Events meeting these criteria are then fit with a transit model and those with sufficient quality fits are flagged as monotransit candidates. It should be noted that individual events are treated separately allowing this method to locate multiple monotransit events in a single light curve if present.

2.2 Photometry

Photometry, from the Greek ‘photo’ (light) and ‘metry’ (measure) is a technique used in astronomy to measure the flux radiated by astronomical bodies. Its basic principles are to capture and record the intensity of light from a source using a photosensitive instrument. The light observed from a target over a given amount of time determines its magnitude and changes in this value with time can be used to infer stellar variability or the presence of stellar companions or planets [Deeg and Alonso, 2018].

2.2.1 Technique

Astronomical photometry in the modern era is usually carried out using Charge-Coupled Devices, or CCDs (although the popularity of complementary metal–oxide–semiconductor, or CMOS, detectors has risen in recent years) [Waltham, 2010]. CCDs work on the principle of the photo-electric effect. Photons striking individual pixels in a CCD array cause the charge on that pixel to increase which can be read out. At the end of a given time span (the exposure time) the charge on each pixel is read and recorded giving a 2D array of pixel values, proportional to the intensity of light that fell on each pixel. CCDs are so useful because the output voltage is linear with respect to the incoming flux across a wide range of intensities. The values are measured in Analogue-to-Digital units (ADUs) and then converted into fluxes or magnitudes [Howell, 2000]. Within the conversion of charge to counts a scale factor is used known as the gain. The gain is a factor used to convert the charge into a count number that can be read by a machine. This is done by choosing a multiplicative factor that converts the maximum charge per pixel to the maximum count number based on the integer format being used. For example, if a CCD can hold 350,000 electrons of charge and the count system being used is limited at 65,000 (16 bit) then a
gain value of 350,000/65,000 = 5.4 is used. This means that one count per 5.4 electrons is measured [Newberry, 1999].

**Magnitudes**

The brightness of a star is usually defined by its magnitude which is a log scale measure of its flux. The difference in magnitudes between two stars, \( m_1 \) and \( m_2 \), is a function of their fluxes, \( f_1 \) and \( f_2 \), and is given by

\[
m_1 - m_2 = -2.5 \log_{10} \frac{f_1}{f_2}.
\] (2.1)

On this scale a difference of \( n \) magnitudes translates to a factor of \( 2^{0.5n} \) in flux (e.g. a change of 5 magnitudes gives almost a factor of 100 change in flux). Additionally, this formalism means that brighter stars have lower magnitude values. The reasoning behind this scale is mainly historical thanks to Hipparchus and, later, Norman Pogson, an English astronomer [Henden and Kaitchuck, 1982].

It must also be noted that flux (and therefore magnitude) is not constant with wavelength. Some stars emit more radiation at particular wavelength thus magnitudes are measured according to a photometric pass-band used (i.e. the range of wavelengths considered). Some key pass-bands used are the UBVRI system, the JHK system (for near infrared) and the ubvyβ Strömgren system [Bessell, 2005]. Each pass-band has a specified flux value which produces a magnitude of 0 allowing for relative magnitudes in different bands to be measured. The primary calibration source is the star Vega which is defined to have a zero magnitude at all wavelengths.

**Reduction**

After observation, but before analysis, CCD photometry data generally require a number of data reduction steps in an attempt to remove instrumental effects. These often take the form of three types of calibration frames, frames that are taken in addition to the science frames which record photometric data.

The first of these are the **Bias Frames** which are zero second exposures under dark conditions used to measure the zero noise level of a CCD. This is the noise present in CCD pixels under no external photons thus must be removed from science frames. Usually multiple bias frames are taken and averaged to give a master bias frame which can then be subtracted from science images [Davenhall et al., 2001].

Second are the **Dark Frames**. These are also taken under dark conditions but have a non-zero exposure time. These frames measure the dark current which is caused by the moment of electrons that are agitated by thermal events as opposed to incident photons; it
is for this reason that most astronomical CCDs are cooled to a constant operational temperature. Ideally, dark frames are taken with the same exposure as the science frames but can be scaled to different exposure times. In a similar way to master bias frames, master dark frames are usually created from averages of multiple individual dark frames. It should also be noted that dark frames contain the signal measured in the bias frames so must themselves be bias subtracted [Davenhall et al., 2001].

Finally are the Flat Field Frames. Flat fields measure inter-pixel sensitivity changes as well as optical occlusions such as dust on the detector. Flat fields also allow the measurement and mitigation of vignetting effects which are the result of the reduction in light reaching the edges of the detector compared to the centre. Flat fields are taken by exposing the detector to a uniform light source, typically the inside of a telescope dome or the twilight sky. As before, master flat fields are created by averaging multiple flat field frames [Davenhall et al., 2001].

The effects of bias, dark and flat field frames must be accounted for and removed from science frames so as to not influence the measured pixel values. Figure 2.3 shows a visual flowchart detailing the results of these calibration steps.
Aperture photometry

Once the data has been sufficiently calibrated one can attempt to extract the relevant information. The method most commonly used for the work in this thesis is aperture photometry [Lahe et al., 2012]. Aperture photometry is one of the simplest photometric methods and consists of placing an aperture, often circular, but theoretically of any shape, over the target of interest and counting up the measured flux with that area [Aime, 2005]. This area then contains the light from the star as well as the background light that is present across the sky. To account for this another area of the image, often a circular annulus around the target, is chosen and its average pixel value is found (usually the median to reduce the effect of bright pixels in the annulus). This value is assumed as the background value for each pixel in the image and is subtracted from each pixel in the target aperture.

Though broadly simple there are subtleties in this method such as the choice of size and shape of the aperture. Stars are point light sources but, due to instrumental and atmospheric effects, are blurred out according to a Point-Spread Function (PSF) which can vary between instruments and be affected by temporal atmospheric conditions. It is important to choose an aperture that contains as much of the light from the target as possible while minimising the light from nearby sources. This is especially tricky for instruments with large pixels or for targets in crowded fields. If this is not done carefully it is quite possible to incorrectly interpret astrophysical signals from a neighbour as coming from the target star [Lupo et al., 2020].

Differential photometry

Once the aperture is chosen and the light has been measured and background-subtracted we are left with a time variable flux measurement, or light curve. The problem is, it is not always clear whether variations are astrophysical in nature, and therefore important, or instrumental/atmospheric, and thus should be removed. A way around this is to use differential photometry. This is a technique that reproduces the aperture photometry on other non-target sources known as reference stars [see, e.g., Young et al., 1991]. If we assume these stars to be similar to the target and non-variable then any variations shared by both reference and target are likely due to instrumental or atmospheric effects since the synchronisation of astrophysical variability in individual stars is highly unlikely. Using multiple reference stars decreases this chance even further and, by comparing reference stars to each other, allows us to choose the least variable stars as our reference points. Dividing the target light curve by the light curve of a non-variable reference star (or average of many reference stars) ensures that any variability we see should be real and astronomically caused by the star in question [Vigan et al., 2010]. An example of the effect is shown in Figure 2.4.
attempt should be made to choose reference stars as similar to the target as possible since some systematic effects may differ for stars of different colours or magnitudes (see section 2.2.1 below).

Atmospheric extinction

One of the key systematics for photometric measurements, and worth noting here explicitly, is that caused by atmospheric extinction. Atmospheric extinction is the dimming of starlight as it passes through the Earth’s atmosphere, caused by scattering and absorption, and will vary depending on the volume of atmosphere that the light has passed through, defined as the air mass. Air mass is a function of zenith distance, \( z \); the angular distance from the object to the zenith. As \( z \) increases the volume of atmosphere between the target and the observer increases as well causing the light to dim to a greater degree [Green, 1992]. Air mass is normalised to be 1 at the zenith and can approximated by \( \sec(z) \) [Young and Irvine, 1967]. The magnitude of a target as seen from the ground, \( m(\lambda) \) in relation to the magnitude that would be observed outside the atmosphere, \( m_0(\lambda) \) is given by the following equation;

\[
m(\lambda) = m_0(\lambda) + \kappa(\lambda)X(z)
\]

where \( X(z) \) is the air mass at zenith distance \( z \) and \( \kappa(\lambda) \) is the extinction coefficient at wavelength \( \lambda \). \( \kappa(\lambda) \) is a wavelength dependent quantity and is greater for blue light than red [Patat et al., 2011]. This is why it is important to use similar colour stars as references when conducting differential photometry. Depending on the site in question a reasonable upper limit to air mass is around 2.5, or a zenith angle of around 65°.
Seeing

Though stars are in reality extended masses, they are sufficiently far from us that they can be described as point sources. When carrying out photometry on point sources however, their light cannot be focused onto a single point, even under theoretically perfect conditions. Instead, what is seen is an Airy disk pattern. This is a bulls-eye shaped pattern with peak brightness at the centre and then alternating dark and bright annuli \[ \text{[Romanishin, 2002]} \]. The angular size of this pattern is set by the telescope diameter, $D$, and the wavelength, $\lambda$. The angular radius of the first dark annulus, $\Theta$ is given by

$$\Theta = \frac{1.22\lambda}{D}. \tag{2.3}$$

This is the theoretical limit in resolution and telescopes that reach this are called diffraction-limited. In practice however, this is only achievable for space-based telescopes since ground-based observers are affected by the larger impact of seeing. When light travels through the Earth’s atmosphere it is scrambled slightly, impacting the parallel wave-fronts and causing the image to be seen as a blob of light, centrally concentrated and radially decreasing, as opposed to the sharp diffraction pattern described by theory \[ \text{[Romanishin, 2002]} \]. In practice, seeing is a larger effect than that described by equation 2.3 and ground-based telescopes are typically seeing-limited. Seeing is exacerbated by turbulent air, and especially parcels of air with different temperatures. Thus it is preferable for an observing site to have smooth airflow and for telescopes to be cooled to the ambient temperature. At some of the best sites in the world the seeing can regularly be as good as 0.5 arcsec but lesser sites may record values of $\sim$ 2 arcsec or greater.

2.2.2 Ground-based

The vast majority of photometric observatories are ground-based. Ground-based observatories are cheaper and quicker to design and build, easier to maintain and may continue operations much longer than space-based telescopes. There are, however, a number of disadvantages to ground-based observatories. One of the key disadvantages is that many wavelengths of photometry are completely inaccessible during the day due to the intense brightness of the sun. Additionally, the rotation of the Earth means that targets are only accessible for half a day at most. These two effects can be reduced somewhat by building observatories near the poles, though this shifts the issue to a summer/winter one \[ \text{[Kenyon et al., 2006; Carlstrom et al., 2011]} \]. Another key issue is that observing from the ground requires looking through the atmosphere first which has numerous negative effects on photometry. Firstly, our atmosphere is opaque to some wavelengths. Gamma, X-ray and UV light, as well as most infrared and the longest radio waves, are blocked by our atmosphere,
limiting ground-based photometry to visible light and the shorter radio waves [Duband, 2015]. However, even those wavelengths that reach the ground are affected. Turbulent motions in the atmosphere cause low-level refraction of incident light resulting in scintillation, which causes point sources to ‘twinkle’ when observed from the ground, and seeing effects [Dravins et al., 1997]. These many effects lead to a reduction in ground-based photometric precision compared to space-based but the field is progressing constantly. Using sophisticated techniques [McGraw et al., 2010; Stefansson et al., 2017] and next-generation observatories [e.g. Wheatley et al., 2018] ground-based photometry has succeed in discovering transit signals down to only 0.13% [West et al., 2019]. Some key ground-based photometric surveys are detailed here.

**WASP**

The Wide-Angle Search for Planets [WASP, Pollacco et al., 2006] is comprised of two observatory locations. SuperWASP-North at the Roque de los Muchachos Observatory, La Palma (operating since 2004) and WASP-South at the South African Astronomical Observatory (operating since 2006). Each observatory contains 8 200 mm camera lenses capable of covering 482 deg$^2$ with each exposure at an angular scale of 13.7$''$ pixel$^{-1}$. The 8 cameras have a robotic mount capable of 30$''$ rms pointing accuracy with tracking accuracy better than 0.01$''$ s$^{-1}$. WASP is the most prolific ground-based transit survey ever, responsible for the detection of over 170 planets [e.g. Smith and WASP Consortium, 2014] many of which have been confirmed with RV follow-up. WASP detections include many firsts such as the first planet found in a retrograde orbit [WASP-17b, Anderson et al., 2010] and the first planet to have observed stratospheric water [WASP-121b, Evans et al., 2017]. An example of a recent WASP discovery paper is given in Chapter 3 with the discovery of WASP-150b and WASP-176b [Cooke et al., 2020a].

**NGTS**

The Next-Generation Transit Survey [NGTS, Wheatley et al., 2013, 2018] is an observatory comprised of 12 cameras located at ESO’s Paranal Observatory, Chile. Operations began in 2016. Each camera is 20 cm, f/2.8 and is mounted on an independent, automated mount. With a combined field of view of 96 deg$^2$ NGTS has a pixel scale of 5$''$. The telescope uses a custom 520-890 nm band-pass, increasing sensitivity to late K and early M dwarfs. Sophisticated tracking and nightly reacquisition of a target to the same pixel keeps stars from moving by more than 0.1 pixels, reducing the effect of red noise. At the time of writing NGTS is responsible for 18 confirmed exoplanet discoveries including the shallowest detection from the ground [West et al., 2019] and the most massive planet found to transit
Targeted surveys

The two ground-based photometric surveys mentioned so far have both been designed to observe hundreds or thousands of stars simultaneously in the hopes that among them there may be a transiting system. Other surveys, however, are more targeted and will photometrically observe individual stars at a time. Some of these surveys are briefly mentioned here.

MEarth [Irwin et al., 2009, 2015] is a photometric survey targeting M dwarfs to find transiting Earth-like planets. MEarth consists of 2 robotically controlled observatories each housing 8 0.4 m telescopes. Each telescope targets multiple stars in a night, switching between them. When evidence of a transit in progress is detected the telescope will remain on the current target for the duration of the transit before returning to the original observing mode. The choice of M dwarfs as the objects of interest is twofold. Firstly, M dwarfs are smaller than solar-type stars meaning transits caused by terrestrial planets are proportionally deeper and thus easier to detect. Secondly, M dwarfs are cooler, meaning even close in planets may reside in the star’s habitable zone. MEarth is responsible for a number of planetary discoveries including LHS1140b [Dittmann et al., 2017] and GJ1132b [Berta-Thompson et al., 2015].

The TRAnsiting Planets and PlanetesImals Small Telescope [TRAPPIST, Jehin et al., 2011] is a 60 cm robotic telescope located at the La Silla observatory in Chile. TRAPPIST is used for multiple purposes including follow-up of known transiting planets, the search for transits of known RV planets and a detailed solar system comet survey. On top of these science cases TRAPPIST also searches for new transiting systems around a small sample of Ultra Cool Dwarfs by observing each for several full nights spread out over a few months. The choice to focus on this subset of stars is similar to that described above for the MEarth mission. TRAPPIST’s most notable discovery is a transiting planetary system hosting 7 temperate terrestrial planets around an ultra cool dwarf [Gillon et al., 2016, 2017]. Four of these planets are believed to lie in or near the stars habitable zone [de Wit et al., 2018].

2.2.3 Space-based

The are many pros and cons to space-based observatories. They are expensive to design and build and difficult, if not impossible, to maintain [Haas et al., 2014]. Additionally, as our use of space increases, it may become difficult to access without risk of collisions during launch or operation [Kessler and Cour-Palais, 1978]. Space observatories are also more exposed
to cosmic rays without the protection of Earth’s atmosphere and the enhanced number of cosmic ray events can cause electronics to degrade more quickly [Binder et al., 1975]. The positive sides of space-based photometry however are obvious. In space, and in particular at the L2 Lagrange point, there is no day/night cycle so targets may be observed continuously with much longer baselines, vital for detecting longer period systems [Mountain et al., 2009]. Operating in space also avoids all issues created by the Earth’s atmosphere such as opacity and scintillation, as discussed above. This means space observatories can access the full range of the electromagnetic spectrum that cannot be observed from the ground as well as being fully diffraction-limited. The constant space environment additionally means that space observatories should remain stable over long timescales, assuming sufficient fuel for orbital manoeuvring. This results in well-defined and stable PSFs [Mountain et al., 2009]. Finally, space telescopes are generally designed to reach a Strehl ratio of >80%. This value is defined as the ratio between peak PSF flux and peak flux in a purely diffraction limited system [Perrin et al., 2003]. High values are easier to achieve in space due to the lack of necessary adaption to atmospheric effects when on the ground. Some of the key space-based photometric surveys are described here.

**Kepler/K2**

Kepler [Borucki et al., 1997, 2010], first conceived in 1997 and launched in 2009, is a space-based 95 cm aperture telescope. Kepler’s mission specifications were to monitor around 150,000 pre-selected main sequence stars for an observing campaign of 4 years or more. Kepler has a field of view of 115 deg² with a scale of 3.98” pixel⁻¹. The field of view is separated into 21 CCD modules (each containing two CCDs) laid out in a grid pattern. Kepler’s observations are separated into quarters with the telescope exhibiting a roll every 93 days which reorients the stars on the CCD pixels, creating light curve discontinuities (this roll is necessary to reorient the satellites solar panels towards the sun). Kepler is the most prolific transit survey ever with more than 2300 validated exoplanet detections [e.g. Morton et al., 2016]. Within this number are many firsts including the first circumbinary planet [Doyle et al., 2011], the first Earth sized exoplanets [Fressin et al., 2012], and the first system with multiple transiting planets [Ragozzine and Holman, 2019] (see Lissauer and Eisberg [2018] for a full list). The well characterised Kepler targets allowed for multiple statistical analyses and determinations of planetary occurrence rates [e.g. Fressin et al., 2013; Dressing and Charbonneau, 2013].

In 2013, after just over 4 years of observations, Kepler experienced the failure of a second reaction wheel meaning it could no longer maintain pointing stability along its third axis. However, by using solar radiation pressure to maintain pointing the mission was re-purposed to observe fields along the ecliptic equator and was re-branded as K2 [Howell
Figure 2.5: Physical layout and sky coverage of the TESS satellite. Taken from Figure 7 of Ricker et al. [2015].

et al., 2014], resuming observations in 2014. In this new mode of operations K2 observes fields for 80 days before moving on. This gives a greater sky coverage than Kepler but with much shorter observing baselines. K2 is responsible for 370 confirmed exoplanet detections including the first habitable zone exoplanet with atmospheric water vapour [Montet et al., 2015; Benneke et al., 2019]. As a secondary bonus K2’s position along the ecliptic means it is well placed to observe small solar system bodies as well [Dotson et al., 2018].

TESS

The Transiting Exoplanet Survey Satellite [TESS, Ricker et al., 2015] is a space-based transit survey launched in 2018. TESS is an all-sky survey with the primary mission of observing nearly the entire sky (∼85%) over a two year time span. TESS has four 10.5 cm aperture cameras, each 24×24 deg giving a total instantaneous field of view of 96×24 deg. During the primary mission TESS is aligned along the ecliptic plane with its camera oriented vertically with camera four located on the ecliptic pole. Every 27 days TESS rotates, moving on to new areas of the sky. Depending on the distance from the ecliptic equator the new position will overlap creating regions with much longer baselines, up to 1 year at the ecliptic poles. TESS’s major trade off for its vast sky coverage is its pixel scale with each pixel being 21” per side, this makes observations of crowded fields challenging. Figure 2.5 shows the arrangement of the cameras and the primary mission sky coverage.

TESS observes around 200,000 target stars at a cadence of 2 minutes (Postage Stamps, PS) and the rest of the sky at 30 minute cadence (Full Frame Images, FFIs) [Sullivan et al., 2015]. Each 27 day sector is separated into 2 TESS orbits and interrupted briefly to download the data. TESS has now progressed passed its nominal mission lifetime into an extended mission where it will repeat its full sky observations at an improved FFI cadence of 10 minutes before covering the ecliptic equator in year 5. At the time of writing TESS
has produced 122 confirmed exoplanets with many more candidates. Alongside transit photometry TESS has been used for many other science areas including solar system science [Pál et al., 2018], asteroseismology [Campante et al., 2016] and flare detection Günther et al. [2020].

CHEOPS

The CHaracterising ExOPlanet Satellite [CHEOPS, Fortier et al., 2014; Rando et al., 2019] was launched in late 2019. CHEOPS is an ultra-high precision photometry mission designed to follow-up known transiting systems discovered from other surveys in an attempt to provide more accurate radii. CHEOPS has a single f/5 telescope with effective aperture of 300 mm and a field of view of 0.32 deg². CHEOPS has a pixel scale of ∼1”. Since CHEOPS targets will be distributed over the whole sky its orbital design ensures that at least 50% of the sky is accessible for a cumulative 50 days per year [Benz et al., 2020]. CHEOPS is designed for a 3.5 year primary mission with an extended mission lifetime up to 5 years. CHEOPS has already completed its first transit observations [Lendl et al., 2020a] and simulations suggest it will make improvements to large numbers of planetary radii as part of its primary mission goals [Futyan et al., 2020] as well as having the potential to contribute to secondary science [Barceló Forteza et al., 2020].

2.3 Spectroscopy

Spectroscopy is one of the most important tools of the observational astronomer. Spectroscopy is based on the fact that astronomical bodies do not emit light at a single wavelength but in fact emit across a whole spectrum. Using a prism or similar tools observers can separate the emitted light as a function of wavelength and reveal the absorption or emission features within. Light from a particular source, for example a star, is emitted in a continuous spectrum but specific elements in the stellar composition can absorb specific wavelengths meaning they are not detected by observers. These absorption lines allow observers to determine the composition of astronomical bodies and, due to the Doppler effect, also allow for the measurement of line-of-sight velocities of the objects.

A key aspect of spectroscopic observing, especially when compared to photometry, is that it can be used in less favourable observing conditions. That is, a night can be of sufficient observing quality for spectroscopy but not photometry, usually due to the presence of clouds which are better at blocking visible light than some other wavelengths. Observatories often list higher atmospheric observing constraints for photometry than spectroscopy (see the European Southern Observatory for an example⁴).

⁴https://www.eso.org/sci/observing/phase2/ObsConditions.html
2.3.1 Technique

Astronomical spectrographs can differ in a range of specifics but at the basic level, all will contain the same elements. A slit, or fibre, on to which the light from a telescope is focused, a collimator, which transforms the incoming light into a parallel wavefront, a disperser, usually a grating or prism, to separate the light by wavelength and finally, a camera, which focuses the light on to a detector, usually a CCD [Massey and Hanson, 2013]. A simple schematic of this set up is given here in Figure 2.6.

Design

The disperser most commonly used in spectrographs is a grating which consists of a high number of grooves per mm (usually 100-1,000). The wavelength seen from a grating is dependent of the angle of the grating to the incoming light, $i$, and the angle from the grating to the observer, $\theta$. This is governed by the grating equation,

$$ml = \sigma (\sin i + \sin \theta),$$  \hspace{1cm} (2.4)

where $m$ is the order in which the grating is used and $\sigma$ is the separation between grating grooves [Massey and Hanson, 2013; Wise et al., 2005]. The dispersion of the light with respect to wavelength is then the differential of this equation, which, at a given value of $i$, is given by

$$\frac{d\theta}{d\lambda} = \frac{m}{\sigma \cos \theta}.$$  \hspace{1cm} (2.5)

At small $\theta$ this simplifies to $d\theta/d\lambda \sim m/\sigma$ meaning that, to increase dispersion we must either increase $m$ or decrease $\sigma$. The grating order $m$ must usually be kept low to avoid overlap from multiple orders at once except in the case of echelle spectrographs which operate at $m \gtrsim 5$ thanks to the extra step of using cross-dispersal after the grating.
which separates out each order across the detector [Queloz, 1995; Piskunov and Valenti, 2002].

Spectral resolution, or spectral resolving power, is the ability of the spectrograph to separate between similar wavelengths separated by $\Delta \lambda$. This is generally quoted as either $\Delta \lambda$ (measured in nm or Å) or by the dimensionless quantity $R = \lambda / \Delta \lambda$. Spectrographs with $R < 1,000$ are low resolution and cannot measure specific spectral lines whereas instruments with $1,000 < R < 10,000$ are intermediate level and can measure broad spectral lines. Spectrographs with $R > 10,000$ are high resolution, capable of measuring even narrow spectral lines in detail.

Data reduction

The data reduction steps involved in spectroscopic data include many of the steps previously discussed for photometry in section 2.2.1. Bias, dark and flat field frames are all necessary for the reduction of spectroscopic data since they correct for effects inherent to the use of CCDs, which are often used in spectroscopic data collection [Clayton and Holloway, 1998]. There are additional steps needed specific to spectroscopy however. Spectrographs also require what is known as comparison arcs. These are short exposures of discharge tubes with known spectra, used to give a wavelength scale to the measured stellar spectra [Massey and Hanson, 2013]. Stellar and reference spectra are measured simultaneously ensuring the best comparison between relative spectral positions. The discharge lamps chosen depend on the dispersal level at which one is working. HeNeAr lamps are sparse in lines and useful at low and moderate dispersal ranges whereas ThAr are more rich in lines and useful at high dispersal ranges. From the reference spectra one can measure a wavelength to pixel conversion which can then be reversed to determine wavelength values for the science spectra [Baranne et al., 1996].

At each value of wavelength the sum of the pixels at that wavelength are added together to find the flux as a function of wavelength. In theory this is simple but in practice, not all pixels at a given wavelength position have the same flux as the signal is spread out perpendicular to the wavelength direction. This can be accounted for by taking a weighted sum, increasing the weighting towards the centre of the image and decreasing towards the wings [Clayton and Holloway, 1998]. The final step is to normalise the measured spectrum since the usual goal of spectroscopy is analysis of the absorption/emission features, not the underlying continuum. This is usually done by fitting a low order polynomial to the data and dividing out the continuum. In theory this leaves just the spectral features. In practice however, deep or wide features must first be masked else they will affect the continuum fit and thus spoil the normalisation [Clayton and Holloway, 1998].
Spectrograph drift

A specific source of spectrographic noise that must be accounted for is spectrograph drift. For a perfectly ideal spectrograph the light from a target falls on the same pixels at every point in time. In practise however, the exact pixels that the light falls out can drift over a night. This is generally due to instrumental effects and changes in detector performance [Min et al., 2014]. The cause of these changes are typically mechanical in nature, such as the relative movement of various parts of the instruments or changes in the temperature and pressure environment of the instrument. Because of the way pixels are converted to wavelength, as discussed above, changes in pixel position can lead to changes in wavelength and thus measured radial velocities. These drifts are typically less than a pixel over a night but can translate to a few tens of centimetres per second in measured RVs [Ma and Ge, 2019].

In the first instance these drift effects are best removed entirely and this is attempted using strict environmental control [Robertson et al., 2019]. The most precise spectrographs have environmental control systems in place to maintain constant temperature and pressure. Different instruments control this in different ways, some choose to house the entire instrument in a vacuum whereas some attempt to maintain constant atmospheric pressures. The materials chosen for the construction of the instrument are also carefully selected to be durable, and resist degradation or transformation over time and under other physical changes. The choices are made by balancing the required spectrograph precision with other concerns such as ease of access (for potential maintenance or upgrades) or expense [Siverd et al., 2017].

If perfect environmental control is not achieved (and in theory small changes will always remain) a way of removing these effects in the data must be found. This is usually approached in one of two ways. The first of these is to simultaneously measure a reference spectrum with the target spectrum. Since this reference spectrum is well understood drifts in the spectrograph can be measured and then, under the assumption that reference and target spectra are observed under the same conditions, these drifts can be subtracted from the target observations. This should then leave a clean target spectrum, unaffected by drifts.

The second method is to take environmental readings of the spectrograph, either at the start and end of a night of observing, or continuously during a night [Ma and Ge, 2019]. The observed spectra can then be correlated against these environmental recordings and the drifts modelled and removed. Figure 2.7 shows an example of this method used on the Keck II NIRSPEC spectrometer [McLean et al., 1998] where the drift is calibrated against temperature.
Figure 2.7: Spectrograph drift removal on the Keck II NIRSPEC spectrometer. The left panel shows the instrument drift across 1.2 hours in orange and the drift residuals after reduction using the temperature. The temperature chosen is the average of multiple temperature reading shown in the right panel (as the black line), each standardised to the individual sensors. Taken from Figure 6 of Yi et al. [2016].

Absorption and emission lines

The result of this data collection and reduction is a plot of flux as a function of wavelength. An example of part of the Solar spectrum is shown in Figure 2.8.

The absorption lines seen in this figure are the result of the presence of these elements in the Solar atmosphere. This is a direct result of the quantised nature of atomic energy levels. Electrons in atoms lie in distinct energy levels and can only move to other, equally well defined, levels. When a photon is incident on an electron it may be absorbed by said electron only if the energy of the photon is equal to the energy required to transition to a higher energy level. Too high or low and the absorption will not happen. Therefore, photons at particular energies are absorbed and others are not. Converting energy to wavelength \( E = \frac{hc}{\lambda} \) means that the presence of a particular atom can result in the reduction of emitted photons at a particular wavelength and therefore the production of absorption lines in the spectrum. Different atoms produce absorption features at different wavelengths.

Spectral emission happens in much the same way. An electron in a non-ground state of an atom may fall back down to an unoccupied energy level by spontaneous emission of a photon of energy equal to the difference between the two energy levels. Therefore atoms can only emit electrons of discrete energies (or wavelengths). The presence of peaks in flux at particular wavelengths can be the key to determining the existence of specific atoms that are emitting photons of a discrete energy.
Figure 2.8: Part of a reduced and normalised Solar spectrum. Individual absorption features can be seen and are captioned. Taken from http://spiff.rit.edu/classes/phys230/lectures/spec_interp/spec_interp.html.
Cross-correlation

The presence of the spectral features, and the precise knowledge of what wavelengths they occur at (in the laboratory rest frame), allows us to carry out cross-correlation and measure the velocity of the astronomical source as discussed in section 1.2.2. Generally this is carried out using a stellar template spectrum but can also be done using a time averaged spectrum of the target in question. In its simplest form cross-correlation is a measure of the overlap, or correlation, between the template and observation [Baranne et al., 1996; Pepe et al., 2002b]. The template is held still while the observation is shifted in wavelength and at each step a measure of the difference between the two spectra is made, usually some form of chi-squared value. As the observed spectrum is shifted in wavelength the measured value goes up and down giving a plot as a function of wavelength shift. The point at which the cross-correlation function is minimised gives the wavelength shift between the observed spectrum and the template. The wavelength offset can then be converted into a line-of-sight velocity using equation 1.8. Care must be taken to account for velocities added by the motion of the Earth around the Sun but this technique offers a method of determining the relative line-of-sight velocity of astronomical bodies [Malavolta et al., 2017].

2.3.2 Ground-based

As with photometry, spectroscopy can be carried out both on the ground and in space and many of the same advantages and disadvantages affect spectrographs. Ground-based instruments are easier and cheaper to build, maintain and upgrade compared to space-based instruments. Spectrographs on the ground can also be transferred between different telescopes; for example ELODIE was replaced by SOPHIE on the 1.93 m reflector telescope at Haute-Provence Observatory, France without necessitating a new telescope be built [Perruchot et al., 2008]. Many of the same adverse effects are still applicable though. Since spectroscopy is so dependent on wavelength the fact that our atmosphere is opaque to large regions of the electromagnetic spectrum completely removes some regions of wavelength space from consideration from the ground [Duband, 2015]. Additionally to this, even the regions that are not fully opaque are impacted. Since our atmosphere contains a large range of elements and molecules it induces multiple absorption and emission signatures into the light coming from distant objects [Catanzaro, 1997]. These telluric signals must be removed to properly characterise the astronomical spectra. This can be done by comparing the observations with observations of standard spectral stars which can be used to identify the telluric additions. Another mitigating strategy is to use the Doppler shift of the target light. The spectral features in the Earth’s atmosphere should have a different Doppler shift to those from the target and this can be identified leading to their removal [Rudolf et al.,
2016]. Only two spectrographs, both ground-based, are used repeatedly in this thesis, and are discussed below.

**CORALIE**

CORALIE [Queloz et al., 2000b] is a fibre-fed echelle spectrograph installed on the Euler telescope 1.2 m telescope in La Silla, Chile. Achieving first light in 1998 CORALIE is an improved version of the ELODIE spectrograph [Baranne et al., 1996], the instrument responsible for the detection of the first planet around a sun-like star [Mayor and Queloz, 1995]. CORALIE covers the wavelength range from 380 to 690 nm and was installed with a resolution of 50,000 which has since been upgraded to 60,000 [Pepe et al., 2018]. CORALIE is contained in an isolated and temperature controlled room, fed by two fibres allowing for simultaneous Thorium-Argon referencing. CORALIE has a 2" FOV and uses a 2k x 2k CCD with 15 µm pixels which allow it to reach a higher resolution that the original ELODIE instrument [Queloz et al., 2000a]. Multiple upgrades over its lifetime including a prism replacement in 2007 and the change to octagonal fibres in 2014 have allowed the instrument to reach an improved precision of typically 3 ms$^{-1}$ [Pepe et al., 2018]. CORALIE’s primary mission is to search for planets using the radial velocity method and is responsible for 111 detections over the last 2 decades. CORALIE has contributed to planets across a range of periods and masses [Mayor et al., 2004; Ségransan et al., 2010], allowed for statistical population analysis [Santos et al., 2004b] and assisted in follow-up of photometrically discovered systems [e.g. Lendl et al., 2020b].

**HARPS**

The High Accuracy Radial Velocity Planet Searcher [HARPS, Pepe et al., 2002b] is a spectrographic instrument designed with the goal of reaching sub-ms$^{-1}$ precision radial velocity measurements. HARPS is a fibre-fed cross-dispersed echelle spectrograph mounted on the 3.6 m ESO telescope in La Silla and saw first light in 2003. HARPS makes a number of improvements over CORALIE and similar instruments in its attempt to reach such high precision. HARPS covers the same wavelength range as CORALIE and uses the same simultaneous Thorium-Argon referencing as CORALIE but has a smaller FOV, only 1", allowing it to reach a resolution of 115,000 [Mayor et al., 2003]. HARPS has also made a number of mechanical improvements to increase long-term stability and thus precision of the instrument. The entire HARPS instrument is housed in a vacuum to remove effects of pressure changes during observations. Additionally the surroundings are temperature controlled to better that 0.1K [Queloz et al., 2001b]. Finally, the optical design is simplified as much as possible and contains no moving parts with the number of degrees of freedom kept
to an absolute minimum. In fact, HARPS operated from 2003 to 2015 without ever being opened [Pepe et al., 2018]. In practice, the pioneering design of the HARPS instrument has led to a precision of 0.5-1 ms$^{-1}$. The success of the instrument led to a twin facility, HARPS-North, installed on the 3.58 m TNG, La Palma which saw first light in 2012 [Latham and HARPS-N Collaboration, 2013]. Combined, these two HARPS instruments are responsible for more than 200 confirmed planet discoveries. Amongst these discoveries are several worthy of note including one of the first Super-Earths [Santos et al., 2004a], many multi-planet systems [e.g. Udry et al., 2007; Pepe et al., 2011] and a planet around our nearest neighbour, Proxima Centauri b [Anglada-Escudé et al., 2016]. As well as specific systems the HARPS instruments have contributed to many population studies [e.g. Mayor et al., 2011; Adibekyan et al., 2012], as well as additional science goals such as atmospheric transmission spectroscopy [Wyttenbach et al., 2015] and the measurement of reflected light from an exoplanet [Martins et al., 2015]. HARPS-N specifically has also been used to extensively study the radial velocity signals of our Sun [Dumusque et al., 2020].

2.3.3 Space-based

Though data from space-based spectrographs are not used in this thesis some details are mentioned briefly here. As discussed above, space-based instruments are more expensive to build and launch and much harder to upgrade or maintain (though not impossible e.g. Nurre et al. [1995]). However there are numerous advantages as well, especially for spectrographs. Firstly, and most importantly, spectrographs in space can access regions of the electromagnetic spectrum impossible to observe from Earth. This allows for UV and X-ray spectroscopy which cannot be done from the ground. Additionally, the vacuum of space means the spectrographs can be very stable on long time periods. This is ideal to reduce instrumental effects in the spectral data and much work is needed on the ground to replicate this stability (see section 2.3.1). Some noteworthy space-based spectrograph instruments are the Space Telescope Imaging Spectrograph (STIS) on the Hubble Space Telescope [Gregg and Minniti, 1997], the Chandra and XMM-Newton X-ray observatories [Paerels and Kahn, 2003] and the Hershel space observatory [Pilbratt et al., 2010]. In the future the James Webb Space Telescope (JWST, Gardner et al., 2006) will be able to carry out space-based spectroscopy using a whole suite of spectrographs covering a range of wavelength regions. Finally, due to the vastly oversubscribed nature of space spectrographs, there is increasing interest in the idea of low-cost, small scale space-based spectroscopy instruments using specially designed CubeSats [Raskin et al., 2018].
Chapter 3

WASP-150b & WASP-176b

Note

The following chapter is a reproduction of the paper ‘Two Transiting Hot Jupiters from the WASP Survey: WASP-150b and WASP-176b’ [Cooke et al., 2020a] published in The Astronomical Journal in June 2020. The majority of the paper including introduction, analysis, discussion and conclusions were carried out and written up by me. This paper included external contributions in the form of observational data and descriptions of instruments. Specifically, Sections 3.3.1, 3.3.2, 3.4.1, 3.4.2 and 3.5 contain details written by co-authors. Additionally, co-author contributions consisted of discussions and improvements regarding methodology and results as well as general proof reading.

This paper lays out the discovery, confirmation and characterisation of two hot Jupiter exoplanets from the WASP survey, WASP-150b and WASP-176b.

Abstract

We report the discovery of two transiting exoplanets from the WASP survey, WASP-150b and WASP-176b.

WASP-150b is an eccentric ($e = 0.38$) hot Jupiter on a 5.6 day orbit around a $V = 12.03$, F8 main-sequence host. The host star has a mass and radius of $1.4 M_\odot$ and $1.7 R_\odot$ respectively. WASP-150b has a mass and radius of $8.5 M_J$ and $1.1 R_J$, leading to a large planetary bulk density of $6.4 \rho_J$. WASP-150b is found to be $\sim 3$ Gyr old, well below its circularisation timescale, supporting the eccentric nature of the planet.

WASP-176b is a hot Jupiter planet on a 3.9 day orbit around a $V = 12.01$, F9 sub-giant host. The host star has a mass and radius of $1.3 M_\odot$ and $1.9 R_\odot$. WASP-176b has a mass and radius of $0.86 M_J$ and $1.5 R_J$ respectively, leading to a planetary bulk density of $0.23 \rho_J$. 
3.1 Introduction

As of October 2019, over 4000 exoplanets have been verified\textsuperscript{1}. Of these planets, over 3000 have been discovered using the transit method. These results have shown that exoplanet populations are both very common and very diverse [Batalha, 2014], with a wide range of system parameters found thus far. The transit discoveries have been made using both space-based surveys, for example Kepler/K2 [Borucki et al., 2010; Howell et al., 2014], and ground-based surveys, including WASP [Pollacco et al., 2006], HATNet/HATSouth [Bakos, 2018], KELT [Pepper et al., 2007], TRAPPIST [Jehin et al., 2011] and NGTS [Wheatley et al., 2018]. This field is expected to be expanded upon even further in the coming years with the yield of TESS [Ricker et al., 2014] and other future planned missions such as JWST [Gardner et al., 2006], PLATO [Rauer et al., 2016] and ARIEL [Pascale et al., 2018].

Within this sample of exoplanets lies the region corresponding to hot Jupiters. A type of planet noticeably absent from our own solar system, hot Jupiters are planets with masses comparable to Jupiter but with orbital periods on the order of days. Current estimates put the occurrence rate of giant planets within 5 - 10 au around FGK stars at 10 - 20% [Cumming et al., 2008; Mayor et al., 2011]. Further detections and precise characterisations of hot Jupiter exoplanets will improve our ability to study this population using statistical methods, and hence allow for a better understanding of this unique type of planet.

Exoplanets are broadly characterised using a combination of the transit method [Henry et al., 2000; Charbonneau et al., 2000; Winn, 2010a] and radial velocity measurements. When combined these methods allow for the measurement of radius, mass and density which can lead to inferences of composition.

This paper discusses the discovery and characterisation of the transiting hot Jupiters WASP-150b and WASP-176b. Sections 3.2, 3.3 and 3.4 detail the WASP discovery, spectroscopic follow-up and photometric follow-up, respectively. Section 3.5 discusses the high-spatial-resolution follow-up of WASP-150. Section 3.6 outlines the analysis and derived parameters. Finally, section 3.7 summarises the discovery findings.

3.2 SuperWASP discovery photometry

The WASP project (now decommissioned) was split into north and south facilities with telescopes located at the Isaac Newton Group (ING) at the Observatorio del Roque de los Muchachos, La Palma, Spain and at the South African Astronomical Observatory (SAAO), Sutherland, RSA respectively. Both facilities consisted of 8 Canon 200 mm f/1.8 lenses, each linked to an Andor e2v 2048 × 2048 pixel CCD. Each camera had a total field of view

\textsuperscript{1}https://exoplanetarchive.ipac.caltech.edu/
Figure 3.1: (a) Upper panel: Phase folded WASP data for WASP-150b binned to 10 minutes. Data are shown in orange with the best fit model derived from MCMC analysis shown in red. Lower panel: Residuals from MCMC fit.
(b) BLS periodogram of WASP photometry. The period from our MCMC analysis is highlighted.

of $7.8^\circ \times 7.8^\circ$ with a pixel scale of $13.7''$ [Pollacco et al., 2006].

For WASP-150, a total of 99,892 photometric data points were taken between 14 May 2004 and 4 August 2011. These data ranged across 38 transits. WASP-150b was flagged as a high priority candidate on 17 February 2014 and confirmed as a planet on 25 June 2015. Figure 3.1 shows the WASP discovery curve folded by the best fit period and binned to 10 minutes. Additionally, we show the best fit transit model from our MCMC.

For WASP-176, a total of 23,082 photometric data points were taken between 26 May 2004 and 6 October 2010. These data ranged across 30 transits. WASP-176b was flagged as a high priority candidate on 12 February 2014 and confirmed as a planet on 23 January 2018. Figure 3.2 shows the WASP discovery curve folded by the best fit period and binned to 10 minutes. Additionally, we show the best fit transit model from our MCMC.

The SuperWASP data were reduced using the standard SuperWASP pipeline as described in Pollacco et al. [2006]. Analysis of the light curve was then carried out using the
Figure 3.2: (a) *Upper panel*: Phase folded WASP data for WASP-176b binned to 10 minutes. Data are shown in orange with the best fit model derived from MCMC analysis shown in red. *Lower panel*: Residuals from MCMC fit. (b) BLS periodogram of WASP photometry. The period from our MCMC analysis is highlighted.
Table 3.1: Photometric properties

<table>
<thead>
<tr>
<th>Parameter</th>
<th>WASP-150</th>
<th>WASP-176</th>
</tr>
</thead>
<tbody>
<tr>
<td>R.A.</td>
<td>17:37:03.14</td>
<td>20:54:44.94</td>
</tr>
<tr>
<td>Dec.</td>
<td>+53:01:16.4</td>
<td>+09:10:44.5</td>
</tr>
<tr>
<td>$V$</td>
<td>12.03</td>
<td>12.01</td>
</tr>
<tr>
<td>$J$</td>
<td>11.06</td>
<td>10.99</td>
</tr>
<tr>
<td>$G$</td>
<td>11.92</td>
<td>11.94</td>
</tr>
</tbody>
</table>

box least-squares (BLS) fit method, as in Kovács et al. [2002], and the SysRem detrending algorithm, described in Tamuz et al. [2005]. The results of BLS searches on the detrended data are shown in Figures 3.1 and 3.2. The data were searched with a transit-search algorithm [Collier Cameron et al., 2007b] and flagged as belonging to a planetary candidate. System parameters were then estimated from catalogue data and a Monte Carlo simulation [Collier Cameron et al., 2006]. These initial estimates produced a period of 5.644 days, a depth of 3.2 mmag and a width of 2.4106 hours for WASP-150b and a period of 3.899 days, a depth of 3.4 mmag and a width of 4.5292 hours for WASP-176b. Additional follow-up spectroscopy and photometry were then obtained to confirm and characterise the planets.

3.3 Spectroscopic follow-up

3.3.1 WASP-150b

WASP-150 was observed with the SOPHIE spectrograph, first to establish the planetary nature of the transiting candidate, then to characterize the secured planet by measuring in particular its mass and orbital eccentricity. SOPHIE is dedicated to high-precision RV measurements at the 1.93 m telescope of the Haute-Provence Observatory [Perruchot et al., 2008; Bouchy et al., 2009] and is widely used for WASP follow-up [e.g. Collier Cameron et al., 2007a; Hébrard et al., 2013; Schanche et al., 2019]. We used its High-Efficiency mode with a resolving power $R = 40,000$ and slow readout mode. We obtained 22 observations between May 2014 and April 2015. Depending on weather conditions, exposure times ranged between 8 and 33 minutes in order to maintain a signal-to-noise ratio as constant as possible among observations.

The spectra were extracted using the SOPHIE pipeline [Bouchy et al., 2009], and the radial velocities were measured through weighted cross-correlation with a numerical mask [Baranne et al., 1996; Pepe et al., 2002a]. They were corrected for the CCD charge transfer inefficiency [Bouchy et al., 2009], and their error bars were computed from the cross-correlation function (CCF) using the method presented by Boisse et al. [2010]. The monitoring of constant stars revealed no significant instrumental drifts during the epochs of
Figure 3.3: Upper panel: Phase folded SOPHIE radial velocity observations of WASP-150. Data are shown in blue with the best fit model derived from MCMC analysis shown in red. Lower panel: Residuals from RV fit.

observation, and none of the spectra were significantly affected by any moonlight or other sky background pollution.

The resulting CCFs have full width at half maximum (FWHM) of 14.1 ± 0.1 km/s, and contrast that represents ~14% of the continuum. The lines are slightly broader than what is usually measured in High-Efficiency mode due to the stellar rotation of WASP-150 (we measured a projected rotational velocity $v \sin i = 8.3 \pm 1.0$ km/s from the parameters of the CCF using the calibration of Boisse et al. [2010]).

The RVs have uncertainties around 19 m/s. They show large variations in phase with the SuperWASP transit ephemeris for an eccentric orbit, and a semi-amplitudes of the order of 800 m/s. This would correspond to a companion in the massive-planet regime. The SOPHIE RVs are shown in Figure 3.3 with the best fit model and residuals from our MCMC analysis (see section 3.6.1). Data files can be found at https://github.com/BenCooke95/W150-W176.

Radial velocities measured using different stellar masks (F0, G2, K0, or K5) produce variations with similar amplitudes, so it is unlikely that these variations are produced by blend scenarios composed of stars of different spectral types. Similarly, the measured CCF bisector spans quantify possible shape variations of the spectral lines. They show a low dispersion of 27 m/s, which agrees with their expected accuracy and is tiny by compar-
Figure 3.4: SOPHIE RV bisector span as a function of radial velocity for WASP-150. The solid black line shows the best weighted linear fit to the data. The lack of any significant gradient supports the assumption that the RV signal is produced by a planetary companion. The aspect ratio of the two axes is unity.

ison to the large RV variations. We can thus conclude that the RV variations are not due to spectral-line profile changes attributable to blends or stellar activity, but rather to Doppler shifts due to a massive, eccentric planetary companion. Figure 3.4 shows the bisector spans.

3.3.2 WASP-176b

High resolution spectroscopy for WASP-176 was performed with the CORALIE spectrograph [Queloz et al., 2000b; Pepe et al., 2018] on the Swiss 1.2 m telescope at La Silla Observatory (Chile). In total we obtained 26 measurements between July 2014 and July 2018. RVs were computed with the standard CORALIE data reduction pipeline by cross-correlating the spectra with a binary G2 mask [Pepe et al., 2002a].

The reduced CORALIE data are displayed in Figure 3.5 along with the best fit model and residuals from our MCMC analysis (see section 3.6.2). Data files can be found at https://github.com/BenCooke95/W150-W176.

Additionally, the line bisector was analysed to ensure the observed signal was indeed from an orbiting body and not a blended binary [Queloz et al., 2001a]. Figure 3.6 shows the results of this analysis using the bisector velocity span as a function of RV. No significant correlation is seen (evidenced by the best fit linear slope). Thus, this analy-
sis supports the detection of a planetary companion to WASP-176. Using the Grubbs test for outliers [Grubbs, 1950] we found exactly one outlier at 95%. This point was removed before calculating the line bisector correlation.

### 3.4 Photometric follow-up

#### 3.4.1 WASP-150b

A number of telescopes were used to gather the necessary follow-up photometry for WASP-150b. These are summarised in Table 3.2 along with observation dates, photometric filters and transit notes. The following sections detail the follow-up. The lightcurves are shown in Figures 3.7 and 3.8. Data files can be found at [https://github.com/BenCooke95/W150-W176](https://github.com/BenCooke95/W150-W176).

**IAC80**

A full transit of WASP-150b was observed on 7 July 2015 using the CAMELOT\(^2\) (CAmara MEjorada Ligera del Observatorio del Teide) camera installed on the IAC80\(^3\) telescope.

---

\(^2\)[http://vivaldi.11.iac.es/OOCC/iac-managed-telescopes/iac80/camelot/]

\(^3\)[http://www.iac.es/OOCC/instrumentation/iac80/]
Figure 3.6: CORALIE RV bisector span as a function of radial velocity for WASP-176. The solid black line shows the best weighted linear fit to the data. The lack of any significant gradient supports the assumption that the RV signal is produced by a planetary companion. The aspect ratio of the two axes is unity.

Table 3.2: Photometric follow-up of WASP-150

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Date (night of)</th>
<th>Filter</th>
<th>No. points</th>
<th>Average cadence (s)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>NITES</td>
<td>11/09/2014</td>
<td>None</td>
<td>599</td>
<td>26</td>
<td>Noisy egress</td>
</tr>
<tr>
<td>RISE</td>
<td>23/05/2015</td>
<td>V + R</td>
<td>14640</td>
<td>1</td>
<td>No pre-transit OOT</td>
</tr>
<tr>
<td>IAC80</td>
<td>07/07/2015</td>
<td>Johnson-Bessel B</td>
<td>230</td>
<td>75</td>
<td>Full transit observed</td>
</tr>
<tr>
<td>TCS</td>
<td>24/07/2015</td>
<td>Johnson-Cousins R</td>
<td>13500</td>
<td>1</td>
<td>Full transit observed</td>
</tr>
<tr>
<td>RISE</td>
<td>24/07/2015</td>
<td>V + R</td>
<td>12141</td>
<td>1</td>
<td>Missed egress</td>
</tr>
<tr>
<td>Cassini</td>
<td>27/08/2015</td>
<td>Johnson I</td>
<td>482</td>
<td>47</td>
<td>Full transit observed</td>
</tr>
<tr>
<td>CAHA</td>
<td>27/08/2015</td>
<td>Cousins I</td>
<td>421</td>
<td>46</td>
<td>Full transit observed</td>
</tr>
</tbody>
</table>
at the Observatorio del Teide, Tenerife, Spain. CAMELOT contains a $2048 \times 2048$ back-illuminated e2v CCD, providing a field-of-view of $10.4'\times10.4'$, with a pixel scale of 0.304''.

Data were reduced using standard routines of IRAF [Tody, 1986, 1993]. All images were bias and flat subtracted and differential photometry was carried out using IDL DAOPHOT-Type Photometry Procedures$^3$ [Stetson, 1987]. Among the several stars appearing on the FOV of the camera, those showing less dispersion were selected to produce an average reference star and obtain the final light curve. We used a fixed aperture radius of 13 pixels, which minimised the RMS scatter in the out of transit data.

**CAHA 1.23 m**

The transit on 27 August 2015 was observed with the DLR-MKIII camera fed by the Zeiss 1.23 m CAHA telescope. During the observations the sky was clear besides a couple of intervals when some thin clouds passed in front of the target. The observations were carried out using the defocusing technique, which allowed the use of longer exposures compared to in focus observations, without the risk of saturation [Southworth et al., 2009].

The data reduction was performed using standard methods making use of the defot pipeline [Southworth et al., 2009, 2014]. In brief, each scientific image was calibrated using a master-bias and a master-flat. The fluxes detected from the target and comparison stars were then obtained via aperture photometry, selecting the aperture sizes that minimised the scatter of the light curves. A relative-flux light curve was then obtained for the target star relative to an optimally-weighted composite comparison star constructed from the light curves of comparison stars present in the same field of view, to account for atmospheric and instrumental changes occurring during the observations.

**Cassini 1.52 m**

On 27 August 2015 WASP-150b was simultaneously observed with the Cassini 1.52 m telescope at the Astronomical Observatory of Bologna in Loiano (Italy), thus performing the two-site observational strategy [Ciceri et al., 2013]. The Cassini has a focal ratio of f/8, a focal length of 12 m, and is equipped with a back-illuminated CCD with $1300 \times 1340$ pixels and a pixel size of 20$\mu$m. A focal reducer makes the telescope an f/5, so that its plate scale is 0.58'' pixel$^{-1}$ and the FOV is $13' \times 12.6'$. The autoguided observations were performed with the defocusing technique (to improve the photometric precision), though the level of defocusing was limited in order to avoid blending from a fainter star a few arcsec away. The data were reduced using the defot code, as described in the preceding section.

$^3$https://idlastro.gsfc.nasa.gov/contents.html

58
RISE

Two partial transits of WASP-150b were observed on 23 May 2015 and 24 July 2015 with the RISE\(^5\) (Rapid Imaging Search for Exoplanets) optical camera, installed on the 2-meter robotic Liverpool Telescope [Steele et al., 2008; Gibson et al., 2008] at the Observatorio del Roque de los Muchachos, La Palma, Spain. The camera is a frame transfer e2v CCD of 1024 × 1024 pixels, which has a FOV of 9.2′ × 9.2′.

Bias and flat reduced images were provided by the telescope pipeline. The light curves were extracted following the standard procedures described in section 3.4.1, using a fixed aperture radii of 5.5 and 7 pixels, for the first and second night respectively.

TCS

On 24 July 2015, WASP-150b was simultaneously observed with the Telescopio Carlos Sánchez\(^6\) (TCS) a 1.52-metre telescope installed at the Observatorio del Teide, once again performing the two-site observational strategy [Ciceri et al., 2013]. We used the Wide FastCam camera, a 1024 × 1024 pixels EMCCD detector, coupled to an optical design [Murga et al., 2014] that is able to provide ~ 8′ × 8′ FOV. This camera provides small readout times and low electronic noise, which allow us to precisely measure exoplanet transit timings.

All images were bias and flat subtracted, and light curves were extracted using similar procedures to those described in section 3.4.1. In this case, a fixed radius aperture of 11.5 pixels was selected.

NITES

A transit of WASP-150b was obtained on 11 September 2014 using the Near Infrared Transiting ExoplanetS telescope [NITES, McCormac et al., 2014] on La Palma. The data were reduced in Python using ccdproc [Craig et al., 2015]. A master bias, dark and flat were created using the standard process. Non-variable nearby comparison stars were selected by hand, and aperture photometry extracted using sep [Barbary, 2016; Bertin and Arnouts, 1996].

TESS

The Transiting Exoplanet Survey Satellite [TESS, Ricker et al., 2015] observed WASP-150b during its northern hemisphere campaign. WASP-150b was observed in the full frame images of sectors 14 and 20. To account for the blending in the TESS lightcurve (TESS

---

\(^5\)https://telescope.livjm.ac.uk/TelInst/Inst/RISE/
\(^6\)http://www.iac.es/00CC/instrumentation/telescopio-carlos-sanchez/
Table 3.3: Photometric follow-up of WASP-176

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Date (night of)</th>
<th>Filter</th>
<th>No. points</th>
<th>Average cadence (s)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPECULOOS-Io</td>
<td>14/06/2018</td>
<td>Sloan $z'$</td>
<td>1127</td>
<td>21</td>
<td>Missed egress</td>
</tr>
<tr>
<td>TRAPPIST-North</td>
<td>26/06/2018</td>
<td>$I + z$</td>
<td>708</td>
<td>28</td>
<td>Full but no out of transit (meridian flip)</td>
</tr>
<tr>
<td>Cassini</td>
<td>30/06/2018</td>
<td>Johnson $V$</td>
<td>123</td>
<td>119</td>
<td>Missed ingress (cloud)</td>
</tr>
<tr>
<td>NITES</td>
<td>04/08/2018</td>
<td>Johnson-Bessel $R$</td>
<td>340</td>
<td>32</td>
<td>Missed ingress</td>
</tr>
<tr>
<td>NITES</td>
<td>08/08/2018</td>
<td>Johnson-Bessel $I$</td>
<td>676</td>
<td>32</td>
<td>Missed ingress</td>
</tr>
<tr>
<td>RISE-2</td>
<td>24/09/2018</td>
<td>$V + R$</td>
<td>1929</td>
<td>11</td>
<td>Missed ingress</td>
</tr>
</tbody>
</table>

pixels are 21″ square) this data was included with a variable dilution term in the MCMC modelling for WASP-150b. This light curve was extracted from the full frame images using Eleanor extraction pipeline [Feinstein et al., 2019] utilising the background subtraction and systematics removal packages.

3.4.2 WASP-176b

To gather the necessary follow-up photometry for WASP-176b, a number of telescopes were used. These are summarised in Table 3.3 along with observation dates, photometric filters and transit notes. The following sections detail the follow-up. The lightcurves are shown in Figures 3.9 and 3.10. Data files can be found at https://github.com/BenCooke95/W150-W176.

Cassini 1.52 m

A partial transit of WASP-176b was recorded with the Cassini 1.52 m telescope on 30 June 2018. Unfortunately, cloud prevented observing the start of the transit. The observations were performed by defocusing the telescope for improving the photometric precision and using autoguiding. The data were reduced as discussed in Section 3.4.1. We detrended the light curve to remove slow instrumental and astrophysical trends, by fitting a straight line to the out-of-transit data.

NITES

A total of 2 transits were obtained using NITES [McCormac et al., 2014] on La Palma. The data were reduced in Python using ccdproc [Craig et al., 2015]. A master bias, dark and flat were created using the standard process on each night. Twenty one images of each type were used for the master calibration frames. Non-variable nearby comparison stars were selected by hand, and aperture photometry extracted using sep [Barbary, 2016; Bertin and Arnouts, 1996]. The aperture photometry radii were chosen to minimise the dispersion in the data points out of transit.
Figure 3.7: Upper panel: Individual WASP-150 lightcurves binned to 10 minutes (TESS data are left unbinned). From top down the light curves are from NITES, RISE, RISE, TESS, TCS, Cassini, CAHA and IAC80. The red curves show the best fit from MCMC. Lower panel: Best fit residuals coloured as in the upper panel.
RISE-2

WASP-176 was observed with RISE-2 mounted on the 2.3 m telescope situated at Helmos observatory in Greece. RISE-2 has a CCD size of 1K×1K with a pixel scale of 0.51″ and a field of view of 9′×9′ [Boumis et al., 2010]. The data were reduced using master bias and flat frames, created using the standard process on each night. Non-variable nearby comparison stars were selected by hand and aperture photometry extracted using sep [Barbary, 2016; Bertin and Arnouts, 1996]. The aperture photometry radii were chosen to minimise the dispersion in the data points out of transit.

TRAPPIST

TRAPPIST-North [Gillon et al., 2017; Barkaoui et al., 2019] observed one full transit of WASP-176b on 26 June 2018. TRAPPIST-North is a 60 cm robotic telescope installed in spring 2016 at Oukaimeden Observatory in Morocco. TRAPPIST-North is a northern twin of TRAPPIST-South [Jehin et al., 2011; Gillon et al., 2011]. TRAPPIST-North has an f/8 Ritchey-Chretien optical design. It is equipped with a thermoelectrically-cooled 2048 × 2048 deep-depletion Andor iKon-L CCD camera that has a pixel size of 13.5 µm, which translates into a 0.60″ pixel⁻¹ image scale and a field of view of 19.8′×19.8′. Data
reduction consisted of standard calibration steps (bias, dark and flat-field corrections) and subsequent aperture photometry using IRAF/DAOPHOT [Tody, 1986]. Extraction of fluxes of selected stars using aperture photometry was performed with IRAF/DAOPHOT (as described in Gillon et al. 2013).

During the TRAPPIST-North observations of WASP-176 the telescope underwent a meridian flip at JD 2458296.6355. To counter this problem the pre- and post-meridian flip data are treated separately.

SPECULOOS

One partial transit of WASP-176b was observed with SPECULOOS-Io, one of the four telescopes of the SPECULOOS-South facility [Gillon, 2018; Delrez et al., 2018; Burdanov et al., 2018; Jehin et al., 2018], which is located at ESO Paranal Observatory (Chile). Each telescope is a robotic Ritchey-Chretien (f/8) telescope of 1m diameter. They are equipped with Andor iKon-L Peltier-cooled deeply depleted 2K×2K CCD cameras, with good sensitivities in the very-near-infrared up to 1 \( \mu \text{m} \). The field of view of each telescope is 12’ × 12’ and the pixel scale is 0.35” pixel\(^{-1}\). The calibration and photometric reduction of the data were performed as described in Gillon et al. [2013].

3.5 High-spatial-resolution follow-up

WASP-150 was observed on four occasions: 15 November 2015, 9 March 2016, 10 March 2016 and 6 May 2016; using the FastCam camera [Oscoz et al., 2008] installed on the 1.52 m Telescopio Carlos Sánchez (TCS) located at the Observatorio del Teide, Tenerife, Spain. FastCam is a EMCCD camera with 512 × 512 pixels, with a physical pixel size of 16 \( \mu \text{m} \), which makes a FOV of 21.2” × 21.2”. Thanks to the very low noise and fast readout speed of the EMCCD array, this camera is appropriate for Lucky Imaging (LI) observations.

During each of the four observing nights, 10,000 individual frames of WASP-150 were collected in the Johnson-Cousins I filter, except on the 15 November 2015 night, for which clear filter was used due to the dusty weather conditions (\textit{calima}). Each individual frame had an exposure time of 50 ms. In total, 500 seconds and 1,500 seconds of data were gathered of WASP-150, with the clear and I-band filters, respectively.

Using the FastCam dedicated software developed at the Universidad Politécnica de Cartagena [Labadie et al., 2010; Jódar et al., 2013], each individual frame was bias-subtracted, aligned and co-added and then processed in order to construct a high resolution, long-exposure image. For each night’s data set, we took a high resolution image constructed by co-addition of the best 30% of images, thus making a 150 seconds total exposure time.

No close companion was detected, only a \( \Delta m_I = 2.17 \pm 0.03 \) mag fainter star at a distance
Figure 3.9: *Upper panel:* Individual WASP-176 lightcurves binned to 10 minutes. From top down the light curves are from NITES, TRAPPIST, NITES, SPECULOOS, Cassini and RISE2. The red curves show the best fit from MCMC.

*Lower panel:* Best fit residuals coloured as in the upper panel.
of 10.58 ± 0.05 arcsec. Figure 3.11 shows the contrast light curve that was computed based on the scatter within the annulus as a function of angular separation from the target centroid [see, e.g., Gauza et al., 2015].

3.6 Results

3.6.1 WASP-150b

Stellar parameters

To determine the stellar parameters of WASP-150 a spectral analysis of the SOPHIE spectra was carried out. The results of this analysis are shown in Table 3.4.

Additionally, the second data release of Gaia [Gaia Collaboration et al., 2016, 2018b] provided some stellar parameters presented in Table 3.5. Gaia was also searched for close companions of WASP-150. No significant companion was found.

The open-source stellar modelling code bagemass\(^7\) [Maxted et al., 2015] was then used to estimate the age and mass of WASP-150. Bagemass uses the Garstec stellar evolution code [Weiss and Schlattl, 2008] to calculate model grids of individual stars. A Bayesian

\(^7\)https://sourceforge.net/projects/bagemass/
Figure 3.11: I-band magnitude contrast as a function of angular separation up to 4.0″ of WASP-150. The solid line indicates the 5-σ detection limit for primary star. The inset shows the 2.0 × 2.0″ combined image of WASP-150. North is up and East is left.

Table 3.4: Stellar parameters from spectral analysis

<table>
<thead>
<tr>
<th>Parameter (unit)</th>
<th>WASP-150</th>
<th>WASP-176</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral type</td>
<td>F8</td>
<td>F9</td>
</tr>
<tr>
<td>$T_{\text{eff}}$ (K)</td>
<td>6250 ± 80</td>
<td>6100 ± 100</td>
</tr>
<tr>
<td>log g</td>
<td>4.23 ± 0.13</td>
<td>4.0 ± 0.2</td>
</tr>
<tr>
<td>$[Fe/H]$</td>
<td>0.18 ± 0.11</td>
<td>0.15 ± 0.08</td>
</tr>
<tr>
<td>$v \sin i$ (kms$^{-1}$)</td>
<td>8.82 ± 0.95$^a$</td>
<td>3.8 ± 1.0$^b$</td>
</tr>
</tbody>
</table>

$^a$ Assuming a microturbulence of 0.78 ± 0.05 kms$^{-1}$ from Doyle et al. [2013] calibration.

$^b$ Assuming a macroturbulence of 5.1 ± 0.7 kms$^{-1}$ from Doyle et al. [2013] calibration.
Table 3.5: Stellar parameters from Gaia DR2

<table>
<thead>
<tr>
<th>Parameter (unit)</th>
<th>W ASP-150</th>
<th>W ASP-176</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallax (mas)</td>
<td>1.865 ± 0.020</td>
<td>1.731 ± 0.036</td>
</tr>
<tr>
<td>Distance (pc)</td>
<td>536 ± 6</td>
<td>578 ± 12</td>
</tr>
<tr>
<td>PM R.A. (mas yr(^{-1}))</td>
<td>−4.289 ± 0.041</td>
<td>−6.192 ± 0.064</td>
</tr>
<tr>
<td>PM Dec. (mas yr(^{-1}))</td>
<td>7.000 ± 0.040</td>
<td>−4.954 ± 0.062</td>
</tr>
<tr>
<td>(T_{\text{eff}}) (K)</td>
<td>6093(^{+188}_{-70})</td>
<td>5902(^{+20}_{-36})</td>
</tr>
<tr>
<td>(R_{\star}) (R(_{\odot}))</td>
<td>1.706(^{+0.040}_{-0.101})</td>
<td>1.945(^{+0.024}_{-0.013})</td>
</tr>
<tr>
<td>(L_{\star}) (L(_{\odot}))</td>
<td>3.616 ± 0.0073</td>
<td>4.136 ± 0.130</td>
</tr>
</tbody>
</table>

Table 3.6: Stellar properties from \textsc{bagemass}

<table>
<thead>
<tr>
<th>Parameter (unit)</th>
<th>W ASP-150</th>
<th>W ASP-176</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass (M(_{\odot}))</td>
<td>1.346 ± 0.029</td>
<td>1.270 ± 0.025</td>
</tr>
<tr>
<td>(\tau_{\text{iso}}) (Gyr)</td>
<td>2.950 ± 0.229</td>
<td>4.810 ± 0.191</td>
</tr>
<tr>
<td>([Fe/H]_{\text{init}})</td>
<td>0.204 ± 0.079</td>
<td>0.215 ± 0.069</td>
</tr>
</tbody>
</table>

method then samples the posterior distributions on mass and age. The results of this analysis are presented in Table 3.6.

The best-fit evolutionary track and isochrone produced from the \textsc{bagemass} analysis are shown in Figure 3.12 along with the 1\(\sigma\) uncertainties. This plot also includes the posterior distribution produced by the EXOFASTv2 analysis (see section 3.6.1).

Figure 3.12 includes a small collection of points above the main area of convergence. These data are not affected by increasing the run-time of the MCMC analysis or by increasing the burn-in period thus we do not believe them to be an artefact of unfinished fitting. However, since the significance of this region is \(\sim 10\) times lower than the peak we do not find it impactful.

**EXOFASTv2 analysis**

To perform simultaneous fitting of the SuperWASP detection, the SOPHIE RVs and the follow-up photometry we used the fitting code EXOFASTv2 [Eastman, 2017; Eastman et al., 2019]. This tool is designed to fit all the available data and ensure consistency between derived stellar and planetary parameters. EXOFASTv2 explores the given parameter space through a differential evolution Markov chain method using 30,000 steps. We use the Gelman-Rubin statistic [Gelman et al., 2003] to check the mixing of the chains as proposed by Ford [2006]. We fit a total of 50 free parameters, these are the parameters presented in Table 3.7 as well as limb darkening parameters for each band used and offset baselines and variances to account for any errors in normalisation.

For the EXOFASTv2 input parameters we take period and epoch from the initial
Figure 3.12: *bagemass* analysis of WASP-150. The black line shows the Zero Age Main Sequence (ZAMS). The orange line shows the best fit mass track with the dotted lines showing the $\pm1\sigma$ mass tracks. The red line shows the isochrone relating to the best fit age with the dotted lines showing the $\pm1\sigma$ isochrones. Finally, the density of EXOFASTv2 samples is shown in the colour scale of the plotted posterior distribution.
SuperWASP discovery photometry. Metallicity and effective temperature are taken from the results of spectroscopic analysis. Additionally, we use stellar radius and luminosity as well as parallax and distance from Gaia DR2. Finally we impose a prior on the V-band extinction from Schlegel et al. [1998] and Schlafly and Finkbeiner [2011] which are used to model the stellar properties through SED fitting. Within EXOFASTv2 we use the MESA Isochrones and Stellar Tracks [MIST, Dotter, 2016; Choi et al., 2016] to model the star. This produces an age of $2.18^{+0.58}_{-0.65}$ Gyr, which is in reasonable agreement with the bagemass age in Table 3.6. The best fit values, along with uncertainties, are presented in Table 3.7.

3.6.2 WASP-176b

Stellar parameters

Similar to WASP-150, the stellar parameters of WASP-176 are calculated via a spectral analysis of the CORALIE spectra. The results of this analysis are shown in Table 3.4.

Once again, the second data release of Gaia [Gaia Collaboration et al., 2016, 2018b] provided some additional stellar parameters presented in Table 3.5. Gaia was also searched for close companions of WASP-176.

To determine the age of WASP-176 we again ran bagemass and show the results in Table 3.6.

The best-fit evolutionry track and isochrone produced from the bagemass analysis are shown in Figure 3.13 along with the 1σ uncertainties. This plot also includes the posterior distribution produced by the MCMC analysis (see section 3.6.2).

Once again, we see that Figure 3.13 includes a few points above the main peak. As before we find that these data are not affected by increasing the run-time of the MCMC analysis or by increasing the burn-in period so, again, we do not believe them to be an artefact of unconvurged fitting. Since the significance of this region is $\sim$10 times lower than the peak we do not find it impactful.

EXOFASTv2 analysis

As for WASP-150b we once again use EXOFASTv2 to derive the quoted parameters for WASP-176b. We use the same number of steps, 30,000, and utilise the Gelman-Rubin statistic to check the mixing of the chains. We fit 43 free parameters, these are the parameters presented in Table 3.7 as well as limb darkening parameters for each band used and offset baselines and variances to account for any errors in normalisation. In the same way as for WASP-150b we take the EXOFASTv2 input parameters from a combination of the WASP discovery photometry (period and epoch), spectroscopic analysis (metallicity and
Figure 3.13: bagemass analysis of WASP-176. The black line shows the Zero Age Main Sequence (ZAMS). The orange line shows the best fit mass track with the dotted lines showing the ±1σ mass tracks. The red line shows the isochrone relating to the best fit age with the dotted lines showing the ±1σ isochrones. Finally, the density of MCMC samples is shown in the colour scale of the plotted posterior distribution.
effective temperature) and Gaia DR2 (stellar radius, luminosity, parallax and distance). Additionally we impose a V-band extinction prior from Schlegel et al. [1998] and Schlafly and Finkbeiner [2011]. We use the MESA Isochrones and Stellar Tracks [MIST, Dotter, 2016; Choi et al., 2016] to model the star. This produces an age of $3.69^{+1.9}_{-0.83}$ Gyr, which is in good agreement with the bagemass age in Table 3.6. The best fit values, along with uncertainties, are presented in Table 3.7.

The WASP-176b analysis was first conducted allowing for an eccentric orbit. This resulted in $\chi^2_{ecc} = 11.423$ Repeating the analysis, this time forcing a circular orbit gives $\chi^2_{circ} = 11.358$. Since the discrepancy between these values is negligible it was decided there was insufficient evidence to support eccentricity and thus a circular orbit was assumed.

### 3.7 Discussion and conclusions

#### 3.7.1 WASP-150b

WASP-150b is a high density hot Jupiter on a 5.644207$^{+0.000003}_{-0.000004}$ day orbit around its F8 host. With a mass of of $8.46^{+0.28}_{-0.2}$ M$_J$ and a radius of $1.07^{+0.024}_{-0.025}$ R$_J$ WASP-150b has a density of $6.44^{+0.50}_{-0.47}$ $\rho_J$ placing it amongst the highest density planets known. Figure 3.14 shows a plot of all the confirmed exoplanets from the NASA exoplanet archive with radius and mass known to an accuracy of $\leq 10\%$ and periods of $\leq 10$ days. Though large, this density is in line with expectations based on the planets mass [Chabrier et al., 2009].

WASP-150b is also relatively eccentric ($e = 0.3775^{+0.0038}_{-0.0029}$). Using equation (3) from Adams and Laughlin [2006] (and assuming $Q_p \sim 10^5$) we see that the approximate circularisation time-scale for this system is $\sim 5$ Gyr. Our analysis of this system gives the age of this star as $2.950 \pm 0.229$ Gyr, well below the circularisation time-scale, showing that the eccentricity is compatible with the other system parameters. On top of this it has been shown that more massive planets are more inclined to larger eccentricities [Ribas and Miralda-Escudé, 2007] although their origin is still an open question. Additionally, we can use a simple scaling relation to predict the main-sequence lifetime of WASP-150. The following relation, $t_{MS} = t_{MS,\odot} (M/M_\odot)^{-2.5}$, where $t_{MS,\odot}$ is the main-sequence lifetime of the sun ($\sim 10$ Gyr), predicts the main-sequence lifetime of WASP-176 as $\sim 4.3$ Gyr, much longer than its estimated age.

For some examples of comparable high-mass, hot Jupiters in eccentric orbits see WASP-8b [2.2 M$_J$, $e = 0.31$; Queloz et al., 2010], Kepler-75 [9.9 M$_J$, $e = 0.57$; Hébrard et al., 2013], WASP-162b [5.2 M$_J$, $e = 0.43$; Hellier et al., 2019] and HAT-P-2b [8.74 M$_J$, $e = 0.52$; Bakos et al., 2007]. HAT-P-2b was the first exoplanet discovered in this class and has been extensively studied in regards to planet-star interactions. The large mass and eccentricity of WASP-150b, coupled with its high eccentricity, make it an interesting addition
<table>
<thead>
<tr>
<th>Parameter (unit)</th>
<th>WASP-150b</th>
<th>WASP-176b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_0$ (HJD)</td>
<td>7217.2614$_{-0.0004}^{+0.0004}$</td>
<td>8234.1771$_{-0.0007}^{+0.0007}$</td>
</tr>
<tr>
<td>$P$ (days)</td>
<td>5.644207$_{-0.000004}^{+0.000003}$</td>
<td>3.899052$_{-0.000005}^{+0.000005}$</td>
</tr>
<tr>
<td>$\Delta F$</td>
<td>0.0044$_{-0.0001}^{+0.0001}$</td>
<td>0.0064$_{-0.0002}^{+0.0002}$</td>
</tr>
<tr>
<td>$T_{14}$ (days)</td>
<td>0.1299$_{-0.0015}^{+0.0001}$</td>
<td>0.2147$_{-0.0019}^{+0.0021}$</td>
</tr>
<tr>
<td>$b$</td>
<td>0.758$_{-0.014}^{+0.011}$</td>
<td>0.347$_{-0.12}^{+0.098}$</td>
</tr>
<tr>
<td>$i$ (°)</td>
<td>84.01$_{-0.2}^{+0.25}$</td>
<td>86.7$_{-1.1}^{+1.3}$</td>
</tr>
<tr>
<td>$e$</td>
<td>0.3775$_{-0.0029}^{+0.0038}$</td>
<td>0.0 (Fixed)</td>
</tr>
<tr>
<td>$M_*$ ($M_\odot$)</td>
<td>1.394$_{-0.049}^{+0.007}$</td>
<td>1.345$_{-0.13}^{+0.08}$</td>
</tr>
<tr>
<td>$R_*$ ($R_\odot$)</td>
<td>1.651$_{-0.03}^{+0.024}$</td>
<td>1.925$_{-0.044}^{+0.047}$</td>
</tr>
<tr>
<td>$\rho_*$ ($\rho_\odot$)</td>
<td>0.439$_{-0.025}^{+0.032}$</td>
<td>0.263$_{-0.032}^{+0.03}$</td>
</tr>
<tr>
<td>log $g_*$ (cgs)</td>
<td>4.147$_{-0.02}^{+0.027}$</td>
<td>3.995$_{-0.053}^{+0.037}$</td>
</tr>
<tr>
<td>$T_{eff}$ (K)</td>
<td>6218.0$_{-45.0}^{+49.0}$</td>
<td>5941.0$_{-79.0}^{+77.0}$</td>
</tr>
<tr>
<td>[Fe/H]</td>
<td>0.156$_{-0.089}^{+0.1}$</td>
<td>0.164$_{-0.082}^{+0.081}$</td>
</tr>
<tr>
<td>$M_{pl}$ ($M_J$)</td>
<td>8.46$_{-0.2}^{+0.28}$</td>
<td>0.855$_{-0.069}^{+0.072}$</td>
</tr>
<tr>
<td>$R_{pl}$ ($R_J$)</td>
<td>1.07$_{-0.025}^{+0.024}$</td>
<td>1.505$_{-0.045}^{+0.05}$</td>
</tr>
<tr>
<td>$\rho_{pl}$ ($\rho_J$)</td>
<td>6.44$_{-0.47}^{+0.50}$</td>
<td>0.234$_{-0.032}^{+0.032}$</td>
</tr>
<tr>
<td>log $g_{pl}$ (cgs)</td>
<td>4.263$_{-0.023}^{+0.024}$</td>
<td>2.972$_{-0.053}^{+0.047}$</td>
</tr>
<tr>
<td>$T_{pl}$ (K)</td>
<td>1460.0$_{-11.0}^{+11.0}$</td>
<td>1721.0$_{-21.0}^{+28.0}$</td>
</tr>
<tr>
<td>$a$ (AU)</td>
<td>0.0694$_{-0.0008}^{+0.0011}$</td>
<td>0.0535$_{-0.0019}^{+0.0001}$</td>
</tr>
</tbody>
</table>
Figure 3.14: Scatter plot of all confirmed exoplanets from the NASA exoplanet archive with radius and mass known to an accuracy of $\leq 10\%$ and periods of $\leq 10$ days. The points are then coloured by density in cgs units. WASP-150b is shown as a red diamond and Jupiter is shown as a black diamond for reference (WASP-176b is shown as an orange diamond). The plot also includes lines of constant density to guide the eye.
to similar studies [Cébron et al., 2013; Salz et al., 2016; de Wit et al., 2017]. In particular
the short period of this system allows for the collection of entire phase curves which will en-
able studies of the evolution of the planetary flux as a function of orbital phase [Lewis et al.,
2013]. Additionally, the eccentricity means that the planet cannot rotate synchronously.
Thus the atmosphere should display time-dependent effects such as atmospheric radiative
time constants and tidal luminosities [Laughlin and Lissauer, 2015]. Future studies may
even be able to detect the manifestation of time-dependent storms [Laughlin et al., 2009].

3.7.2 WASP-176b

WASP-176b is a slightly inflated hot Jupiter orbiting an F9 host star on a 3.899052 ± 0.000005 day
orbit. The MCMC best-fit results presented here give a planetary mass of 0.855$^{+0.072}_{-0.066}$ M$_J$
and a planetary radius of 1.505$^{+0.045}_{-0.045}$ R$_J$ leading to a density of 0.234$^{+0.032}_{-0.032}$ ρ$_J$.
Using the same scaling relation as above predicts the main-sequence lifetime of WASP-176 as ∼4.77 Gyr.
Using the isochronal age estimated in section 3.6.2 indicates that the star has evolved past
the end of its main-sequence life. This comparison, combined with the stellar surface gravity
and density given in Table 3.7 indicates that the host is a sub-giant. This is further
supported by its location on the HR diagram as seen in Figure 3.13. The planetary radius
found here is approximately 25% larger than expected for a coreless planet, when predicted
using the planetary evolution models from Fortney et al. [2006].

WASP-176b is similar to other hot Jupiters (WASP-54b: Faedi et al. [2013]; WASP-
78b and WASP-79b: Smalley et al. [2012]; WASP142b: Hellier et al. [2017]; WASP-
136b: Lam et al. [2017]) in that it receives stronger irradiation from its F-type host than
it would from a G-type star, thus leading to inflation. Based on this increased radiation
we suggest that the inflation mechanism behind WASP-176b may be due to the deposit of
stellar irradiation into the planetary core, consistent with the Class I model presented in
Lopez and Fortney [2016]. If so, then this planet may have experienced increasing inflation
as its host moved off the main sequence. However, this is only a possible solution, additional
characterisation may give more information to this end.

WASP-176b is a potential target for atmospheric characterisation via transmission
spectroscopy due to its low density and high equilibrium temperature. If we assume an
atmosphere similar in composition to Jupiter ($μ = 2.2u$, where $u = 1.66\times10^{-27}$ kg is the
atomic mass unit) then the scale height is found to be $H = k_B T / μg ∼ 500$ km which is
smaller, but comparable, to values seen in recent successful atmospheric retrievals [Kirk
et al., 2019]. However, with a magnitude of only $V = 12.01$, the follow-up would be a sig-
nificant challenge.
Chapter 4

TESS primary mission monotransit simulation

Note

The following chapter is a reproduction of the paper ‘Single Site Observations of TESS Single Transit Detections’ [Cooke et al., 2018] published in Astronomy & Astrophysics in November 2018. The simulations, results and analysis presented here were all designed and carried out by me with the exception of the custom tool referred to in Section 4.5 which was written by James McCormac. The writing was all done by myself. Co-author contributions consisted of discussions and improvements regarding methodology and results as well as general proof reading.

This paper describes a simulation to predict the number of monotransit candidates that will be found in the TESS primary mission. Additionally, it goes on to examine the feasibility of ground based follow-up of the targets from a single site. Some of the simulation details are reused and updated in Chapters 5, 6 and 8.

Abstract

Context. TESS has been successfully launched and has begun data acquisition. To expedite the science that may be performed with the resulting data it is necessary to gain a good understanding of planetary yields. Given the observing strategy employed by TESS the probability of detecting single transits in long period systems is increased. These systems require careful consideration.

Aims. We aim to simulate the number of TESS transit detections during its two year mission with a particular emphasis on single transits. We also aim to determine the feasibility of
ground-based follow-up observations from a single site.

**Methods.** A distribution of planets was simulated around the approximately four million stars in the TESS candidate target list. These planets were tested for detectable transits and characterised. Based on simulated parameters the single transit detections were further analysed to determine which are amenable to ground-based follow-up.

**Results.** TESS will discover an approximate lower bound of 4700 planets with around 460 being single transits. A large fraction of these will be observable from a single ground-based site. This paper finds that, in a single year, approximately 1000 transit events of around 320 unique TESS single transit detections are theoretically observable.

**Conclusions.** As we consider longer period exoplanets the need for exploring single transit detections increases. For periods $\gtrsim 45$ days the number of single transit detections outnumber multitransits by a factor of three (82±18 and 25±7 respectively) a factor which only grows as longer period detections are considered. Therefore, based on this paper, it is worth expending the extra effort required to follow-up these more challenging, but potentially very rewarding, discoveries. Additionally, we conclude that a large fraction of these targets can be theoretically observed from a single ground-based site. However, further work is required to determine whether these follow-up efforts are feasible when accounting for target specific criteria.

### 4.1 Introduction

As of 28 June 2018 there are 3735 confirmed exoplanet discoveries [NASA Exoplanet Archive\(^1\), Akeson et al., 2013] with 2934 having been discovered via the transit method. The vast majority of these detections come from a small number of efficacious surveys including Kepler [Borucki et al., 2010, 2327 detections], K2 [Howell et al., 2014, 294 detections], the Super Wide Angle Search for Planets [SuperWASP, Pollacco et al., 2006, 125 detections], the Hungarian Automated Telescope survey [HAT, Bakos et al., 2004, 101 detections], the Convection, Rotation and planetary Transits satellite [CoRoT, Baglin, 2003, 30 detections] and the Kilodegree Extremely Little Telescope survey [KELT, Siverd et al., 2012, 19 detections]. The Transiting Exoplanet Survey Satellite [TESS, Ricker et al., 2014] is expected to discover in the order of $10^3$ exoplanets [Sullivan et al., 2015] which will make it the most prolific exoplanet discovery mission ever.

**TESS** launched on 18 April 2018 and will search for transiting planets around bright, nearby stars [Ricker et al., 2014]. TESS will be unrivalled in its ability to find hundreds of sub-Neptune radius planets around bright stars that provide amenable targets for follow-up and atmospheric characterisation [Crouzet et al., 2017]. Thanks to its large viewing

\(^1\)https://exoplanetarchive.ipac.caltech.edu/index.html
area (>85% of the sky will be observed to some degree) current simulations predict TESS will discover thousands of planets orbiting stars with $m_V \leq 12$ [Sullivan et al., 2015; Bouma et al., 2017; Ballard, 2019; Barclay et al., 2018; Villanueva et al., 2019]. Subsequent follow-up observations will enable the determination of planet mass via radial velocity measurements [Cloutier et al., 2018] and the characterisation of planet atmosphere via transmission spectroscopy [Kempton et al., 2018].

The observing strategy employed by TESS, combined with the fact that some areas of the sky will be observed for long baselines, means that the chance of detecting the transits of long period planets is increased compared to shorter duration missions. For example, each K2 field is observed for \( \sim 75 \) days [Howell et al., 2014] whereas some regions of TESS fields will reach baselines of >300 days [Ricker et al., 2014]. However, long period orbits then mean that it is likely that only a single transit event will be observed. Traditionally, transit detections are confirmed using many repeated transits folded along the period of the orbit leading to improved signal detection and signal to noise ratio (S/N). Obviously this is not possible with single transits and so these need to be treated differently to avoid losing potential discoveries. Using Kepler and K2 as a guide it has been shown that observations of just a single transit can still lead to accurate orbital parameters on the 10 – 20\% level at 30-minute cadence [Yee and Gaudi, 2008; Osborn et al., 2016] meaning they are still worth the additional effort required to confirm. Follow-up however, will need to be more rigorous, since the ephemerides are less constrained than for a system with multiple observed transits, and will require significant follow-up effort from high-precision ground-based facilities, such as the Next Generation Transit Survey (NGTS, Wheatley et al. 2018) and space-based observatories such as the James Webb Space Telescope (JWST, Gardner et al. 2006).

To this end a detailed simulation of the number of TESS single transits must be carried out and then extended to determine how many of the events may be observed from the ground to confirm the periodic nature of the system. We therefore created a simulation of TESS detections of single transits and then determined their follow-up observability from a single ground-based site.

We have set out the paper in the following way. Section 4.2 discusses the stellar population used in the project with Sect. 4.3 describing the generation of a planetary population. Next, Sect. 4.4 determines the detectability of these simulated planets by TESS and Sect. 4.5 details their potential for single site follow-up observations. Section 4.6 displays the simulation results and finally, Sect. 4.7 outlines some points of discussion and the conclusions.
4.2 Stellar population

Our simulation was built using the 3.8 million stars in the TESS input catalogue (TIC) candidate target list (CTL) available from the Mikulski Archive for Space Telescopes (MAST\(^2\)). The key parameters within the CTL are stellar properties including TESS magnitude \((m_{\text{TESS}})\), mass \((M_*)\), radius \((R_*)\), effective temperature \((T_{\text{eff}})\) and luminosity \((L_*)\) as well as positional parameters such as R.A. & Dec., latitude & longitude (ecliptic) and contamination ratio (for a full description of the CTL and its parameters see Stassun et al. 2018). Additionally, it was necessary to know for how long each star (and therefore each potential planet) will be observed for. To determine this baseline value we simulated 13 sectors (per ecliptic hemisphere) on the sky, each one a \(24^\circ \times 96^\circ\) rectangular region [Ricker et al., 2014]. Each CTL star was then checked using ecliptic coordinates and it was determined how many sectors would cover this region and which sectors (numbered from 1-13) they would be. It was assumed that the 13 sectors were evenly distributed around the sky and reflected exactly about the ecliptic equator. Additionally, since the initial spacecraft pointing is not yet known [Bouma et al., 2017], it was decided that the first sector would be centred on the line of longitude \(= 0^\circ\). Figure 4.1 shows a plot of all CTL stars coloured by the number of regions for which they will be observed. The observing baseline for each target was then taken to be 27.4 days multiplied by the number of sectors which observe that target [Ricker et al., 2014]. As an additional caveat, TESS will spend \(~6.5\) hrs per 13.7 day orbit transferring data during which time data acquisition will halt. This time was included in these simulations as non-observing time. We were then able to simulate a synthetic population of planets around these stars and the systems were checked for detections.

Figure 4.2 shows the percentage of sky covered by combinations of sectors; from zero along the ecliptic equator, up to 13 in the TESS continuous viewing zone (CVZ). This is then shown alongside the percentage of CTL stars observed by the same number of sectors. Discrepancies between the two sets of results is due to the non-uniformity of CTL stars when distributed across the celestial sphere (for example, stellar density is increased in the galactic plane).

As an additional step we estimated the spectral type of each star based on their \(T_{\text{eff}}\) values. The chosen \(T_{\text{eff}}\) thresholds come from Pecaut and Mamajek [2013] and are given in Table 4.1. The spectral type is needed to more accurately determine the occurrence rate of planets, as discussed in Sect. 4.3.

Figure 4.3 shows the stellar distribution of TESS targets on an HR diagram. The distribution is split up based on observing cadence used for each target (as described in Sect. 4.4).

\(^2\)https://archive.stsci.edu/tess/bulk_downloads.html
Figure 4.1: All TIC CTL stars in the northern ecliptic hemisphere coloured by number of observing regions. The north ecliptic is shown but the pattern is mirrored on the southern ecliptic hemisphere.

Figure 4.2: TESS sky coverage showing area of sky observed by combinations of regions. Sky areas observed by more than 4 and less than 12 regions are combined as individually these regions are very small. The plot shows sky area in red and number of stars in blue, both given as a percentage of the total.
Table 4.1: Spectral type classification as a function of $T_{\text{eff}}$.

<table>
<thead>
<tr>
<th>Spectral Type</th>
<th>$T_{\text{eff}}$ Range (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>7330 - 10050</td>
</tr>
<tr>
<td>F</td>
<td>5980 - 7330</td>
</tr>
<tr>
<td>G</td>
<td>5310 - 5980</td>
</tr>
<tr>
<td>K</td>
<td>3905 - 5310</td>
</tr>
<tr>
<td>M</td>
<td>2285 - 3905</td>
</tr>
</tbody>
</table>

Figure 4.3: TIC CTL stellar distribution used in this simulation. Stars observed at two-minute cadence (postage stamps) are shown in blue and stars observed at 30-minute cadence (full frame images) are in green.
4.3 Planet population

The simulated planet population was created by drawing from two distributions according to spectral type. The initial step was to determine the planet occurrence rates. The occurrence rate of planets, $\mu$, was taken to be 2.5 for M stars [Dressing and Charbonneau, 2015] and 0.689 for AFGK stars [Fressin et al., 2013]. Taking Barclay et al. [2018] as a guide, the number of planets around each star was then predicted by drawing from a Poisson distribution centred on $\mu$ and is thus dependent on spectral type and therefore $T_{\text{eff}}$.

Once the number of planets per star was decided the radius and period of these planets were drawn from two separate distributions depending on whether the star was an M type (in which case the distribution presented by Dressing and Charbonneau [2015] was employed) or an AFGK type (in which case the distribution presented by Fressin et al. [2013] was used). The distributions presented in these papers give results in terms of period and radius bins. To choose parameters a bin was chosen at random, weighted by the number of systems within that bin divided by the total number of systems in the sample. Once a bin was chosen the precise value of radius and period were chosen from a distribution between the limits of the bin. For radius the distribution was uniform between the bin limits whereas for period a uniform logarithmic distribution was used to give greater weighting to shorter period planets for which the period distributions are more complete [Villanueva et al., 2019]. Figure 4.4 shows the input distributions in period and radius from Dressing and Charbonneau [2015] and Fressin et al. [2013]. We note that the input distributions used are only complete up to 85 and 200 days for AFGK and M stars respectively. To explore longer periods the final period bins were extended up to 1000 days (as demonstrated in Villanueva et al. [2019]).

For each simulated planet the orbital separation was then calculated using Kepler’s third law,

$$a = \sqrt[3]{\frac{GM_\star P^2}{4\pi^2}}. \tag{4.1}$$

Next, values for the inclination, $i$, eccentricity, $e$, and periastron angle, $\omega$, were chosen. The periastron angle was drawn from a uniform distribution between $-\pi$ and $\pi$ whereas eccentricity was drawn from a beta distribution with $\alpha = 1.03$ and $\beta = 13.6$ which has been found to be a valid assumption for transiting exoplanet populations [Van Eylen and Albrecht, 2015]. Inclination was stored as $\cos i$ and was chosen from a uniform distribution between 0 and 1. Inclination was assumed to be the same for all planets in the same system. Finally, the transit duration $T_{\text{dur}}$ was calculated using equation (1) from Barclay et al. [2018];
Figure 4.4: Input planet distributions used to simulate TESS planets. The distributions for M stars are shown in red and AFGK stars are in blue. Panel a: Input distribution of planets with period. The dotted lines show the extension of the last period bins up to 1000 days. Panel b: Input distribution of planets with radius.

\[ T_{dur} = \frac{P}{\pi} \times \arcsin \left( \frac{R_* a}{1 - e^2} \times \frac{\sqrt{1 + \frac{e^2}{R_*^2} - b^2}}{\sqrt{1 - \cos^2 i}} \right) \]  

(4.2)

### 4.4 Detectability

To determine whether a planet will have a detectable transit during its observations by TESS it must first be determined whether the planet transits at all. To determine this the impact parameter was calculated using an equation from Winn [2010b] reproduced here,

\[ b = \frac{a \cos i}{R_*} \times \frac{1 - e^2}{1 + e \sin \omega} \]  

(4.3)

A transit is defined as occurring for \(|b| < 1\). This is a good approximation for the majority of systems excluding large planets around M stars. However, since these systems are relatively rare [Dressing and Charbonneau, 2015] the approximation can be used without significant effects on the simulation outcome [Barclay et al., 2018].

Next it must be determined whether a transit will occur during the observation of its host. A random number, designated \(T_0\), was generated from 0 to \(P\) (where \(P\) is the orbital period). If \(T_0 < T_{obs}\) a transit will be observed. Here \(T_{obs}\) is the length of time for which a planet system is observed by TESS. Further comparison of \(T_0\) to \(T_{obs}\), accounting for period, \(P\), shows how many transits are observed during the mission, and this is denoted as \(n\). The final step was then to determine whether the transit signal will be detectable above
the noise. The S/N is given by

$$S/N = \frac{\delta_{eff}}{\sigma_{1hr}} \sqrt{\frac{T_{dur}}{\Delta T}} \sqrt{n}, \quad (4.4)$$

where $\delta_{eff}$ is effective transit depth, $\sigma_{1hr}$ is the total noise measured in 1 hour of data, $n$ is number of observed transits, $T_{dur}$ is transit duration in hours and $\Delta T$ is observing cadence. The factor of $\Delta T$ is included to account for different levels of noise that will be present depending on the cadence of the observations. To be consistent with previous work this is the same S/N prescription as employed by Villanueva et al. [2019] and Barclay et al. [2018].

Following the example of Stassun et al. [2018] we calculated the noise in 1hr using a fifth-order polynomial approximation,

$$\ln(\sigma_{1hr}) = C_1 + C_2m_{TESS} + C_3m_{TESS}^2 + C_4m_{TESS}^3 + C_5m_{TESS}^4 + C_6m_{TESS}^5, \quad (4.5)$$

which is a direct function of the TESS magnitude, $m_{TESS}$. The constants are given as $C_1 = 3.2969$, $C_2 = 0.8500$, $C_3 = -0.2850$, $C_4 = 0.0396$, $C_5 = -0.0022$ and $C_6 = 4.7351 \times 10^{-5}$. Figure 4.5 shows the noise profile employed here (for $3 \leq m_{TESS} \leq 17$) with some of the contributing sources.

The effective transit depth $\delta_{eff}$ was found using

$$\delta_{eff} = \left(\frac{R_p}{R_*}\right)^2 \times \frac{1}{1 + C}, \quad (4.6)$$

where $C$ denotes the contamination ratio from background stars, defined as the ratio of background flux to source flux, $C = F_b/F_s$ [Stassun et al., 2018]. Finally, it must be decided whether each target is observed at two-minute cadence (Postage Stamps) or 30-minute cadence (Full Frame Images) [Ricker et al., 2014]. The choice of cadence value is based on a priority metric;

$$\frac{\sqrt{N_s}}{\sigma_{1hr}R_*^{3/2}}, \quad (4.7)$$

where $N_s$ is the number of sectors for which a target is observed, $\sigma_{1hr}$ is the noise in 1hr and $R_*$ is the stellar radius. This priority metric is provided in the CTL but we recalculated it in this paper to allow for changes resulting from the observing sector locations. The top ranked 200,000 stars were then assigned two-minute cadence with all others being observed at a cadence of 30-minutes. From equation 4.7 it is clear that two-minute cadence targets are biased towards the TESS CVZ through the factor of $\sqrt{N_s}$. A transit was determined to be detectable when it had $S/N > 7.3$, a value chosen so as to be consistent with
Figure 4.5: Total noise in 1 hour as a function of TESS magnitude calculated using equation 4.5. Also shown are individual contributions from star noise, read noise and systematic noise as well as a predicted saturation cut off [Sullivan et al., 2015]. We note that the total noise budget will saturate at 60ppm.

Table 4.2 shows some of the actual numbers of various subsets of the simulated planets. These numbers are determined as the mean of multiple simulation runs with the uncertainties defined as the standard deviation of all runs. Key in these numbers is the separation based on the number of TESS sectors which will observe each potential host. From this table (as well as from Fig. 4.2) we see that the number of stars within each region is almost exactly proportional to the area of sky covered by that region. As before, the minor differences are due to the CTL stars not being exactly uniform in their distribution across the celestial sphere. The values presented in Table 4.2 show that approximately four million planets are simulated with $4721 \pm 181$ being detectable by TESS and $464 \pm 41$ being both detectable and single transiting. These numbers are further broken down to show the contributions from the two-minute and 30-minute cadence targets. This is a vital statistic as the cadence may heavily influence the accuracy with which a single transit signal can be modelled, which in turn, influences the reliability of follow-up efforts. For a discussion of the prediction of orbital parameters based on single transit events see Osborn et al. [2016]. We note that, due to the factor of $\sqrt{N_s}$ in equation 4.7, there are no two-minute cadence
Table 4.2: Numbers of simulated planets. This table displays planets by number of sectors that observe them as well as making splits based on detectability, single transit nature and observing cadence. Targets observed by 5-11 sectors are combined as individually they comprise very small fractions of the total. Also included is the percentage of sky area observed by each number of sectors and the percentage of total simulated planets in each region. The given values are averaged from multiple simulation runs.

<table>
<thead>
<tr>
<th>Number Of Sectors Observed By (Baseline)</th>
<th>All Planets</th>
<th>Detectable Planets</th>
<th>Detectable Single Transits</th>
<th>Sky Area (%)</th>
<th>Planets (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 min</td>
<td>30 min</td>
<td>2 min</td>
<td>30 min</td>
<td></td>
</tr>
<tr>
<td>0 (0.0)</td>
<td>0 ± 0</td>
<td>0 ± 0</td>
<td>0 ± 0</td>
<td>0 ± 0</td>
<td>14.60 ± 14.29</td>
</tr>
<tr>
<td>1 (27.4)</td>
<td>101633±158</td>
<td>2546775±1445</td>
<td>535±27</td>
<td>2231±38</td>
<td>92±9 ± 317±18</td>
</tr>
<tr>
<td>2 (54.8)</td>
<td>45123±226</td>
<td>551382±352</td>
<td>289±15</td>
<td>737±21</td>
<td>15±4 ± 35±6</td>
</tr>
<tr>
<td>3 (82.2)</td>
<td>13083±73</td>
<td>89545±265</td>
<td>92±5</td>
<td>158±10</td>
<td>1±1 ± 3±1</td>
</tr>
<tr>
<td>4 (109.6)</td>
<td>3790±57</td>
<td>18629±87</td>
<td>30±6</td>
<td>35±7</td>
<td>0±0 ± 1±1</td>
</tr>
<tr>
<td>5 – 11 (137.0 – 301.4)</td>
<td>14256±87</td>
<td>35008±85</td>
<td>119±8</td>
<td>73±9</td>
<td>0±1 ± 0±0</td>
</tr>
<tr>
<td>12 (328.8)</td>
<td>397±7</td>
<td>540±18</td>
<td>4±1</td>
<td>1±1</td>
<td>0±0 ± 0±0</td>
</tr>
<tr>
<td>13 (356.2)</td>
<td>34925±208</td>
<td>38586±173</td>
<td>338±27</td>
<td>80±6</td>
<td>0±0 ± 0±0</td>
</tr>
<tr>
<td>Cadence Total</td>
<td>213236±815</td>
<td>3863513±2799</td>
<td>1407±89</td>
<td>3314±92</td>
<td>108±15 ± 356±26</td>
</tr>
<tr>
<td>Full Total</td>
<td>4076748±3614</td>
<td>4721±181</td>
<td>464±41</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

stars in the area of the sky not observed by TESS. Obviously, the numbers presented here are dependent on TESS working to specifications. Any departure from this, for example increased background due to scattered light, would affect the detection rate of planets, especially those around fainter hosts.

We also note the effect of the contamination ratio, C, on the detectability of targets. Since TESS pixels are relatively large (21 × 21 arcsec/pixel) contamination is a non-trivial issue [Stassun et al., 2018]. Figure 4.6 shows a plot of transit depth, δ, against effective transit depth, δ_{eff}, showing the impact of background contamination.

From this Fig. 4.6 it is seen that the majority of detections are found around weakly contaminated hosts meaning that δ ≈ δ_{eff}. Therefore, it is obvious that efforts to reduce, or correct for, contamination effects may lead to large increases in the number of exoplanet detections by TESS.

### 4.5 Single site follow-up

After simulating all detectable transiting planets around CTL stars cuts were made requiring \( n = 1 \) to select only single transit detections. Figure 4.7 shows a plot of all simulated planets with an overlay of these single transits on an ecliptic sky map.

Next, the dates at which TESS will observe these transits were found. The dates
Figure 4.6: Dilution of simulated planet hosts. Lower values correspond to more diluted signals.

Figure 4.7: All simulated planets coloured by number of sectors observed for. The detectable single transits for a single simulation run (490 targets) are overlaid in pink.
were calculated relative to the beginning of TESS observations and were then arbitrarily offset to begin at the start of 2019. Once the dates and times of TESS single transits are found it must be determined which events can be observed from a single ground-based site. Here the chosen site is Paranal, Chile but the process can be generally extended to any desired site with known geographical coordinates. This is achieved by predicting the times of central transit using our custom tool. Obviously, this process assumes that the epoch and period of a transit event can be accurately predicted based on a single TESS light curve. In reality there will be uncertainties in these measurements meaning that a target will most likely need to be observed for some time prior and post predicted transit time to ensure the transit is observed while allowing for uncertainties in ephemeris, this is further discussed in Sec. 4.6.3.

An analysis of the improvement of these ephemeris uncertainties is carried out by Yao et al. [2019]. This paper simulates TESS single transit signals and injects them into KELT light curves. The paper then determines how many of the signals are recoverable from the KELT light curves and finds that the number of recoverable signals is dependent on the planet radius and period with up to 80% being recoverable for the largest radius planets, falling to 18% for the smallest. For signals not recoverable in the KELT data the period uncertainty will remain around 10% [Osborn et al., 2016] but for signals that can be seen in KELT the additional data will better constrain the ephemeris.

4.6 Analysis and results

4.6.1 TESS single transit demographics

In Figs. 4.8 and 4.9 we show some of the demographics represented in the simulated planets around CTL stars. In each plot the distribution of all simulated planets is shown as well as the sub-distributions corresponding to all planets for which a transit is observed at $S/N > 7.3$ and all single transits which are observed at $S/N > 7.3$. The four plots shown in Fig. 4.8 show the planet distributions as functions of TESS magnitude, planet radius, effective transit depth and $S/N$. The distributions shown in each subplot are the result of averaging multiple simulation runs.

The distribution with TESS magnitude is seen in Fig. 4.8(a). For bright stars ($m_{TESS} \leq 12$) the number of both multiple and single transit detections increases with magnitude. This is a function of the increased numbers of stars with larger magnitude values (approximately an exponential increase). For these bright stars the number of detections and single transit detections comprise approximately the same fraction of total planets at each magnitude ($\sim 1\%$ and $\sim 0.1\%$ respectively). For dimmer stars however, this fraction falls off rapidly. This is partially due to a decrease in CTL stars with $12 < m_{TESS} < 14$, but
Figure 4.8: Simulated planet demographics. Grey shows all simulated planets, red shows all planets for which a transit is observed at $S/N > 7.3$ and blue shows all single transits with $S/N > 7.3$. 
mainly due to the increased difficulty with which transits can be detected around dimmer stars, a direct consequence of the effect of magnitude on noise shown in equation 4.5 which leads to a decrease in $S/N$.

Figure 4.8(b) shows the distribution as a function of planet radius. The overall distribution of simulated planets (in grey) is dictated by the input distribution seen in Fig. 4.4(a). At large planet radii the number of detectable transits is proportional to total number of planets with approximately 3.5% of planets being detectable and 0.5% being detectable single transits. Below $\sim 6R_\oplus$ however, this is no longer true. Though the number of simulated planets increases as the planet radius falls the number of detections remains constant, thus the detection rate falls. This is a simple consequence of the fact that smaller radius planets produce smaller transit signals which are harder to detect above to noise.

Figure 4.8(c) shows an important distribution. The pattern here shows that detectable transits peak with an effective transit depth at $\sim 0.001$, this pattern is closely matched by the single transits. The effective transit depth here is only relevant in its relation to the photometric precision of TESS. For follow-up however, transit depths that are too small will not be observable from ground based instruments which will have higher noise thresholds than space-based missions like TESS. When looking specifically at the detection of single transits, Osborn et al. [2016] suggest that signals at the level of 1 mmag are potentially discoverable. Thus, we predict that the number of detectable single transits found by TESS with effective depth $\delta_{\text{eff}} \geq 0.001$ will be $350 \pm 25$.

Finally, Fig. 4.8(d) shows the distribution of simulated planets with $S/N$. The limit of $S/N > 7.3$ for a detection used here may seem overly optimistic. From this figure it can be seen that non-negligible numbers of single transits will be detected up to much larger $S/N$ values. For example, the number of detectable single transits with $S/N \geq 20$ is $202 \pm 15$.

An additional key demographic is the distribution of orbital period shown in Fig. 4.9. Figure 4.9(a) shows the distribution of planets with orbital period. The distribution of all simulated planets (in grey) is due to the input distributions employed from Dressing and Charbonneau [2015] and Fressin et al. [2013] (seen in Fig. 4.4(b)). The number of detectable transits shows a similar distribution. There are obviously no single transits below $\sim 15$ days as these would all be observed for a minimum of 2 transits, even if only seen in a single TESS observing sector (with a baseline of 27.4 days). Figure 4.9(b) shows a close up of periods with $P \geq 25$ days. From this figure it can be seen that between 25 and 50 days approximately 70% of all detections will come from single transits with the number of single transit detections falling off slower than the total number of detections. Above 50 days this fraction approaches 100%. The number of single transits detectable by TESS with $P \geq 25$ days is found to be $315 \pm 22$. For periods longer than 50 days the
number is $44 \pm 7$. The full breakdown of detections with period is shown in Table 4.3. This table shows the number of TESS detections in 20 period bins, looking at multiple and single transit detections separately. Also shown in Table 4.3 is the subset of detections which will be found with sub-Neptune radii ($R_p \leq 4R_\oplus$).

Additionally important, especially for follow-up of candidates, is the accuracy with which planetary and orbital parameters can be predicted from a single observed transit [Osborn et al., 2016]. To better constrain these values requires more in transit observations. The number of in transit observations, $N$, is given by

$$N = \frac{T_{dur}}{\Delta T} \times n,$$

where $\Delta T$ is the cadence at which a given host is observed and $n$ is the number of observed transits. Therefore, Fig. 4.10 shows how the detectable planets break down as a function of the number of in transit observations, $N$.

The red distributions in Fig. 4.10 show the multitransit detections with the two-minute cadence targets in the darker shade of red. The blue distributions show the single transit detections with the two-minute cadence targets in the darker shade of blue. We see that there is a very obvious bi-modality in the distribution based on observing cadence. This is to be expected. Observations at two-minute cadence will have, assuming identical values of transit duration, 15 times more in transit observations; $N_{2\text{min}} \approx 15N_{30\text{min}}$. This effect is seen in both multiple and single transits but is less obvious for multitransits as
Table 4.3: Number of detectable planets found in each of 20 logarithmic period bins. Planets are separated into multitransiting planets and single transiting planets. Also shown is the number of sub-Neptune planets ($R_p \leq 4R_\oplus$). Planets with $P < 1$ day are omitted. Given values and uncertainties are the result of averaging multiple simulation runs.

<table>
<thead>
<tr>
<th>Period Bin (days)</th>
<th>Multitransits ($R_p \leq 4R_\oplus$)</th>
<th>Single Transits ($R_p \leq 4R_\oplus$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 - 1.4</td>
<td>107±8 (59±5)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>1.4 - 2.0</td>
<td>135±18 (87±13)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>2.0 - 2.8</td>
<td>414±20 (212±15)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>2.8 - 4.0</td>
<td>518±29 (216±16)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>4.0 - 5.6</td>
<td>601±26 (219±10)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>5.6 - 7.9</td>
<td>658±18 (354±15)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>7.9 - 11.2</td>
<td>574±15 (290±10)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>11.2 - 15.8</td>
<td>525±23 (234±23)</td>
<td>11±3 (3±1)</td>
</tr>
<tr>
<td>15.8 - 22.4</td>
<td>307±20 (159±8)</td>
<td>95±12 (27±4)</td>
</tr>
<tr>
<td>22.4 - 31.6</td>
<td>141±8 (77±10)</td>
<td>150±10 (47±5)</td>
</tr>
<tr>
<td>31.6 - 44.7</td>
<td>88±5 (46±4)</td>
<td>127±12 (35±5)</td>
</tr>
<tr>
<td>44.7 - 63.1</td>
<td>19±4 (10±2)</td>
<td>54±7 (12±3)</td>
</tr>
<tr>
<td>63.1 - 89.1</td>
<td>3±1 (2±1)</td>
<td>13±4 (1±1)</td>
</tr>
<tr>
<td>89.1 - 125.9</td>
<td>2±1 (1±1)</td>
<td>9±4 (1±1)</td>
</tr>
<tr>
<td>125.9 - 177.9</td>
<td>1±1 (0±0)</td>
<td>3±1 (0±1)</td>
</tr>
<tr>
<td>177.9 - 251.2</td>
<td>0±0 (0±0)</td>
<td>2±1 (0±0)</td>
</tr>
<tr>
<td>251.2 - 354.8</td>
<td>0±0 (0±0)</td>
<td>1±1 (0±0)</td>
</tr>
<tr>
<td>354.8 - 501.2</td>
<td>0±0 (0±0)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>501.2 - 707.9</td>
<td>0±0 (0±0)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>707.9 - 1000.0</td>
<td>0±0 (0±0)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>4093±200 (1965±133)</strong></td>
<td><strong>464±55 (126±22)</strong></td>
</tr>
</tbody>
</table>

91
Figure 4.10: Distribution of detectable planets with $N$, averaged from multiple simulation runs. This plot shows the distribution broken down by number of transits observed (multiple or single transits) and observing cadence (two or 30 minutes). The four distributions are stacked so the maximum height of each bar shows the total detections for each $N$.

additional transits also contribute to increasing the value of $N$. The number of detectable single transits with $N \geq 10$ is found to be $336 \pm 26$ (with two thirds coming from 30-minute cadence targets) and with $N \geq 100$ is $74 \pm 5$ (with all but one target being at two-minute cadence).

**Two-minute cadence targets**

Previous studies have found that when attempting to model orbital parameters from single transit events the sparsity of sampling within the transit is a big factor into the accuracy with which parameters can be obtained [Osborn et al., 2016]. Because of this effect, transits sampled at two-minute cadence will be better constrained. In the white-noise limit we assume the following relation,

$$\sigma_{T_{\text{dur}}} \propto \frac{1}{\sqrt{N}} \propto \Delta T.$$  \hspace{1cm} (4.9)

This means that TESS candidates observed at two-minute cadence may have periods constrained a factor of up to 15 better than those observed at 30-minute cadence. Therefore Table 4.4 shows the same information as Table 4.3 but this time only planets sampled at a
cadence of two minutes are included.

Table 4.4 shows a decrease in the number of single transits as a fraction of all detectable planets when compared to both two and 30 minute cadence targets. All cadence targets show single transits to be 10.2% of all detections whereas looking at only two minute targets shows single transits as 9.7% of detections.

4.6.2 Single site observations

When analysing the single transit detections from TESS in terms of follow-up from a single site (here chosen to be Paranal) we found 1000±83 transit events observable in a single arbitrary year (320±21 unique targets) as seen in Fig. 4.11. These events are broken down into the following categories due to how much of the transit occurs during a night. Firstly, 269±28 events where only the mid transit is observed (165±15 unique targets). Secondly, 298±29 events where either an ingress or egress is seen (202±15 unique targets). Thirdly, 315±28 events when an ingress or egress is seen along with the mid transit point (201±19 unique targets). Finally, 117±12 events where the full transit is observed (83±10 unique targets). Figure 4.11 shows the distribution of transit events in an arbitrary year (binned into weeks) coloured by how much of the transit is observed.

These events are spread through a year of observations (January to December) with noticeable changes towards the centre of the year coinciding with the increased length of a Paranal night. Of course, these observations do not yet take into account depth of the signal, magnitude of the host or moon illumination. The only requirement for an event to be listed here is that one transit is observable by TESS and that a transit event is theoretically visible from Paranal during a night and at ≥ 30° above the horizon within a single year of observations.

Since the current follow-up observables are simply dependent on a single TESS observation an additional cut was made to focus on signals that are realistically observable from ground-based sites. In the advent of the latest ground-based transit surveys such as NGTS, sub-percent transit depths are possible from single transits. Therefore, a cut was made at δ_{eff} ≥ 0.005, or a 0.5% transit depth. The remaining candidates are shown in Fig. 4.12.

As can be seen from Fig. 4.12, there are still a significant number of detections even when only looking at relatively deep transits. We also note that at long periods (P ≥ 40 days) there are no multitransit detections (0±0) but a significant number (30±5) of single transits. This confirms that study of this subset of exoplanets will require follow-up of TESS single transit detections. Additionally, even at shorter period ranges (20 ≤ P < 40 days) the brightest detections will come from single transits. Of the brightest (m_{TESS} ≤ 11) detections in this period range 18±3 out of 18±3 will come from single transits (values
Table 4.4: Same information is shown as in Table 4.3 except that only 2-minute cadence targets are included. Planets with \( P < 1 \) day are omitted. Given values and uncertainties are the result of averaging multiple simulation runs.

<table>
<thead>
<tr>
<th>Period Bin (days)</th>
<th>Multitransits ( (R_p \leq 4R_\oplus) )</th>
<th>Single Transits ( (R_p \leq 4R_\oplus) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 - 1.4</td>
<td>26±5 (24±4)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>1.4 - 2.0</td>
<td>47±8 (44±8)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>2.0 - 2.8</td>
<td>114±7 (101±7)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>2.8 - 4.0</td>
<td>115±17 (98±16)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>4.0 - 5.6</td>
<td>115±11 (92±8)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>5.6 - 7.9</td>
<td>210±12 (191±13)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>7.9 - 11.2</td>
<td>185±18 (164±19)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>11.2 - 15.8</td>
<td>167±24 (146±21)</td>
<td>3±1 (3±1)</td>
</tr>
<tr>
<td>15.8 - 22.4</td>
<td>125±13 (110±10)</td>
<td>23±5 (20±4)</td>
</tr>
<tr>
<td>22.4 - 31.6</td>
<td>73±9 (61±7)</td>
<td>38±4 (33±3)</td>
</tr>
<tr>
<td>31.6 - 44.7</td>
<td>49±3 (39±4)</td>
<td>32±6 (24±5)</td>
</tr>
<tr>
<td>44.7 - 63.1</td>
<td>12±4 (8±2)</td>
<td>10±5 (7±3)</td>
</tr>
<tr>
<td>63.1 - 89.1</td>
<td>2±0 (2±1)</td>
<td>1±2 (1±1)</td>
</tr>
<tr>
<td>89.1 - 125.9</td>
<td>1±1 (1±1)</td>
<td>1±1 (1±1)</td>
</tr>
<tr>
<td>125.9 - 177.9</td>
<td>1±1 (0±0)</td>
<td>1±1 (0±0)</td>
</tr>
<tr>
<td>177.9 - 251.2</td>
<td>0±0 (0±0)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>251.2 - 354.8</td>
<td>0±0 (0±0)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>354.8 - 501.2</td>
<td>0±0 (0±0)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>501.2 - 707.9</td>
<td>0±0 (0±0)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>707.9 - 1000.0</td>
<td>0±0 (0±0)</td>
<td>0±0 (0±0)</td>
</tr>
<tr>
<td>Total</td>
<td>1242±133 (1078±120)</td>
<td>133±24 (88±19)</td>
</tr>
</tbody>
</table>
Figure 4.11: Distribution of *TESS* single transits visible from Paranal for an arbitrary year. The data has been binned into weeks. The different colours correspond to different amount of the transit being visible. Colours blue, green, yellow and red corresponded to transits of the types 1, 2, 3 and 4 as described in the text. Additionally, the thick black line indicates the length of a Paranal night in hours. The HJD value is offset to run from 0-365 to cover 1 year of observations.

Figure 4.12: Panel a: All *TESS* detections with at least 0.5% transit depth. Additionally, this plot is restricted to $m_{TESS} \leq 13$ and $P \geq 10$ days. The red circles denote multitransit detections with blue diamonds showing single transits. Panel b: Corresponding histogram. Colours and limits are as before.
determined by averaging across multiple simulation runs). Bright targets are of particular interest as they will be most amenable to atmospheric characterisation via transmission spectroscopy with JWST.

In the work laid out so far we have assumed perfect knowledge of candidate period from a single TESS transit observation. In the following section we relax this assumption and consider realistic ephemeris uncertainties. We also now consider only the best TESS single transit detections, that is, those targets with a transit depth of at least 0.5% and a stellar magnitude less than 13 in TESS band.

### 4.6.3 Realistic follow-up

Since the targets we now consider have relatively deep transits and are hosted by relatively bright stars we assume that the uncertainty on this value is likely to be approximately 10% [Osborn et al., 2016] (this value applies to favourable single transits but here we only consider the best subset). Based on this value we made some predictions regarding the time required to follow-up a single transit detection. To effectively follow-up a candidate we assumed that it is necessary to capture the entire duration of the transit event. To calculate this time allowing for uncertainties in measured parameters we first assumed a flat error of 10% on period. Next we made the assumption that the follow-up efforts are carried out on the next transit of a given planet (a challenge, but feasible, especially for the longer period targets). Therefore, assuming the ephemeris of the TESS detection is known to sufficient accuracy, the total time required to observe a single candidate effectively is equal to its transit duration plus a buffer period at either end of the predicted ephemeris to account for the 10% uncertainty in period. Thus the total time required to effectively follow-up a candidate, $T_{\text{total}}$, is given by

$$T_{\text{total}} = 2 \left( \frac{P}{10} \right) + T_{\text{dur}}. \quad (4.10)$$

Calculating this value for the best TESS single transit detections averaged over multiple simulation runs (130±13 targets) gives a total of 22498±2867 hours or an average of 173±9 hours per candidate. This is a significant amount of time but it can easily be reduced via a few simple measures. For example, since the majority of this time is caused by the uncertainty in period, large amounts of observation time will be out of transit and show no variation. Thus, multiple targets can be observed concurrently using an on-off technique, switching between them until one begins to transit at which point that target is observed continuously. Figure 4.13 shows a plot of the distribution of observable transit events from Paranal in a year as in Fig. 4.11. This plot however, considers only the best TESS candidates as described above and uses the total time required for follow-up (Eq. 4.10) instead of the
exact transit duration.

From Fig. 4.13 it can be seen that there are many more events (due to the increased effective transit length). In fact, there are never fewer than seven events per night and up to 17 in some nights with an average of 12. The large number of simultaneous events means that the on-off technique described above will be able to monitor multiple transit events concurrently using a single telescope, reducing the total telescope time needed.

Additionally, we may utilise existing archival data (such as the publicly available photometry published by SuperWASP and NGTS) as shown by Yao et al. [2019]. Using this technique may drastically improve the ephemeris and then allow for a more accurate prediction of subsequent transits. The best targets we have selected here have the added bonus that they will most likely produce signals observable in archival data meaning they are the most likely targets to receive reduced ephemeris uncertainties via this method. Finally, this time may be further reduced via the use of spectroscopic follow-up, although we note that spectroscopy alone will not provide accurate period estimates, only allow for better timing of ground-based photometry efforts.

Beyond selecting Paranal as the single site this paper does not go into any specifics about the potential observatory that may be used for these follow-up efforts. Obviously the
observatory will have an impact on the observations due to factors including field of view, photometric precision and availability of telescope time. Not discussing these details here is a deliberate choice to keep the results of this paper non-specific. Therefore, these results can then be applied to any observatory and refined as necessary.

4.6.4 False positives

In this paper we do not conduct any analysis of false positive single transit detections in the TESS data. For a proper analysis of sources and rates of false positives see Sullivan et al. [2015]. The results of Sullivan et al. [2015] imply a number of false positives comparable to the number of genuine detections. The choice of S/N value used in this paper is chosen as an antidote to these spurious detections. This threshold is designed to result in 1 statistical false positive per $2 \times 10^5$ light curves when looking at multiply transiting planets [Bouma et al., 2017]. For single transit detections it will be harder to separate false positives, therefore Fig. 4.8(d) is of particular use, allowing us to see how increasing the required threshold will affect our single transit yields. Additionally, the best targets considered in Fig. 4.12 are all relatively deep, further reducing the risk of them being false positives. Since the time required to follow-up our candidates is significant it is important to be able to separate real signals from false so as to avoid wasted time. However, the results of this simulation show an appreciable number of detections, such that it should be feasible to select only the most convincing signals and still maintain a sufficiently large sample. Finally, Fig. 4.13 shows that multiple targets may be observed using the on-off method described in Sec. 4.6.3, further reducing the impact of false detections.

4.7 Discussion and conclusions

The TIC CTL contains approximately 3.8 million stars, after making two cuts for $m_{\text{TESS}}$ and $T_{\text{eff}}$ approximately 3.6 million stars remain. We simulated approximately 4 million stars orbiting the 3.8 million TIC CTL stars. Of these planets more than 4700 will have detectable transits during TESS observations and around 460 will transit only once. We find that 320±21 planets will have at least one transit event observable from Paranal within a 1 year period. As discussed in Sect. 4.6.2 it can be seen that there are an average of $\sim 3.0$ transit events per night throughout 1 year (however, this rate is not uniform, and peaks around the centre of the year). It is found that the number of full transits seen is $\sim 0.3$ per night (this value also peaks towards the centre of the year).

However, it is also important that the numbers presented in this report are all lower bounds. The simulation presented above only simulates planets around known CTL stars, however each full frame image will contain stars absent from the CTL, any one of which
could also display a transit during observations, thus the expected number of transits, and therefore single transits, is higher than those presented here. The CTL stars have been chosen as the best candidate hosts from the TIC [Stassun et al., 2018] but it is possible for non-CTL stars to exhibit observable, and detectable, transits. As a comparison, Sullivan et al. [2015] conduct a TESS detection simulation using a simulated population of stars rather than a known catalogue. From this data they expect over 20,000 detectable TESS transits. Assuming single transits scale as they do in this paper (~ 10% of detections) the actual number of single transits found by TESS may be closer to 2000. Additionally, this study neglects secondary transits. Obviously these are much smaller and harder to detect than primary transits but have the effect of doubling the potential events in the same time period. It is possible that the best targets may not exhibit a primary transit during observations but may host a planet whose secondary transit is observable at the photometric precision of TESS.

As well as the raw numbers of single transits that will be found with TESS it is important to consider these numbers in comparison to the multitransits found. Single transits are harder to follow-up as it is more difficult to obtain accurate transit parameters based on a single transit, although it has been shown that accurate predictions can be made, especially if circular orbits are assumed [Osborn et al., 2016]. With this in mind the most important result of this simulation is that presented in Table 4.3 where multiple and single transit detections are compared in period bins. From this table we see that for periods of \( P \gtrsim 30 \) days the number of single transits outweighs the multitransit detections, though the number of multitransit detections is still significant. For periods \( P \gtrsim 45 \) days however the numbers begin to seem more favourable to single transit follow-up. In fact, by the time we reach \( P \gtrsim 60 \) days TESS is predicted to find only 6±3 multitransit detections but 28±11 single transit detections. The difference between these two numbers is much more significant and is a strong argument that following up these single transit detections will afford opportunities to explore science that would not be possible if only multitransit TESS detections are followed-up.

One possible avenue of science that may be better explored is the discussion of long period (\( P \gtrsim 45 \) days) sub-Neptune or super-Earth planets. Recent results have repeatedly shown a dearth of planets around 2\( R_\oplus \) [Lundkvist et al., 2016; Fulton et al., 2017; Van Eylen et al., 2018], potentially a consequence of photoevaporation [Owen and Wu, 2013; Lopez and Fortney, 2013; Owen and Wu, 2017]. Therefore, TESS detections of planets within this region may be able to aid in this discussion since they will likely be detected around relatively bright hosts. To this end, Table 4.3 includes an additional break down showing numbers of simulated TESS detections which are predicted to be found with \( R_p \leq 4R_\oplus \). When looking at this particular period and radius range TESS multitransits will yield 13±4
targets and TESS single transits will yield $14 \pm 6$. The increase in a statistical sample from 13 to 27 is significant and is a good reason to spend additional effort following up TESS single transits. We also note that these detections are doubly important as, since they are around TESS candidate hosts, they are brighter than average which will lead to improved prospects for follow-up characterisation including atmospheric identification with future missions [Louie et al., 2018; Batalha et al., 2017].

The numbers presented in this paper are in excellent agreement with those presented in Barclay et al. [2018]. We find an approximate 10% increase in TESS detections which is as expected as these additional detections come from the extension of the simulation up to longer periods and the inclusion of single transiting planets as candidates. When compared with another recent estimation of TESS single transits, Villanueva et al. [2019], we find fewer detections. However, this is most likely due to the different approaches used. Whereas this paper employs a Monte Carlo simulation (similar to the method employed by Barclay et al. 2018), Villanueva et al. [2019] use a more analytic integration based approximation. In said paper the authors acknowledge that they may over estimate detections compared to a more realistic simulation as employed in Barclay et al. [2018] (and thus, in this paper) therefore we do not find the discrepancies between results to be an issue.
Chapter 5

TESS extended mission monotransit simulation

Note

The following chapter is a reproduction of the paper ‘An Examination of the Effect of the TESS Extended Mission on Southern Hemisphere Monotransits’ [Cooke et al., 2019] published in Astronomy & Astrophysics in November 2019. The simulations, results and analysis presented here were all designed and carried out by me unless stated otherwise. The writing was all done by myself. Co-author contributions consisted of discussions and improvements regarding methodology and results as well as general proof reading.

This paper provides an update to the work shown in Chapter 4 as well as extending it forward into the TESS extended mission. It explores the recurrence of TESS primary mission monotransits in the extended mission. Additionally the paper discusses period aliasing effects which are developed and discussed in Chapter 8.

Abstract

Context. NASA recently announced an extended mission for TESS. As a result it is expected that the southern ecliptic hemisphere will be re-observed approximately two years after the initial survey.

Aims. We aim to explore how TESS re-observing the southern ecliptic hemisphere will impact the number and distribution of monotransits discovered during the first year of observations. This simulation will be able to be scaled to any future TESS re-observations.

Methods. We carry out an updated simulation of TESS detections in the southern ecliptic hemisphere. This simulation includes realistic sector window functions based on the first
11 sectors of SPOC 2 min SAP lightcurves. We then extend this simulation to cover the expected Year 4 of the mission when TESS will re-observe the southern ecliptic fields. For recovered monotransits we also look at the possibility of predicting the period based on the coverage in the TESS data.

**Results.** We find an updated prediction of 339 monotransits from the TESS Year 1 southern ecliptic hemisphere, and that approximately 80% of these systems (266/339) will transit again in the Year 4 observations. The Year 4 observations will also contribute new monotransits not seen in Year 1, resulting in a total of 149 monotransits from the combined Year 1 and Year 4 data sets. We find that 75% (189/266) of recovered Year 1 monotransits will only transit once in the Year 4 data set. For these systems we will be able to constrain possible periods, but period aliasing due to the large time gap between Year 1 and Year 4 observations means that the true period will remain unknown without further spectroscopic or photometric follow-up.

### 5.1 Introduction

The Transiting Exoplanet Survey Satellite [TESS, Ricker et al., 2014] has recently completed its first year of observations of the southern ecliptic hemisphere. It will now move its observations to the north, continuing its full-sky survey. At the time of writing TESS has already successfully discovered and confirmed 29 new transiting planets [NASA Exoplanet Archive\(^1\), Akeson et al., 2013] including π Mensae c [Huang et al., 2018a; Gandolfi et al., 2018], HD 202772A b, [Wang et al., 2019b], LHS 3844 b [Vanderspek et al., 2019] and HD 219666 b, [Esposito et al., 2019].

Due to the all-sky nature of TESS, as well as its relatively short baselines for much of the sky, it produces a large yield of mono-transiting systems: planets which will exhibit only a single transit during TESS observations. The predicted yield of these monotransits has been estimated in Cooke et al. [2018] and Villanueva et al. [2019]. Recently it has been announced that the TESS mission will be extended, and as part of the extension plans it is envisaged that TESS will re-observe the southern ecliptic hemisphere in the fourth year of its mission (see NASA 2019 senior review\(^2\) and NASA response\(^3\)). Theoretical extended mission templates have been suggested by Huang et al. [2018b] and Bouma et al. [2017]. At the time of writing some of the specifics of this extended mission are unknown, but it is known that after finishing Year 2 (northern ecliptic) TESS will spend approximately one year observing the ecliptic equator before returning to the south and repeating its Year 1 observations. This is very similar to the suggested ‘hemi’ scenarios presented in both

---

\(^1\)[https://exoplanetarchive.ipac.caltech.edu/index.html]

\(^2\)2019 Senior Review Subcommittee Report

\(^3\)NASA Response to the 2019 Senior Review of Operating Missions
Huang et al. [2018b] and Bouma et al. [2017] except we now include an additional inter-
mediary year. An extra detail now currently known is that during the extended mission Postage Stamp (PS) cadence will remain at 2 mins but Full-Frame Images (FFIs) will have an improved cadence of 10 mins (c.f. 30 mins for Year 1 and Year 2 observations).

This extended mission will affect all areas of TESS science, but perhaps most notice-
able, the yield of monotransits. Therefore, an extended simulation is required to explore these effects. Since only the return to the southern ecliptic is confirmed we constrain our simulation to this hemisphere under the following assumptions:

• In Year 4, TESS will observe each southern ecliptic sector exactly three years after the sector was observed in Year 1.

• Each Year 4 sector will match up exactly with the Year 1 sectors.

• PS cadence will remain at 2 min, but FFI cadence will be improved to 10 min.

• Photometric data quality and time coverage will remain constant between Years 1 and 4.

This letter is set out in the following way. Section 5.2 outlines the simulation, particularly focusing on the improvements made to that presented in Cooke et al. [2018]. Section 5.3 outlines the key results, first giving an update to the Cooke et al. [2018] results then moving on to the Year 4 results. Finally, Section 5.4 sums up the key findings and outcomes.

5.2 Methodology

5.2.1 Stellar population

We set up our simulation using a similar method as discussed in Cooke et al. [2018]. We begin with our stellar catalogue, the TESS Input Catalogue (TIC) Candidate Target List (CTL) available from the Mikulski Archive for Space Telescopes (MAST\footnote{http://archive.stsci.edu/tess/tic_ctl.html}). This simulation employs the latest version, version 8, of the CTL cross-matched against the TIC as described by Stassun et al. [2019]. We then filter this sample using TESS-band magnitude cuts, $3.0 \leq m_{TESS} \leq 17.0$, and effective temperature cuts, $2285 \leq T_{\text{eff}} \leq 10050$. Taking only those stars in the southern ecliptic hemisphere produces a final input stellar catalogue of 4789372 potential hosts.

For each target we then predict how many TESS sectors will observe it. We simulate 13 sectors, each $24^\circ \times 96^\circ$ on the sky, and, using ecliptic coordinates, find the number of sectors overlapping each star.
5.2.2 Planetary population

With the stellar catalogue known we then simulate a planetary population around the stars in the same way as discussed in Cooke et al. [2018]. Occurrence rates as functions of radius and period are taken from Dressing and Charbonneau [2015] and Fressin et al. [2013] depending on spectral type and the hosts are randomly populated with planetary companions. Transit parameters such as depth and duration are calculated following the formalism set out in Cooke et al. [2018], which use the equations from Winn [2010b] and Barclay et al. [2018].

5.2.3 Detectability

$S/N$, and thus detectability, is a function of transit depth, stellar contamination, instrumental noise, and number of data points in transit. We calculate this value using the equations presented in Cooke et al. [2018], employing the $5^{th}$ order polynomial noise approximation given in Stassun et al. [2018].

Lastly, we determine host cadence using a priority metric,

$$\frac{\sqrt{N_s}}{\sigma_{1hr}R^3/2},$$

(5.1)

where $N_s$ is the number of sectors for which a target is observed, $\sigma_{1hr}$ is the photometric noise in an hour and $R_*$ is stellar radius. The top ranked 200,000 stars are given an observing cadence of 2 min, corresponding to the Postage Stamp (PS) targets, with the rest getting 30 min cadence during the first year of observing and 10 min cadence when TESS returns to the south hemisphere in year 4 corresponding to the Full-Frame Images (FFIs).

Although it is now known exactly which stars in the TIC have been observed at 2 min cadence in Year 1 the metric used here should produce an almost identical population. Since we do not yet know the 2 min targets for the extended mission it was decided to continue with the metric for both the primary and extended mission.

5.2.4 Sector Window Functions

To make this simulation more realistic we create and use sector window functions from the TESS data currently available for the southern sectors. In the simulation presented in Cooke et al. [2018] we assumed that each sector was two 13.7 day orbits with 6.5 hrs of downtime per orbit. In reality the true amount of usable observing time is less than this due to a variety of reasons including momentum dumps, high scattered light, and spacecraft/camera downtimes. To take this into account we create window functions for each sector and assume that only data points falling within these window functions are reliable observations.
To create these window functions we examine the publicly available 2 min lightcurves from the TESS Science Processing Operations Center pipeline [SPOC, Jenkins et al., 2016]. We download 16 representative lightcurves from each sector, one each from each of the four CCDs on each of the four cameras. For these lightcurves we then compare the Simple Aperture Photometry (SAP) and Pre-search Data Conditioned (PDC, SAP) lightcurves. At each timestamp we compare the two lightcurves, if a timestamp has not been translated over from SAP to PDC, SAP we assume this is due to some instrumental data issue and therefore that data point is unreliable. From this we create a window function of good timestamps, those points which have a corresponding PDC, SAP data point. This results in 16 window functions per sector. Comparisons of these window functions show that differences between CCDs and even between cameras are negligible within a single sector. Thus we use the camera 1, CCD 1 window function as the representative window function for that sector.

Since the data for sectors 12 and 13 are not yet publicly available we copy the mask from sector 11 into the sector 12 and 13 time slots. The coverage fraction for the TESS sectors ranges from 98% (in Sector 2) down to 69% (in Sector 3), with a mean coverage of 91% over the first 11 sectors.

Even during this first year there have been changes in data collection from Sector 1 through to Sector 11. However it is hard to estimate how these improvements will translate into the extended mission, as the spacecraft/cameras will have aged, and this may affect data quality. To produce a baseline study we therefore assume the sector window functions created from the first years data can be exactly mapped into Year 4 by simply increasing the timestamps of the window functions by 1095.75 days (3 years).

These sector window functions then help inform the simulation as to how many transits, and crucially, how many data points in transit, will be observed during the TESS Year 1 and Year 4 observations of the southern ecliptic hemisphere.

For 2 min cadence observations we take the full window function of good timestamps. For the FFI images in Year 1 and Year 4 (30 and 10 min cadence respectively) we take a slightly more involved approach. We assume that, if the data collection were perfect, we would take an FFI data point every 30 or 10 minutes for year 1 and 4. Thus the range of FFI timestamps is found by taking every 15th and 5th timestamp from the full SAP 2 min lightcurves for 30 and 10 min cadence respectively. For each of these times we cross-reference with the window function of good data. For 30 min cadence we require that each simulated timestamp has at least 14 good data points in the 30 minutes surrounding it. Any less and we assume the lack of quality data compromises the point. For the 10 min cadence we require 5 good points in the 10 minutes surrounding each simulated data point. Where this criteria is met we allow the simulated timestamp, where it is not we assume that the timestamp is lost. Thus for each cadence value, 2, 10 and 30 min, we are left with window
functions that we use for the simulation.

Next we simulate the full range of transit times for each planet. We initialise the first transit of a planet at time $T_0$ where $T_0$ is the start of Year 1, Sector 1 observations plus a random fraction of the orbital period. Additional transit times are then simulated up to the end of Sector 13 observations in Year 4. We next use these transit times to reduce our lists of timestamps, keeping only those that land within $T_{dur}/2$ of a transit centre. This process is repeated twice, once using the 2 min list for Years 1 and 4 and once using the 30 min list for Year 1 and the 10 min list for year 4.

This results in a list of $N_{in}$ timestamps per planet where $N_{in}$ is the number of in-transit observations (accounting for observing cadence), spread across $n_{tr}$ transits where $n_{tr}$ is the number of unique transit times that have surviving timestamps within $T_{dur}/2$. This value of $N_{in}$ is then used to calculate $S/N$. Only planets with $S/N \geq 7.3$ are classified as being detected.

We keep the Year 1 and Year 4 results separate. Year 4 observations have no bearing on the S/N, and therefore detectability, of planets observed in Year 1 and vice versa. A Year 1 monotransit is any planet with $n_{tr} = 1$ and $S/N \geq 7.3$ based on Year 1 data alone and the same definition is used for Year 4 monotransits from Year 4 data.

### 5.3 Results

#### 5.3.1 Year 1

Using this simulation we first update the results of Cooke et al. [2018] taking into account the more realistic sector window functions and the updated TIC CTL. As the key results we reproduce Tables 2 and 3 from Cooke et al. [2018] (see Appendix 5.A).

The total planet numbers reported here are noticeably more than expected if one simply halves the values shown in Table 2 of Cooke et al. [2018]. This is in part due to the use of an updated stellar input catalogue and therefore a larger number of simulated systems. Additionally we see more short period monotransits due to the gaps present in the actual TESS window functions.

#### 5.3.2 Year 4

The key issue we seek to address in this work is how an extended mission will affect the population of TESS monotransits. We therefore compare the Year 1 and Year 4 results. Note that both years analyse the same population of stars and planets. The most important result from this analysis is the fraction of simulated Year 1 monotransits that will have a subsequent transit observed in the Year 4 data. We find that $\sim 80\%$ (266 of 339) of the Year
Figure 5.1: The period distribution of Year 1 detectable monotransits (blue) and the fraction that have a recovered transit in the Year 4 data (red). As expected the shorter period systems show a higher fraction Year 4 recovered transits.

1 monotransits will exhibit at least one detectable transit during Year 4. The remaining 20% of systems (73 Year 1 monotransits) will remain as single transit systems even after the extended mission.

Figure 5.1 show a histogram of Year 1 monotransits as a function of period including those that will transit again in Year 4. As expected, it is evident that the short period Year 1 monotransits ($P < 20 \text{ d}$) are largely recovered again in the while the longer period transits ($P > 40 \text{ d}$) are only recovered in about half the cases.

In addition to the monotransits from Year 1, the Year 4 data also contains a population of monotransits that did not transit in the Year 1 data. This population is approximately equal in size to the number of Year 1 monotransits that are not found again in Year 4 meaning that the total number of mono-transiting systems after Year 4 is 149 with approximately half being observed to transit in Year 1 and half in Year 4.

Figure 5.2 shows the period distribution of monotransiting planets after Year 4. Comparing this with the distribution of Year 1 monotransits shown in Fig. 5.1 shows that this population is smaller and distributed more to longer periods as expected.

It is also interesting to look at how the population of recovered Year 1 monotransits varies with other system parameters. Figures 5.3, 5.4 and 5.5 show the Year 1 monotransits as well as those observed to transit again in Year 4 as functions of planetary radius, stellar...
Figure 5.2: Period distribution of systems that remain as monotransits once the Year 1 and Year 4 data have been obtained. Approximately half of these systems are found in the Year 1 data, and half are in the Year 4 data. Note there are far fewer than the original monotransit yield from the Year 1 alone, and the distribution is shifted towards longer period systems.

These three figures suggest that the recovered population of Year 1 monotransits is broadly independent of planetary radius, stellar magnitude or stellar effective temperature. Logically this makes sense. Targets will be re-observed for the same amount of time in Year 4 as in Year 1. Thus, if detected in Year 1, should be detected again in Year 4 assuming a transit occurs again. Therefore the only meaningful parameter for recovery should be orbital period.

5.3.3 Period alias

Although ~80% (266 of 339) of Year 1 monotransits will been seen to transit in Year 4 (Section 5.3.2), approximately 75% of these (198/266) will only transit once in the Year 4 data. For these systems we will have only two transits, which means rather than being able to determine a single period, there will be a discrete set of allowed periods. If we assume no significant transit timing variations, the period of these systems is given by \( P = \frac{T_{\text{diff}}}{n} \) where \( T_{\text{diff}} \) is the difference between the Year 1 and Year 4 transit times and \( n \) is an integer. For this simulation the average value of \( T_{\text{diff}} \) is 3 years.

Obviously, for real data sets the true value of \( n \), and thus \( P \), is unknown so next
Figure 5.3: The planetary radius distribution of Year 1 detectable monotransits (blue) and the fraction that have a recovered transit in the Year 4 data (red).

Figure 5.4: The stellar magnitude distribution of Year 1 detectable monotransits (blue) and the fraction that have a recovered transit in the Year 4 data (red).
we look at the possibility of using the TESS data to constrain these values. To begin we assume that the smallest possible period for one of these systems is 10 days, any shorter than this and a second transit would have been seen within a single sector, even allowing for gaps in coverage. From this we then create a list of \( n \) values for each system with \( n \) running from 1 to \( n_{\text{max}} \), where \( n_{\text{max}} = T_{\text{diff}}/10 \) (rounded up to the nearest integer). For each value of \( n \) we generate a list of periods and the corresponding times at which the system would transit, including values before and after the observed transits. These lists are then compared against the sector window functions (allowing for which sectors observe each target) and if any transits fall into regions that have TESS coverage (in Year 1 or 4) we rule out that value of \( n \) since a third transit would have been seen. In this way we generate a list of allowed \( n \) values (and hence periods) for each system that we can compare against the true period. In order to compare systems with different periods, we define the true period index \( I \) to be

\[
I = n_{\text{max}} - n_{\text{true}},
\]

where \( n_{\text{true}} \) is the actual \( n \) value for a given system. Thus if the true period is the shortest possible period, the period index will be \( I = 0 \). A system with a period index of \( I = 5 \) would, for example, mean that the true period of this system is the 6\(^{\text{th}}\) shortest possible period.
Figure 5.6: Distribution of true period index, $I$. It can be seen that the distribution is maximal at $I = 0$ (i.e. shortest possible period is the true period) but there are a significant number of systems for which $I > 0$, and is non-negligible even up to large $I$ values.

From these results we test whether the number and range of alias periods allows for the recovery of the true period. Based on the geometric probability of transit, one might assume that the true period is usually the shortest allowed period. However we find this is very often not the case. In Figure 5.6 we show the distribution of the true period index $I$ which is the period index that corresponds to the true period.

5.3.4 Extended mission timeline

Since some of the specifics of the extended mission are still unknown we briefly discuss here the impact that may be expected if TESS were to return to the southern ecliptic hemisphere after two years instead of three years. First and foremost we would not expect a sizeable change in the number of recovered Year 1 monotransits. Each system would still be re-observed for the same amount of time so there is no reason that planets should exhibit significantly more or less additional transits after two or three years.

Part of this simulation that would be affected is the period alias discussion presented in Section 5.3.3. If the southern hemisphere Sectors were re-observed after only two years the set of discrete possible periods for those Year 1 monotransits which show a second transit would be reduced. The average value of $T_{\text{diff}}$ would now be 2 years, or 730.5 days, a fraction of 2/3 as long as before. This then results in approximately 2/3 as many allowed
periods for each system.

Some suggestions for the extended TESS mission include a number of TESS observing sectors placed along the ecliptic plane. Assuming these sectors to be exactly along the equator they will overlap the bottom 6° of the southern hemisphere sectors. This will be \( \frac{1}{16} \) of the southern observations (if the whole ecliptic plane is observed). Without knowing when this is to occur it is hard to estimate its effect but qualitatively we would expect the recovered monotransits to increase by a similar fraction.

5.4 Conclusions

Using the current understanding of the TESS stellar population, the planetary occurrence rates from Kepler, a realistic model of TESS photometric capabilities, and a realistic window function, we expect approximately 339 monotransit targets in the Year 1 data from TESS. Of these targets, 266 will be seen to transit again in the Year 4 monitoring (\( \sim 80\% \)), while 73 will remain as monotransit systems (\( \sim 20\% \)). When this number is combined with our predictions of monotransits found in Year 4 only, we end the extended mission with 149 monotransiting systems with signals above the detection threshold.

The majority of Year 1 monotransits that are recovered in Year 4 will only show one transit in the Year 4 data (198/266, or approximately 75%), leaving a wide range of potential periods. Furthermore these are preferentially the longer period systems, which may be seen as more scientifically valuable as they cover planetary equilibrium temperatures that are interesting for internal structure studies [e.g. Thorngren and Fortney, 2018] or even for habitability studies [e.g. Schwieterman et al., 2019]. To discover the true period either a spectroscopic orbit will be required, or the discrete set of periods can be used to predict transit times for photometric follow-up observations. Which of these avenues is best will depend on a variety of factors including the magnitude of the host star, the transit depth, and the availability of spectroscopic or photometric resources (Cooke et al., in prep).

Both techniques are being used successfully for recovery of TESS monotransit periods; for spectroscopy see Lendl et al. (in prep.) and for photometry see Gill et al. (in prep.). The number of monotransit candidates found so far in TESS is in broad agreement with the numbers presented here. In depth searches for these candidates is underway, as evidenced by these two in prep. papers.

5.A Year 1 updated results
Table 5.1: Numbers of simulated planets. This table displays planets by number of sectors that observe them as well as making splits based on detectability, single transit nature and observing cadence. Targets observed by 5-11 sectors are combined as individually they comprise very small fractions of the total. Also included is the percentage of sky area observed by each number of sectors and the percentage of total simulated planets in each region.

<table>
<thead>
<tr>
<th>Number Of Sectors</th>
<th>All Planets</th>
<th>Detectable Planets</th>
<th>Detectable Single Transits</th>
<th>Sky Area (%)</th>
<th>Planets (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PS</td>
<td>FFI</td>
<td>PS</td>
<td>FFI</td>
<td>PS</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>779971</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>45873</td>
<td>4080171</td>
<td>208</td>
<td>1572</td>
<td>37</td>
</tr>
<tr>
<td>2</td>
<td>23675</td>
<td>1040112</td>
<td>150</td>
<td>542</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>7289</td>
<td>187330</td>
<td>23</td>
<td>119</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2264</td>
<td>40121</td>
<td>11</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>5 – 11</td>
<td>8262</td>
<td>83112</td>
<td>75</td>
<td>55</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>289</td>
<td>1478</td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>21148</td>
<td>111866</td>
<td>142</td>
<td>117</td>
<td>0</td>
</tr>
<tr>
<td>Cadence Total</td>
<td>108800</td>
<td>6324161</td>
<td>615</td>
<td>2437</td>
<td>53</td>
</tr>
<tr>
<td>Full Total</td>
<td>6432961</td>
<td>3052</td>
<td>339</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5.2: Number of detectable planets found in each of 20 logarithmic period bins. Planets are separated into multitransiting planets and single transiting planets. Also shown is the number of sub-Neptune planets ($R_p \leq 4R_\oplus$). Planets with $P < 1$ day are omitted.

<table>
<thead>
<tr>
<th>Period Bin (days)</th>
<th>Multitransits ($R_p \leq 4R_\oplus$)</th>
<th>Single Transits ($R_p \leq 4R_\oplus$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 - 1.4</td>
<td>88 (37)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>1.4 - 2.0</td>
<td>88 (49)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>2.0 - 2.8</td>
<td>274 (95)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>2.8 - 4.0</td>
<td>377 (103)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>4.0 - 5.6</td>
<td>432 (104)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>5.6 - 7.9</td>
<td>428 (160)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>7.9 - 11.2</td>
<td>349 (108)</td>
<td>13 (1)</td>
</tr>
<tr>
<td>11.2 - 15.8</td>
<td>281 (85)</td>
<td>28 (4)</td>
</tr>
<tr>
<td>15.8 - 22.4</td>
<td>150 (60)</td>
<td>74 (11)</td>
</tr>
<tr>
<td>22.4 - 31.6</td>
<td>51 (23)</td>
<td>95 (21)</td>
</tr>
<tr>
<td>31.6 - 44.7</td>
<td>36 (14)</td>
<td>85 (15)</td>
</tr>
<tr>
<td>44.7 - 63.1</td>
<td>13 (4)</td>
<td>22 (3)</td>
</tr>
<tr>
<td>63.1 - 89.1</td>
<td>4 (1)</td>
<td>9 (2)</td>
</tr>
<tr>
<td>89.1 - 125.9</td>
<td>1 (0)</td>
<td>8 (2)</td>
</tr>
<tr>
<td>125.9 - 177.8</td>
<td>1 (1)</td>
<td>4 (0)</td>
</tr>
<tr>
<td>177.8 - 251.2</td>
<td>0 (0)</td>
<td>1 (0)</td>
</tr>
<tr>
<td>251.2 - 354.8</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>354.8 - 501.2</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>501.2 - 707.9</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>707.9 - 1000.0</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>2573 (844)</td>
<td>339 (59)</td>
</tr>
</tbody>
</table>
Chapter 6

TESS & CHEOPS

Note

The following chapter is a reproduction of the paper ‘CHEOPS Observations of TESS Primary Mission Monotransits’ [Cooke et al., 2020b] published in Monthly Notices of the Royal Astronomical Society in March 2020. The simulations, results and analysis presented here were all designed and carried out by me with the exception of the tool used to generate the data presented in Figure 6.1 which is referenced where relevant. The writing was all done by myself. Co-author contributions consisted of discussions and improvements regarding methodology and results as well as general proof reading.

This paper builds on the results of Chapter 5 and explores the feasibility of using a separate photometric mission, CHEOPS, to follow-up these systems. The paper highlights the need for period information and sets up the work shown in Chapter 8.

Abstract

We set out to look at the overlap between CHEOPS sky coverage and TESS primary mission monotransits to determine what fraction of TESS monotransits may be observed by CHEOPS. We carry out a simulation of TESS transits based on the stellar population in TICv8 in the primary TESS mission. We then select the monotransiting candidates and determine their CHEOPS observing potential. We find that TESS will discover approximately 433 monotransits during its primary mission. Using a baseline observing efficiency of 40% we then find that 387 of these (~89%) will be observable by CHEOPS with an average observing time of ~ 60 days per year. Based on the individual observing times and orbital periods of each system we predict that CHEOPS could observe additional transits for approximately 302 of the 433 TESS primary mission monotransits (~70%). Given that
CHEOPS will require some estimate of period before observing a target we estimate that up to 250 (~58%) TESS primary mission monotransits could have solved periods prior to CHEOPS observations using a combination of photometry and spectroscopy.

6.1 Introduction

The CHaracterising ExOPlanets Satellite [CHEOPS, Broeg et al., 2013; Fortier et al., 2014] is an ESA mission dedicated to performing ultra-high precision photometry on known transiting planetary systems. CHEOPS was launched on 18th December 2019 and has a nominal mission lifetime of 3.5 years [Rando et al., 2019]. CHEOPS will be able to very precisely measure radii for the planets it observes due to its photometric precision and short observing cadence (1 minute or better). Additionally, if observing systems with uncertain periods (or those with only bounding limits on the period), CHEOPS will be able to help confirm periods by capturing additional transits or ruling out period aliases.

These characteristics make CHEOPS an attractive tool for better characterising monotransiting systems discovered as part of the TESS primary mission. These are targets identified as transiting planets by the TESS primary mission but with only one observed transit. Without the ability to fold multiple transits these systems require significant effort to properly characterise but are of interest since they generally have longer periods than other TESS detections (those that exhibit multiple transits) [Cooke et al., 2018]. Better characterisation of these systems is vital to help develop our understanding of how planetary characteristics vary with orbital separation [Nottale et al., 2004; Jones et al., 2004; Alibert et al., 2013]. Moving further from the host additionally favours the detection of potential habitable zone planets, especially around solar type stars whose habitable zone periods are on the order of hundreds of days [Kasting et al., 1993; Kopparapu et al., 2013].

The Transiting Exoplanet Survey Satellite [TESS, Ricker et al., 2015] is over a year into its primary mission having already moved into the northern ecliptic hemisphere. As part of its exoplanet yield TESS is expected to discover hundreds of monotransiting systems [Cooke et al., 2018; Villanueva et al., 2019; Cooke et al., 2019]. These systems will be known to host transiting exoplanets so will be valid targets for CHEOPS but will have no known period. These systems require time and effort to characterise but some systems have been successfully recovered with more on the way [Gill et al., 2020a; Lendl et al., 2020b; Gill et al., 2020b].

CHEOPS will not be able to directly observe these systems based on a single TESS transit alone since this would require constant observation until the planet transits again. However, if a prior on the period was available, or a small number of period aliases were found, CHEOPS could instead target specific times and, using its high cadence and preci-
sion, vastly improve the transit, and thus planet, characteristics. To this end it is of interest to know what fraction of TESS primary mission monotransits CHEOPS could observe and for which of these period estimates could be found that would justify the use of CHEOPS time.

We set out our paper in the following way. Section 6.2 describes the simulation population and the TESS and CHEOPS observations. Section 6.3 shows the simulation results for both the TESS and CHEOPS observations. Section 6.4 shows our attempts to estimate periods for the TESS monotransits and Section 6.5 gives our conclusions.

6.2 Methodology

6.2.1 Simulation population

We produce our stellar and planetary populations as in Cooke et al. [2018]. Our stellar populations we take as the TESS Input Catalogue (TIC) Candidate Target List (CTL) version 8 [Stassun et al., 2019] available from the Mikulski Archive for Space Telescopes (MAST1). The sample is filtered by magnitude in the TESS-band, $m_{\text{TESS}}$, and effective temperature, $T_{\text{eff}}$, using $3.0 \leq m_{\text{TESS}} \leq 17.0$, and $2285 \leq T_{\text{eff}} \leq 10050$ K.

Planets are generated around these stars based on occurrence rates as functions of radius and period from Dressing and Charbonneau [2015] (M-stars) and Fressin et al. [2013] (AFGK-stars). We determine transit parameters as in Cooke et al. [2018] using equations from Winn [2010b] and Barclay et al. [2018].

6.2.2 Detectability

For TESS detectability we first determine how many TESS sectors will observe each TICv8 target. From there we use the transit parameters with a noise approximation to calculate a Signal-to-Noise ratio, $S/N$. We take our noise approximation from Stassun et al. [2018] who use a 5th order polynomial fit. We take into account the arguments of Barclay et al. [2018] and require $S/N \geq 10.0$ for a detectable observation. Finally we must determine which stars will be observed at 2 min cadence and which will only receive 30 min. We determine this using the priority metric

$$\sqrt{N_s}$$

$$\sigma_{1\text{hr}} R_\star^{3/2}$$

where $N_s$ is the number of sectors for which a target is observed, $\sigma_{1\text{hr}}$ is the photometric noise in an hour and $R_\star$ is stellar radius. The 200,000 top priority targets receive

1http://archive.stsci.edu/tess/tic_ctl.html
2 min cadence observations. As in Cooke et al. [2019] we continue to use this metric as it is almost identical to the known 2 min sample but remains unbiased into the northern ecliptic where the exact 2 min targets are not yet known for all sectors.

### 6.2.3 TESS observations

To simulate TESS observations of our stellar population we use the same sector window functions method described in Cooke et al. [2018]. This method involves taking the timestamps from actual TESS data for each available sector, stripping out data points with systematics and only simulating observations at the remaining times. This produces a more realistic observing strategy than was used in Cooke et al. [2018]. The improvement between the method employed here and the Cooke et al. [2019] implementation of this method is that data for more TESS sectors is available. We now have timestamps for sectors 1-15. We then extrapolate this data up to sector 26 by replicating the sector 15 timestamps separated by the average inter-sector gap to complete the primary mission. As an additional improvement we offset 6 northern sectors (sectors 14, 15, 16, 24, 25 and 26) to follow the procedures that TESS is employing to mitigate scattered light. Once the simulation is complete we select all those planets with one detectable transit during the TESS primary mission as our monotransit sample.

### 6.2.4 CHEOPS observations

We then looked at whether these monotransits will be observable by CHEOPS. CHEOPS orbital period is 98.6 minutes with an observing cadence of 1 minute or better. The location of a target on the sky will lead to a different amount of interruption per orbit. Figure 6.1 shows three sky plots of CHEOPS coverage showing the total time that each part of the sky can be observed for in one year\(^2\). The difference between the three plots is the minimum amount of observing time required per orbit; 19, 59 and 98 minutes (roughly corresponding to 20%, 60% and 99% of the orbital period). The greater the required observing time per orbit, the smaller the total sky area that can be observed. However, the areas of sky that are observable can be observed for longer per orbit. This leads to longer observing times, but spread across a smaller region of sky.

The specific amount of observing time required per CHEOPS orbit is difficult to estimate due to unknown factors such as the specific variability of the host star, when in the orbital phase CHEOPS observations of a target will begin and what fraction of a transit may be caught. Because of these complications we have elected to require a baseline of

\(^2\)Available from [https://www.cosmos.esa.int/web/cheops-guest-observers-programme/ao-1](https://www.cosmos.esa.int/web/cheops-guest-observers-programme/ao-1)
Figure 6.1: CHEOPS sky coverage for three different values of minimum observing time per orbit. The colours denote total observing time in days (across one year) and the areas in grey are the regions that never reach the required coverage per orbit. Also shown are contours at 20, 40, 60 and 80 days (where applicable).
40% observing efficiency for each TESS primary mission monotransit. That is, if a TESS monotransit falls on to a region of sky for which CHEOPS could observe with an efficiency of \( \geq 40\% \) we consider that monotransit to be observable with CHEOPS. 40% efficiency gives \( \sim 40 \) minutes of observation per orbit and was chosen as this matches well with the average length of an ingress/egress of the simulated monotransits. We do note that CHEOPS observability exhibits a slow continuous change between orbits (see Appendix C of Kuntzer [2013]) but, providing we select targets allowing for this it should only increase the number of available targets.

### 6.3 Results

#### 6.3.1 TESS

The results of the TESS part of the simulation are shown here. From the TESS primary mission we predict 433 monotransits. That is, 433 of our simulated planets will be observed to transit once during TESS primary mission observations with an \( S/N \geq 10.0 \). Figure 6.2 shows a plot of all stars in TICv8 coloured by number of sectors with which TESS will observe them. The monotransiting planets are overlaid in blue. This value is comparable to the equivalent values presented in Cooke et al. [2018] and Cooke et al. [2019] allowing for the reduction in TESS coverage in the northern ecliptic hemisphere and an increased \( S/N \).
Figure 6.3: TESS monotransits (blue) plotted over the CHEOPS sky coverage map for a minimum of 39 minutes of observation per orbit (∼ 40% efficiency). The colour shows the total CHEOPS observing time in days across one year.

6.3.2 CHEOPS

Using the specified observation efficiency of 40% as above we plot the 433 TESS primary mission monotransits over the corresponding CHEOPS sky coverage plot.

It can be seen from this plot that, based on the 39 minutes sky coverage, only a small fraction of monotransits fall on uncovered sky. In fact, only 46 out of the 433 monotransits fall outside of the required CHEOPS coverage. This leaves 387 TESS primary mission monotransits observable by CHEOPS, allowing for our observing criteria (∼ 89%). An additional point that can be seen in this plot is the lack of TESS monotransits around the regions of sky which CHEOPS will observe for the longest. This is partially because these areas are along the ecliptic equator which TESS will not observe during its primary mission and partially because the shifted TESS northern sectors happen to coincide with this region of sky.

The full distribution of CHEOPS observing time for the 387 observable TESS monotransits is shown in Figure 6.4.

The distribution of the observing time for the 387 observable targets is seen to peak around 60 days matching the distribution seen in Figure 6.3. This is a result of the majority of TESS monotransits being found in regions of sky with 1 TESS sector of observations, that is, between 6 and 30 degrees in latitude of the equator. These are the areas of sky which
Figure 6.4: Number of days per year for which each TESS primary mission monotransit could be observed by CHEOPS with at least 40% efficiency. We show all 387 targets observable by CHEOPS.

match up with the 60 day contour in Figure 6.3. The sharp drop off above 60 days is again the result of no monotransits being found along the ecliptic equator, combined with the placement of the shifted TESS sectors. The drop off towards less CHEOPS observation is a result of the decreasing number of monotransits that are discovered further from the ecliptic equator, where the sky has multiple TESS sectors of coverage leading to more multitransit detections. This corresponds to the regions where CHEOPS has less coverage as well.

An interesting aspect of this distribution to consider is which regions of the monotransit distribution can be recovered by CHEOPS. Figure 6.5 show the period distribution of all 433 TESS primary mission monotransits as well as the subset which CHEOPS could observe.

From this plot we see that the TESS monotransits period distribution is similar to those seen in Cooke et al. [2018] and Cooke et al. [2019]. The distribution drops off with period due to the underlying occurrence rates used in this simulation, however we still find over 200 detections with \( P \geq 25 \) days and \( \sim 30 \) detections at \( P \geq 50 \) days. The CHEOPS observable subset follows the TESS distribution shape closely but falls off faster with period. For the range \( P < 25 \) days CHEOPS can observe 96% of TESS monotransits. This fraction is 88% for \( 25 \leq P < 50 \) days and 87% for \( 50 \leq P < 75 \) days. This then falls to an average of \( \sim 50% \) recovery for \( P \geq 75 \) days. The drop off towards longer period systems
is explained by the relation between TESS and CHEOPS sky coverage. Longer period monotransits are preferentially found in areas of the sky covered by multiple TESS sectors which are towards the ecliptic poles. However, these are also the areas where CHEOPS has the least coverage (see Figure 6.1).

Figure 6.6 shows the corresponding planetary radius distribution.

As with period the CHEOPS population generally follows the TESS one closely. We see that the fraction of monotransits CHEOPS can recover is independent of planetary radius. This makes sense when comparing the ultra-high photometric precision of CHEOPS with that of TESS. Due to the increased precision the only factor preventing CHEOPS from picking up targets are their locations on this sky and unlike with period, there is no location dependence on planetary radius.

6.3.3 CHEOPS timing

There is still an additional effect to be considered however. Though we have shown which monotransits CHEOPS could observe in theory (those with non-zero sky coverage assuming a required efficiency of 40%), it is not yet known if the times that they could be observed at will match up with additional transits. However, based on the periods of the monotransits and the sky coverage CHEOPS will achieve we can predict the number of monotransits
Figure 6.6: Planetary radius distribution of monotransits. All TESS primary mission monotransits (433) are in blue and the subset that CHEOPS could observe based on our criteria (387) are in red.

that CHEOPS will be able to observe at the correct time to catch an additional transit. The precision of this forecast will be based on how well a period is known prior to using CHEOPS and is therefore an additional argument for photometric or spectroscopic data before CHEOPS is used.

We must first find the number of transits that will occur in a year, $n$, for each TESS monotransit. This is simply found by dividing a year by the orbital period. We then need the fraction of time CHEOPS could observe the target for within a year, $f$, found from the observing time (shown in Figure 6.4). The probability that CHEOPS will then be able to observe at least one additional transit, $F$, is given by the following equation

$$ F = 1 - (1 - f)^n. $$

Using this equation we find that CHEOPS would observe an additional transit for approximately 302 of the 387 monotransits that CHEOPS could observe.

Figures 6.7(a) and 6.7(b) show how this probability changes as a function of orbital period and planetary radius respectively.

It can be seen from these plots that orbital period is the key parameter that determines whether another transit can be observed. We see a general decrease in $F$ with increasing period, as expected from equation 6.2, and no real effect from changing radius.
Figure 6.7: Probability that CHEOPS will observe an additional transit of a TESS primary mission monotransit as a function of orbital period and planetary radius.

In terms of actual probabilities, for periods $\leq 50$ days the majority of monotransits have a greater than 80% chance of having an additional transit observed with only a handful having probabilities less than 60%. At longer periods the distribution becomes more sparse with many monotransits lying around 40-50%.

### 6.3.4 More conservative CHEOPS stability

CHEOPS’ orbit is foreseen to be very stable. Based on this stability it is valid to assume that CHEOPS is capable of detecting a transit from even a temporary drop of flux even if a whole ingress or egress is not seen or if the ingress or egress happens across multiple satellite orbits. However, CHEOPS is yet to return data and it is possible that its instrumental stability may be less than predicted. In this case a more rigorous criteria for a transit detection would be required and we make some comments to this effect here.

If the orbit-to-orbit stability is reduced we would need to be able to detect a transit from a single orbits’ worth of data alone. Therefore we now define a monotransit as observable by CHEOPS only if the observing time per orbit is sufficient to allow for the observation of a full ingress or egress (we assume this is sufficient for a transit detection). We also require that the per orbit observation time is sufficient that an ingress/egress cannot hide in the gap within a single CHEOPS period and be missed. Our minimum required observing time per orbit is then the larger of these two constraints. Based on this criteria there are then 2 cases where CHEOPS will not be able to sufficiently observe a target. First, ingress/egress duration is longer than a single CHEOPS period. Second, in the required coverage map the target coordinates fall in an uncovered region (marked grey in Figure 6.1).
Based on this interpolation we find that 176 TESS primary mission monotransits are observable for the required time per orbit by CHEOPS allowing for our more rigorous observing criteria. This outcome supports the fact that even should CHEOPS stability be less than expected a significant number of TESS primary mission monotransits could still be observed.

### 6.4 Period estimation

As has been mentioned above CHEOPS will not be used as a blind follow-up for monotransits. In other words, CHEOPS will not stare at a monotransiting target waiting for a second transit to confirm its ephemeris, this would be an inefficient use of limited CHEOPS observing time. Therefore we need some estimate of period for these systems before CHEOPS will observe them. Predicting this period relies on obtaining additional observations of the systems in question, either photometric or spectroscopic.

As has been shown in Cooke et al. [2019] ~ 80% of TESS primary mission monotransits will be seen to transit again during the TESS extended mission with ~ 75% transiting only once more. An additional transit during this mission means that we can now constrain the period into a discrete set of period aliases, which include the true period. In what follows we require the data from the TESS extended mission, therefore we limit ourselves to the southern ecliptic hemisphere as this will be re-observed earlier by TESS (July 2020 - June 2021) meaning that observing targets after the extended mission will still be feasible within the nominal CHEOPS mission lifetime (3.5 years). Additionally we focus on those systems that exhibit a single transit during each of the TESS primary and extended missions. This results in 132 planets.

Continuing our simulated TESS observations into the extended mission using the same procedure as Cooke et al. [2019] we find that each system will have an average of 35 aliases when accounting for the TESS coverage and the separation between transits (for full details of this period alias simulation see Cooke et al. 2020, in prep.). 35 aliases is still too many for CHEOPS to reasonably target since most of these will be false and thus reveal no transit. Therefore we require additional observations to rule out these aliases. We simulate a suite of photometric and spectroscopic observations using the Next Generation Transit Survey at Paranal [NGTS, Wheatley et al., 2018] and CORALIE on the Euler 1.2 m telescope [Queloz et al., 2000b] respectively.

For photometric observations we simulate a stare campaign observing the target for a set amount of days using all night time hours as employed by Gill et al. [2020a]. For spectroscopy we simulate one data point taken every 3 days for a set amount of time. After each night of observing we compare the simulated data and coverage with the set of period
aliases based on the TESS transits. For photometry we look at each alias and compare its predicted transit times with the NGTS coverage to rule out aliases. For spectroscopy we fold our data on each alias and see if the phase curve is in line with the predicted shape based on the planet radius and spectrograph noise level (for more details of this period alias simulation see Cooke et al. 2020, in prep.).

The key result of this simulation for influencing CHEOPS follow-up is for what fraction of TESS primary mission monotransits can we determine the period before CHEOPS must observe it. Figure 6.8 shows this distribution as a function of additional observing time.

We show that, depending on the amount of photometric and/or spectroscopic time used, up to 94% of monotransits can have solved periods. Running the simulation for more time shows little change as the fraction of solved systems has plateaued. Using a combination of photometry and spectroscopy we show that 50% of systems can have solved periods after ~ 1 month of additional observations per target. This means that, depending on the amount of photometric and spectroscopic time available, we will be able to target up to ~ 125 TESS primary mission monotransits using CHEOPS, having already determined their
periods. Extrapolating this into the northern ecliptic hemisphere (TESS will re-observe this
hemisphere June 2021 - September 2022, still before the end of CHEOPS’ lifetime) gives
the opportunity for CHEOPS to observe \( \sim 250 \) TESS primary mission monotransits.

## 6.5 Conclusions

We have shown that TESS will discover approximately 433 monotransits during its primary
mission, including south and north ecliptic hemispheres. These systems are distributed
across the sky as shown in Figure 6.2 with period and planetary radius distributions as
shown in Figures 6.5 and 6.6. For more details of the TESS monotransit population see
Cooke et al. [2018] and Cooke et al. [2019]. We have then explored the feasibility of using
CHEOPS to re-observe these monotransiting systems. Using a baseline observing efficiency
of 40% we have shown that 387 of the monotransits could theoretically be observed by
CHEOPS (\( \sim 89\% \)). We have also then shown how this observable distribution depends on
orbital period and planetary radius finding that CHEOPS observations will slightly favour
shorter periods but will not be affected by radius. Of these 387 monotransits we show that
CHEOPS observations will be able to coincide with future transits for 302 (\( \sim 78\% \)) of these
systems.

For CHEOPS to realistically spend time following-up these systems we will require
some knowledge of the period of the monotransit and we have outlined a method in which
this may be obtained. We show that, by combining additional photometry from the TESS
extended mission and NGTS with spectroscopy from CORALIE, up to 250 TESS primary
mission monotransits will be able to have solved periods prior to CHEOPS’ potential obser-
vations of them (for full details of this analysis see Cooke et al. 2020, in prep.). This would
allow an efficient use of CHEOPS observing time to target transits and constrain the char-
acteristics of these interesting long-period systems. The data files and some of the relevant
codes for this project are available at https://github.com/BenCooke95/CHEOPS-TESS.
Chapter 7

SpecPhot

Note

The following chapter is a reproduction of the paper ‘SpecPhot: a Comparison of Spectroscopic and Photometric Exoplanet Follow-up Methods’ [Cooke and Pollacco, 2020] published in Monthly Notices of the Royal Astronomical Society in May 2020. The simulations, results and analysis presented here were all designed and carried out by me unless otherwise stated. The writing was all done by myself. Co-author contributions consisted of discussions and improvements regarding methodology and results as well as general proof reading.

This paper discusses the effectiveness of both photometry and spectroscopy when attempting to follow-up exoplanets. Due to the range of parameters space that must be considered this paper makes some underlying assumptions and approaches the problem in a general manner. The results inform the procedure used in Chapter 8 and help support the methods used by the NGTS monotransit working group, discussed in Chapter 9.

Abstract

We set out a simulation to explore the follow-up of exoplanet candidates. We look at comparing photometric (transit method) and spectroscopic (Doppler shift method) techniques using three instruments: NGTS, HARPS and CORALIE. We take into account precision of follow-up and required observing time in attempt to rank each method for a given set of planetary system parameters. The methods are assessed on two criteria, SNR of the detection and follow-up time before characterisation. We find that different follow-up techniques are preferred for different regions of parameter space. For SNR we find that the ratio of spectroscopic to photometric SNR for a given system goes like $R_p/P^{1/4}$. For follow-up time
we find that photometry is favoured for the shortest period systems (< 10 d) as well as systems with small planet radii. Spectroscopy is then preferred for systems with larger radius, and thus more massive, planets (given our assumed mass-radius relationship). Finally, we attempt to account for availability of telescopes and weight the two methods accordingly.

7.1 Introduction

Exoplanets are being discovered in larger numbers than ever before. Many of these new planets are being detected from the large data sets produced by long photometric campaigns. Thus it is more important than ever to have an efficient plan to best utilise follow-up efforts. As an added impact of these surveys we are beginning to reach into longer period parameter space. Unfortunately, due to the functionality of these photometric surveys, longer period systems have an increased chance of only exhibiting a single transit within a survey leading to significant yields of monotransits [e.g. Cooke et al., 2018; Villanueva et al., 2019]. Therefore these systems have only estimated periods based the shape of the discovery transit [Osborn et al., 2016]. To properly characterise these, and other, systems we need to follow them up. Generally we can follow-up exoplanets using either photometric or spectroscopic techniques.

Photometry as an exoplanet detection method relies on the transit method. That is, the reduction in flux from a host star as an exoplanet passes across an observer’s line of sight. Assuming edge on systems the size of the dip then gives the radius ratio of the star-planet system. Notable photometry missions include the Wide Angle Search for Planets [WASP, Pollacco et al., 2006], Kepler/K2 [Borucki et al., 2010; Howell et al., 2014], the Next-Generation Transit Survey [NGTS, Wheatley et al., 2018] and the Transiting Exoplanet Survey Satellite [TESS, Ricker et al., 2015].

The first detection of a transiting exoplanet was HD 209458b in 1999 [Charbonneau et al., 2000; Henry et al., 2000] (previously discovered using radial velocities) and the first exoplanet discovered via this method was OGLE-TR-56b in 2002 [Udalski et al., 2002]. Photometry can be carried out from space or ground-based with both methods having pros and cons. This paper focuses on ground-based photometric follow-up. Generally space based photometry can detect smaller planetary signals due to reduced instrumental noise, a lack of atmospheric corrections and more consistent sampling but ground-based photometry is improving constantly. The smallest signal discovered from the ground to date is 0.13% [West et al., 2019]. As of November 2019 3150/4093 (∼78%) of confirmed exoplanets have been discovered via this method1.

Spectroscopy uses the radial velocity or Doppler shift technique to detect exoplanets.
ets. As the star and planet system orbit a common centre of mass the light from the star is red/blue-shifted along the observer’s line of sight. Measuring this periodic shift can reveal the presence of an exoplanet. Assuming edge-on, circularly orbiting systems, the size of the motion then gives the ratio of planet mass to total system mass to the power 2/3. Notable spectroscopic facilities include the High-Accuracy Radial-velocity Planetary Search [HARPS, Mayor et al., 2003], CORALIE [Queloz et al., 2000b], and the Echelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations [ESPRESSO, Pepe et al., 2014].

The first planet discovered via this method was 51 Pegasi b in 1995, which was also the first exoplanet to be found orbiting a sun-like star [Mayor and Queloz, 1995]. As of November 2019 779/4093 (~19%) of confirmed exoplanets have been discovered via this method².

Both photometry and spectroscopy require significant time and resources to properly follow-up exoplanet systems. As such it is vital to not waste telescope resources by using an inefficient technique when a different one could be carried out much more cheaply in terms of competitive time. To ensure this it is necessary to know which areas of parameter space lend themselves most naturally to each method. There are two criteria that are notable here, namely instrumental SNR (the reliability with which an instrument can detect a signal of a given magnitude) and follow-up time (how much time would be required for a given technique to successfully follow-up an exoplanet).

In this paper we attempt to answer these questions. We set out our method to choose the best follow-up procedure for an exoplanet that has discovery photometry (i.e. is transiting) but only an estimated or unsure period (a monotransit for example). This is the situation for many systems discovered as part of large photometric surveys. Section 7.2 discusses our methodology, sets out the instruments and methods used and our SNR and follow-up time definitions and criteria. Section 7.3 lays out our results and sections 7.4 and 7.5 give our discussions and the projects conclusions.

7.2 Methodology

To attempt to determine the feasibility of using each method for follow-up of exoplanet candidates we first defined a grid of points covering a range of exoplanetary parameter space. It was decided to define a point by 3 parameters as this lends itself easily to 3D representations of results. Additionally, using a small number of parameters increases the generality of the results. The first parameter chosen is stellar radius. Since we are interested in the follow-up of planetary systems with a discovery transit we assume that stellar radius

²https://exoplanetarchive.ipac.caltech.edu
will be a known parameter for the majority of systems as a result of comprehensive input catalogues for photometric surveys (e.g. the Kepler Input Catalogue [KIC, Brown et al., 2011] and the TESS Input Catalogue [TIC, Stassun et al., 2019]). Our second chosen parameter is planetary radius as this can be inferred from the discovery transit and the stellar radius. Finally we choose period as our third parameter due to the strong dependence of follow-up efforts on period. Additionally, period is a parameter that can be estimated from the discovery transit [Osborn et al., 2016].

The range of parameters chosen were determined from looking at a set of exoplanet parameters drawn from the NASA exoplanet archive [Akeson et al., 2013]. The chosen ranges were \(0.1 \leq R_\star \leq 10.0 \, R_\odot\), \(0.1 \leq R_p \leq 2.0 \, R_{\text{Jup}}\) and \(1.0 \leq P \leq 1000\) days. The simulation could then be carried out on every point in the grid of parameter space and a comparison of the two follow-up methods could be made. We create the parameter grid using 40 points along each parameter axis evenly separated in log-space. For each point in this parameter space two values needed to be determined. Firstly, the possibility that a signal corresponding to a point with these parameters could be detected using either of the two methods. This was defined as the possibility that the signal could be detected with \(S\,\text{NR} \geq 3.0\) for a given instrument. Secondly, it needed to be calculated how long it would take before a point with these parameters could be effectively followed up using either of the two methods.

### 7.2.1 SNR

To calculate the SNR we first needed to calculate the size of the relevant signal based on the three known parameters, \(R_\star\), \(R_p\) and \(P\). For photometry this is trivial if we assume the system is edge on to the observer (a reasonable assumption as we are only concerned with systems that have already been found to transit). The measurable signal is then simply the fractional reduction in flux caused by a transit;

\[
\delta = \left(\frac{R_p}{R_\star}\right)^2,
\]

where \(\delta\) is the size of the photometry signal [Winn, 2010b].

For spectroscopy the process is a little more involved. Under the assumption of an edge on, circular orbit the equation for radial velocity amplitude (the signal that must be detected) is given by;

\[
K = \frac{2\pi a M_p}{(M_\star + M_p) P},
\]

where \(K\) is the RV semi amplitude, \(a\) is the semi-major axis and \(M_\star\) and \(M_p\) are the
masses of the star and planet respectively [Lovis and Fischer, 2010]. The semi-major axis is given by

$$a = \left( \frac{P^2 G (M_\star + M_p)}{4\pi^2} \right)^{\frac{1}{3}},$$

(7.3)

where \(G = 6.67408 \times 10^{-11} \text{m}^3\text{kg}^{-1}\text{s}^{-2}\) is the gravitational constant [Lovis and Fischer, 2010].

Calculating the two masses requires finding a mass-radius relationship for both stars and planets.

**Mass-radius relations**

To determine the best mass-radius relation for stars and planets a set of data was taken from the NASA exoplanet archive. Since this paper focuses on the follow-up of planetary systems already found to transit the planetary data used include only transiting planets. The stellar data however, used all planet hosts from the exoplanet archive. These data were then filtered leaving only those data points for which mass and radius are known to a fractional uncertainty of 30% or better. To determine a mass-radius relation based on these data multiple techniques were tested. Figures 7.1(a) and 7.1(b) shows the results of these efforts. The NASA data is in black with the different colours corresponding to the range of different fits and relations that were tested.

For the stellar relationship the plots show the results of a cubic fit (in log-log space) and two published relations; Demircan and Kahraman [1991] (a one-to-one fit using the empirical values from their Table II) and Moya et al. [2018] (a one-to-one fit using relations 2 & 3 from their Table 10). For the planetary relation we show three published relations; Wolfgang et al. [2016] (a probabilistic relation), Forecaster\(^3\) from Chen and Kipping [2017] (a probabilistic relation) and MRExo\(^4\) from Kanodia et al. [2019] (a one-to-one fit).

For each tested method we computed the sum of the absolute values of the residuals between the method and the NASA data. The methods with the smallest total residuals were then chosen. Based on this test it was found that the stellar mass-radius relation was best approximated by the cubic fit whereas the planetary relation was best approximated using the MRExo relation. Therefore these are the mass-radius relations used in the simulation presented here.

\(^3\)https://github.com/chenjj2/forecaster  
\(^4\)https://github.com/shbhuk/mrexo
Figure 7.1: Mass radius relations, NASA data is in black.
(a) Stellar relations, the colours are as follows; Red squares: cubic fit, Blue diamonds: Demircan [Demircan and Kahraman, 1991], Green crosses: Moya [Moya et al., 2018].
(b) Planetary relations, the colours are as follows; Red squares: Wolfgang [Wolfgang et al., 2016], Blue diamonds: Forecaster [Chen and Kipping, 2017], Green crosses: MRExo [Kanodia et al., 2019].

**Instrumental noise**

Once the size of the photometric or spectroscopic signal has been calculated it is also necessary to determine the amount of noise that would be present on such a signal. Using this the SNR can then be calculated and it can be determined whether the signal would be observed. The levels of noise are instrument specific. The instruments employed in this simulation are NGTS for photometry and HARPS and CORALIE for spectroscopy. For each of these instruments the noise levels are functions of the magnitude of the host star. Figure 7.2 shows the noise thresholds as a function of magnitude for these three instruments. The NGTS noise model (black) is adapted from Figure 14 of Wheatley et al. [2018] and the HARPS and CORALIE noise models (red and blue respectively) are adapted from Figure 10 of Günther et al. [2017].

Using these data and the calculated signal sizes we can now determine the SNR for each instrument at each point in the parameter space to be explored.

**7.2.2 Follow-up times**

The simple calculation of SNR for each of the two methods is not sufficient to determine which method is the best follow-up technique. It is also pertinent to take into account the amount of time required for this follow-up. For example, a factor of 2 difference in SNR using each method is broadly irrelevant if the higher SNR method requires 100 times more
observing hours. To account for this it we attempted to quantify the amount of time each method would require to accurately characterise each system.

**Photometry**

For photometry it was decided that a system was sufficiently followed-up when at least 2 additional transits (or partial transits) had been detected. Here we define this to mean at least two distinct nights in which at least an ingress or egress is seen during the hours of astronomical night. This means we require either the start or end of a transit \((T_c \pm T_{dur}/2)\) to happen during the hours of darkness. The length of a night varies across the year but is, on average, \(\sim 8\) hrs \([\text{Cooke et al., 2018}]\) so we use this value to avoid biasing our results by when the observations are begun. This calculation then requires an additional parameter, transit duration \(T_{dur}\). Assuming an edge-on, circular orbit transit duration is given by

\[
T_{dur} = \frac{P}{\pi} \arcsin \left( \frac{R_\star + R_p}{a} \right),
\]

where symbols are as described above \([\text{Winn, 2010b}]\). Using the calculated value for \(T_{dur}\) together with the period it is straightforward to calculate how long until 2 transits are observed. However, this time will be affected by a number of random processes, specif-
ically, what fraction of the period has elapsed when observations begin and what fraction of nights are adequate for photometric observations. Since there is some randomness involved in the calculation it makes sense to run the process multiple times, taking the average required observing time at the conclusion of the simulation. In this simulation we run the process 10 times (larger numbers of iterations were tested and found to have little effect on the results but a large increase in computing time) for each value of period, planetary radius and stellar radius (and therefore, transit duration). For each possible observed transit we apply a constraint assuming that 20% of nights are inadequate for observations, thus transits which occur on these nights are missed. This fraction is based on a combination of bad weather effects, technical downtime and non-constant access to telescope time (see Figure 9 of Wheatley et al. [2018] for more details). As an approximation we assume this fraction is constant though it is most likely to vary throughout a year. Additionally, we assume that observations begin at a completely random point in the orbital phase.

Each simulation is run for a time of $10P$, where $P$ is the period of the system in question. If the simulation has not observed 2 transits within this time frame we set the follow-up time equal to $10P$ and move on. This has the effect of putting a lower bound on the follow-up time of some systems but reduces the computing time required sufficiently as to make the simulation tractable. As it turns out the vast majority of systems are followed-up before reaching this threshold. This fact, combined with the averaging of 10 runs means this limit has only a marginal impact on the results.

**Spectroscopy**

For spectroscopy we define a system as sufficiently followed up when we can determine its period, RV amplitude and phase to within 5% of true values from spectroscopy measurements alone. To model this we make the assumption that the orbit is circular and can therefore be fit with a sine curve of the form

$$y = K \sin \left( \frac{2\pi}{P} x + \phi \right) + \gamma. \quad (7.5)$$

This assumes that the planetary signal we are searching for is the only cause of RV variations. In reality there may be other signals including additional planets or stellar variability. This assumption is a necessary simplification to achieve the general results of this simulation but it should be noted that spectroscopic follow-up times are generally a lower bound (especially for small amplitude signals) due to these additional effects.

To predict spectroscopic follow-up time $T_{spec}$ we simulate RV data points being taken and continue until the number of data points is sufficient to allow estimation of period and RV semi-amplitude to sufficient precision. We simulate RV observations in somewhat
of a targeted strategy, assuming an observation every \( n \) days where \( n = \frac{P}{30} \) (rounded up to the nearest integer). As with photometry we assume that on 20% of observing nights we cannot obtain data, accounting for weather, technical issues and higher-priority telescope targets. The effect of weather on spectroscopy is lessened compared to photometry but the instruments considered here are likely to have many programs running, reducing availability for specific follow-up. Thus the same 20% was utilised. For each data point obtained we calculate its amplitude using equation 7.5 with \( \phi = \gamma = 0 \). We then introduce a noise to the measurement drawn from a Gaussian distribution with \( \mu \) equal to zero and \( \sigma \) equal to the noise at the chosen magnitude from Figure 7.2.

Every time a new data point is added we create a Lomb-Scargle periodogram [Lomb, 1976; Scargle, 1982] of the current data. If the highest power period is a match to the true period (to within 5%) we then fold the data on this found period and attempt to fit it with a sine wave of the form shown in equation 7.5 (we fit to the folded data as it was found that this method reduces the computational time required). Since the data is already folded on the correct period we can simplify equation 7.5 by setting \( P = 1 \) and \( \gamma = 0 \). The fit utilises a least squares method to fit the equation and upholds the requirement for SNR \( \geq 3.0 \). The fit then returns parameter \( K \) which is the RV semi-amplitude of the signal, and \( \phi \) which is the phase of the orbit. If \( K \) and \( \phi \) are within 5% of the true values we consider the signal recovered and record the time taken. Otherwise we simulate a new data point and repeat the process.

As for photometry, each simulation is run for a time of \( 10P \) where \( P \) is the period of the system is question. If the simulation has not solved the orbit to 5% within this time frame we set the follow-up time equal to \( 10P \) and move on. Once again we find that the vast majority of systems are followed-up before reaching this threshold resulting in only a marginal impact on the results. We make two additional requirements for spectroscopic follow-up. We require at least 5 valid data points and assume the systems cannot be characterised before it has been observed for at least half an orbital period.

Instrument availability

A factor that requires more discussion is that of telescope availability for the type of follow-up envisioned by this simulation. A photometric instrument like NGTS can be operated in a dedicated fashion. Because of this the main impacts that would reduce the instrument time that can be spent on this follow-up are external factors such as weather and instrument failure. These factors have been combined into an average of 20% of nights being removed as photometrically feasible, as mentioned in Section 7.2.2. For the spectroscopic instruments discussed here (and particularly for HARPS) this is less feasible. Observers wishing to use these instruments must contend with potentially disrupted schedules due to the popularity
of the instrument and the process of user queues. This is a hard effect to model, especially into the future when schedules are unknown. Our solution to this problem is to remove a random fraction of nights to simulate these falling into the gaps of the schedules. The benefit of a spectrograph is that they are less affected by bad weather, that is, conditions can be sufficient for spectroscopic observations but not for photometric ones. Because of these two factors combined, and informed by past user queues for these instruments, we choose to use the same fraction of 20% of nights as being unsuitable for spectroscopy.

7.3 Results

The results presented below all show the case for an $V = 11$ magnitude host star.

7.3.1 SNR results

Figure 7.3(a) shows a plot of the area of $R_\ast, R_p, P$ parameter space that can be observed by NGTS and HARPS with an SNR $\geq 3.0$. The colour bar on this plot corresponds to the ratio of the predicted SNR that could be reached with NGTS and HARPS with values $> 1$ being points for which HARPS would reach a better SNR and values $< 1$ indicating points for which NGTS would reach a better SNR. From this plot and the values in Table 7.1 it can be seen that approximately 48% of the physically realistic phase space can be accessed by both NGTS and HARPS. This 48% is then split approximately evenly (42:58) between points for which HARPS would achieve a higher SNR and points for which NGTS would be better. Figure 7.3(b) shows the same range of points as Figure 7.3(a) but this time includes the additional two areas that correspond to the regions of parameter space that only 1 instrument can reach with the required SNR $\geq 3.0$ (NGTS in orange and HARPS in cyan). The plot shows that there is a larger fraction of parameter space accessible to only HARPS (22%) than there is to only NGTS (5%). Key from these plots is that it is the longer period, smaller planet systems which are best followed by NGTS and the shorter period, larger planet systems that are best followed by HARPS. The empty regions of this plot correspond to the parameter space that is associated with unphysical systems (for example, planet larger than the star, or orbital separation less than stellar radius) as well as those regions for which neither NGTS nor HARPS can reach at sufficient SNR.

Figures 7.4(a) and 7.4(b) show the results of the SNR distribution again, this time comparing NGTS to CORALIE. As expected, there is a smaller region of space accessible to both instruments (27%) due to CORALIE’s higher noise values, however the overall trend is the same with the same regions of parameter space lending themselves to photometry and spectroscopy. In Figure 7.4(b) it is clear that there is now significantly more parameter space that can be accessed by NGTS only (25%) than by CORALIE only (8%), a simple result
Figure 7.3: SNR ratio in NGTS/HARPS parameter space. The gradated colour area indicates the region accessible to NGTS and HARPS with the colour indicating the ratio of their respective SNR values (bluer where HARPS is better, redder where NGTS is better). The solid orange region in Figure (b) indicates the region accessible to NGTS only and the solid cyan region in Figure (b) indicates the region accessible to HARPS only. Empty regions are unphysical or inaccessible systems.

of CORALIE being unable to reach as many low radial velocity signal systems as HARPS can.

7.3.2 Follow-up time results

Figures 7.5(a), 7.5(b) and 7.5(c) show plots of period against predicted follow-up time for NGTS, HARPS and CORALIE respectively. These follow-up times indicate the time until either two transits are observed (photometry) or period, amplitude and phase of RV signal are constrained to 5% or better (spectroscopy). These follow-up times are predicted using the methods set out in section 7.2.2. The plots additionally show some linear trends to aid the eye. Figure 7.5(a) includes a trend at $y = 3x$ and Figures 7.5(b) and 7.5(c) include trends at $y = x$.

For photometry it can be seen that the relation between period and follow-up time is approximately linear with the time required for two transits being approximately equal to 3 times the period. It is also seen that, for periods $\gtrsim 20$ days, this is relatively independent of the transit duration, shown in the colour bar. For shorter periods however, we do see a dependence, with shorter transit durations requiring more follow-up time.

For spectroscopy the trend is linear at larger periods ($P \gtrsim 10$ day) for all $K$ but flattens off at shorter periods for large $K$ values. This however is a direct result of our requirement for at least 5 valid spectroscopic points. For spectroscopy there is a dependence of RV amplitude with lower amplitude signals requiring more time to properly constrain as
Figure 7.4: SNR ratio in NGTS/CORALIE parameter space. The gradated colour area indicates the region accessible to NGTS and CORALIE with the colour indicating the ratio of their respective SNR values (bluer where CORALIE is better, redder where NGTS is better). The solid orange region in Figure (b) indicates the region accessible to NGTS only and the solid cyan region in Figure (b) indicates the region accessible to CORALIE only.

seen in the colour bar. Additionally, by comparing Figures 7.5(b) and 7.5(c) it can be seen that, when comparing a system with the same period and RV amplitude, it takes CORALIE longer to constrain a system than HARPS, as is expected from the worse noise performance of CORALIE compared to HARPS. Each plot shows all points accessible to that instrument.

Figures 7.6(a) and 7.6(b) show the same regions of parameter space as shown in Figures 7.3(a) and 7.4(a). This time however, the plots are coloured by a ratio of the predicted follow-up times for NGTS & HARPS and NGTS & CORALIE respectively.

The main point of interest from these plots is that there are regions of parameters space more amenable to follow-up by both methods. For the very shortest period planets ($\lesssim 10$ d) photometry is quicker at follow-up, most likely due to the additional requirements for at least 5 spectroscopic points. For longer periods spectroscopy is quicker for follow-up of those systems with larger planets whereas photometry can be faster for smaller planet systems. The ratio of follow-up times is broadly independent of stellar radius. The same general pattern is seen in both figures with the main difference being a slight preference for photometry in Figure 7.6(b) compared to Figure 7.6(a) as a result of HARPS being quicker to constrain the same system than CORALIE (see Figures 7.5(b) and 7.5(c)). Not shown in these plots are the regions of parameter space that only one instrument can access at the required level of SNR. The instrument that can quicker constrain a given system in these regions is broadly irrelevant as only one instrument would actually be able to confidently measure the required signal.

An additional point that must be mentioned in regards to follow-up time is the fact
Figure 7.5: Predicted follow-up time as a function of period for NGTS, HARPS and CORALIE. For NGTS a colour bar denoting transit duration is shown and for HARPS and CORALIE a colour bar denoting RV amplitude is used. Additionally, Figure (a) includes a linear $y = 3x$ trend and Figures (b) and (c) include linear $y = x$ and $y = 10x$ trends to aid the eye. Each plot shows all points accessible to that instrument.
that stars are only seasonally observable from the ground. To this end, the total follow-up times may be longer than those shown here, particularly for systems with periods above \( \sim 180 \) days. Since the three instruments considered here are all within geographic proximity to one another (all being located in northern Chile) this effect should impact all observations equally. Because the point of interest in this simulation is the relative follow-up time between the different methods the follow-up times are not extended since there would be no change to the relative times. However this is an effect that should be considered when planning real observations.

### 7.3.3 Other magnitudes

The results and plots presented thus far in this paper have looked at a host star of \( V = 11 \) magnitude. These results however, will not be identical for different magnitudes as the noise profiles of the three instruments varies (see Figure 7.2). For this paper we reran our analysis at \( V = 8 \) and \( V = 14 \). The broad patterns in the results are consistent with similar regions of parameter space preferring photometry and spectroscopy as for the \( V = 11 \) case. The main difference between the three cases is the amount of parameter space which is or isn’t available to the three instruments. Where multiple instruments are viable the outcomes are broadly unchanged. Table 7.1 shows how the accessible regions of parameter space vary with magnitude.

As this table shows, and as expected, increasing host magnitude leads to a reduction
Table 7.1: Accessible parameter space as a function of host magnitude. Given as a percentage of the full physically realistic range of points considered (57847 points).

<table>
<thead>
<tr>
<th>Magnitude (V-band)</th>
<th>8</th>
<th>11</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>NGTS (total)</td>
<td>54</td>
<td>53</td>
<td>37</td>
</tr>
<tr>
<td>NGTS (only)</td>
<td>0</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>HARPS (total)</td>
<td>94</td>
<td>70</td>
<td>35</td>
</tr>
<tr>
<td>HARPS (only)</td>
<td>19</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>CORALIE (total)</td>
<td>71</td>
<td>35</td>
<td>7</td>
</tr>
<tr>
<td>CORALIE (only)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NGTS and HARPS</td>
<td>54</td>
<td>48</td>
<td>22</td>
</tr>
<tr>
<td>NGTS and CORALIE</td>
<td>49</td>
<td>27</td>
<td>6</td>
</tr>
<tr>
<td>NGTS and HARPS and CORALIE</td>
<td>49</td>
<td>27</td>
<td>6</td>
</tr>
</tbody>
</table>

in the volume of parameter space available to each instrument. This is a natural consequence of the increased noise that accompanies dimmer targets. The table also confirms that there is no area of parameter space accessible to CORALIE and not HARPS. Once again this is as expected since HARPS has a lower noise floor than CORALIE at all magnitudes.

There is an additional preference towards spectroscopy for brighter targets. Brighter stars lead to lower noise floors and higher SNR values for all instruments but spectroscopic follow-up time is also affected. The search for additional transits is not affected by host magnitude since only transits with $SNR \geq 3.0$ are considered. For spectroscopy however the noise value is built into the simulation of data and fitting routines. These should then be more efficient for brighter hosts leading to a reduced follow-up time for spectroscopy. For a system that can be observed at $SNR \geq 3.0$ at both $V = 8$ and $V = 11$ the follow-up time for photometry is unaffected whereas the follow-up time for spectroscopy is reduced for the brighter target.

### 7.3.4 Online repository

The results presented here in Figures 7.3, 7.4, 7.6 and 7.7 are attempts to display 3D information in a 2D medium. This is obviously non-ideal so we provide a link to an online repository containing the 3D data and plotting script required to create these figures. The details are included to recreate and manipulate all plots for 8th, 11th & 14th V-band magni-
tude hosts. The repository is to be found at https://github.com/BenCooke95/SpecPhot. This repository also contains the codes necessary to reproduce the simulation discussed in this paper.

7.4 Discussion

7.4.1 SNR

The methods used to calculate SNR presented in this paper are fairly simplistic but are a realistic approximation. In reality, of course, SNR can be increased by taking more measurements and binning data points. This can then beat down the noise and allow smaller signals to show through. In this paper however we are interested in the broader picture for which considering SNR as the simple ratio of signal size to noise size is sufficient.

To explain the behaviour of our SNR ratio with $P$, $R_\star$ and $R_p$ we can look at equations 7.1, 7.2 and 7.3. Combining these equations to find the ratio of the signals gives us the relation

$$\frac{K}{\delta} \propto \frac{M_p R_p^2}{R_p^3 (M_\star + M_p)^{\frac{3}{2}} P^{\frac{1}{2}}}.$$  \hfill (7.6)

Taking $M_\star \propto R_\star^3$ (see Figure 7.1(a), the relation used in this paper) and $M_p \propto R_p^3$ (see Figure 7.1(b), an approximation to the relation used in this paper) and assuming $R_\star^3 \gg R_p^3$ we find that

$$\frac{K}{\delta} \propto \frac{R_p}{P^{\frac{1}{3}}}.$$  \hfill (7.7)

This relation now explains our results well. We see clearly that longer period systems favour better photometry SNR than spectroscopy in accordance to the $P^{-\frac{1}{3}}$ term. Additionally, our SNR ratio is broadly independent of stellar radius but generally favours spectroscopy for larger radius planets. Relaxing the approximations reasoned here and allowing for the more realistic planet mass-radius relation used explains the distributions seen in Figures 7.3 and 7.4.

7.4.2 Follow-up time

There are many possible definitions of follow-up time, or how long before a system is sufficiently characterised. In this paper we have defined a successful follow-up campaign as one which either observes two additional transits (photometry) or measures period, RV amplitude and phase to 5% or better (spectroscopy). In reality, the exact conditions required
for a successful follow-up may change in a system dependent way but we consider the criteria stated here as reasonable baselines.

For photometry we see that the follow-up time is approximately equal to $3P$ and largely independent of other parameters. This makes sense since the frequency with which transits occur is $1/P$. A longer transit duration will help to increase the chance that at least part of a transit occurs during a good observing night but this is a small effect compared to the decreasing frequency of events towards longer periods. The fact that the follow-up time is not always $\leq 2P$ is due to the fact that we require part of a transit during a night of good observing quality.

The follow-up time to period relation for spectroscopy is more detailed. For longer period systems the follow-up time is usually $\gtrsim P$ with smaller amplitude signals requiring longer to follow up. This implies that we generally require close to or above 100% phase coverage to accurately measure period and RV amplitude for long period systems. Within this regime higher amplitude signals are characterised with fewer data. The relation is approximately 1:1 for $K$ values of $\sim 20\,\text{ms}^{-1}$ and $\sim 50\,\text{ms}^{-1}$ for HARPS and CORALIE respectively but is closer to 5:1 for $K$ values of $\sim 4\,\text{ms}^{-1}$ and $\sim 10\,\text{ms}^{-1}$ for HARPS and CORALIE respectively. For shorter period systems, less than around 10 days, the linear relation begins to break down (at least for larger $K$). This is due to our requirement for $\geq 5$ data points before attempting characterisation. Even in this regime the effect of larger amplitude signals is still seen, though to a lesser degree due to the clumping of points between 5 and 10 days.

As an additional factor to consider we see that when the period is an integer day the follow-up time can spike to infinity due to every transit occurring during the day (for photometry) or only a limited amount of phase coverage being possible (for spectroscopy). This is a known issue for ground-based observations and the difficulty of characterisation leads to a dearth of integer day periods, not representative of the true distributions.

7.4.3 Weighting

To be able to rank individual parameter combinations in terms of SNR and follow-up time requires combining both terms into a single value. Additionally, for a more realistic comparison we must also take into account the availability of the two methods being compared in terms of telescope time. Finding the appropriate weighting that takes these factors into account is non-obvious and, ideally, bespoke to the needs and resources of a specific follow-up effort or research group. For the purposes of this paper we combine the factors in a simplistic way. We simply multiply the ratio of SNR values and the ratio of follow-up times. This leads to a single number for each set of $R_*, R_p$, and $P$ values with values $> 1$ meaning spectroscopy is preferred and values $< 1$ meaning photometry is preferred. Additionally,
Figure 7.7: Weighted combinations of SNR and follow-up time. These plots are a combination of the distributions in Figures 7.3, 7.4 and 7.6 weighted by a factor of $\frac{1}{12}$. The colour bar shows the weighted value with values $< 1$ indicating a preference for photometry (NGTS) and values $> 1$ indicating a preference for spectroscopy (HARPS or CORALIE).

we multiply by a factor of $\frac{1}{12}$ to reflect that each of HARPS and CORALIE comprise of a single instrument whereas NGTS contains 12 individually operable telescopes [Wheatley et al., 2018]. Figure 7.7 shows the weighted distributions for NGTS & HARPS and NGTS & CORALIE.

Comparing Figure 7.7 to Figures 7.3 and 7.4 we see that the weighting has caused noticeable changes, most significantly for the NGTS/HARPS comparison. For NGTS & HARPS we see that the area of the distribution more favourable to photometry has increased compared to Figure 7.3, especially towards low periods. The factor of $\frac{1}{12}$ would lead to an increased preference for photometry across the distribution but the small period range is more affected due to the increased follow-up time spectroscopy has in this parameter range (seen in Figure 7.5). The NGTS/CORALIE distribution is similarly affected with an enhancement of photometry towards low periods. The distribution is less obviously difference due to the smaller region of parameter space accessible to CORALIE and the already weaker SNR and follow-up times when compared to HARPS.

7.4.4 Importance of results

Follow-up of exoplanetary systems, especially poorly characterised ones, is a detailed undertaking that can involve different criteria based on required outcome and potential resources. The myriad specifics that could potentially be involved are too numerous to explore in a paper such as this so we make some simplifying assumptions and choose to focus on the broadly applicable results. Even at this level of detail, however not all of the results
presented here are obvious. For example, the overall relation describing the ratio of spectroscopic to photometric SNR (equation 7.7) can be simply found by making some assumptions to the common transit and radial velocity equations but the more detailed structure present in Figures 7.3 and 7.4 is not straightforward to predict. The precise dividing line between the two regimes cannot be intuited but may be vital when arguing for one method over another. The follow-up time data is not necessarily obvious either. The fact that short period planets can be quickly confirmed using photometry is easy to explain (multiple transit opportunities within only a few days) but the fact that there are systems up to even the longest periods for which photometry is the quicker follow-up method is surprising.

A simulation like this is a good starting point when attempting to garner time on a particular instrument for follow-up by showing that other methods have been considered and that the method you are targeting is appropriate for your specific case. Securing time on spectroscopy or photometry instruments can be difficult and results like these, showing that one method is preferred or more efficient, are useful evidence.

7.5 Conclusions

The results presented in this paper show how the effectiveness of using photometry or spectroscopy for exoplanet follow-up changes as a function of $R_*$, $R_p$ and $P$. In terms of purely measuring the SNR value of a detection we see that photometry is favoured for longer period systems and for smaller radius planets whereas spectroscopy is favoured for shorter period systems and larger radius planets. Generally both methods perform comparably regardless of stellar radius. In terms of follow-up time, we find that photometry usually follows-up a system faster for small planet systems whereas spectroscopy is faster for larger planets. The region where this becomes less of a factor is towards shorter period systems where follow-up times generally prefer photometry due to the spectroscopic methods departure from a linear relation.

Because of this dependence on period it can be difficult to choose the best method when the period of the system is uncertain (the case for monotransits). As mentioned before, the shape of the discovery transit can inform the period with an accuracy up to 10%, generally this is accurate enough to help determine the chosen follow-up method (especially if relying on the weighted ratios as in Figure 7.7 which are less dependent on period). However we mention here that extra care must be taken for systems with poorly constrained periods.

To choose the best method with which to follow-up these systems requires taking into account both criteria, SNR and follow-up time, with an appropriate weighting based on the specific goals of the follow-up efforts. For borderline signal cases one should most
likely weight the SNR ratio higher, whereas for obvious but scientifically important systems follow-up time is likely to be more important. Additionally, both methods must be considered accounting for availability of telescope time, a shorter amount of required spectroscopic time may be a moot point if photometric time is more readily available. We make an estimation of a typical weighting based on the larger number of NGTS telescopes but bespoke weighting based on specific research goals and resources would doubtless be more realistic.
Chapter 8

Period aliases

Note

The following chapter is a reproduction of the paper ‘Resolving Period Aliases for TESS Monotransits Recovered During the Extended Mission’ [Cooke et al., 2021] published in Monthly Notices of the Royal Astronomical Society in November 2020. The simulations, results and analysis presented here were all designed and carried out by me unless otherwise stated. The writing was all done by myself. Co-author contributions consisted of discussions and improvements regarding methodology and results as well as general proof reading.

This paper uses the results of Chapter 5 and explores the optimal methods of using ground based follow-up data to contain monotransit periods. The results of this chapter (along with those of Chapters 4 and 5) explicitly informed the procedure of the NGTS monotransit working group discussed in Chapter 9.

Abstract

We set out to explore how best to mitigate the number of period aliases for a transiting TESS system with two identified transits separated by a large time period on the order of years. We simulate a realistic population of doubly transiting planets based on the observing strategy of the TESS primary and extended missions. We next simulate additional observations using photometry (NGTS) and spectroscopy (HARPS and CORALIE) and assess its impact on the period aliases of systems with two TESS transits. We find that TESS will detect around 400 exoplanets that exhibit one transit in each of the primary and extended missions. Based on the temporal coverage, each of these systems will have an average of 38 period aliases. We find that, assuming a combination of NGTS and CORALIE over observing
campaigns spanning 50 days, we can find the true alias, and thus solve the period, for up to 207 of these systems with even more being solved if the observing campaigns are extended or we upgrade to HARPS over CORALIE.

8.1 Introduction

The Transiting Exoplanet Survey Satellite [TESS, Ricker et al., 2015] has recently completed its observations of the northern ecliptic hemisphere during the second half of its primary mission. As of 5 July 2020 TESS has begun its extended mission in which it will first carry out a re-observation of the southern ecliptic hemisphere, approximately 2 years after first observing it (for extended mission details see NASA 2019 senior review\(^1\) and NASA response\(^2\)).

TESS has already produced numerous discoveries of planets around bright stars [e.g. Huang et al., 2018a; Gandolfi et al., 2018; Vanderspek et al., 2019; Gilbert et al., 2020] as well as discoveries of systems for which we observe only a single transit, [e.g. Gill et al., 2020a; Lendl et al., 2020b; Gill et al., 2020b], including the first monotonotransiting planet from TESS, Gill et al. [2020c]. The detection of monotonotransits is a natural consequence of the limited temporal coverage of TESS when compared to other surveys such as the Wide Angle Search for Planets [WASP, Pollacco et al., 2006] or Kepler/K2 [Borucki et al., 2010; Howell et al., 2014]. These monotonotransits have been explored via simulations [Cooke et al., 2018; Villanueva et al., 2019], which show we can expect a large number of monotransit candidates in the TESS primary mission data. These predictions have been verified by recent monotransit searches including Montalto et al. [2020] who find 15 southern hemisphere candidates and the NGTS monotonotransit team (see Gill et al. [2020c,b] for a description of the pipeline used) who have identified over 50 candidates from the first few TESS sectors alone. Many of these candidates exhibit additional extended mission transit events as discussed later in this paper. Because of the close dependence of monotransit detections on survey coverage the determination of orbital periods for single-transit candidates will be heavily influenced by the continuation of TESS into its extended mission [Cooke et al., 2019].

With the data from the first year of the extended TESS mission (i.e. Year 3 of TESS), most southern ecliptic TESS targets will be observed for about twice as long. However, these additional observations will be carried out approximately 2 years after the initial observations. Due to this, there will be significant gaps in the photometric baseline which may hide additional transits. This effect is most pronounced for those systems which only

\(^1\)2019 Senior Review Subcommittee Report
\(^2\)NASA Response to the 2019 Senior Review of Operating Missions
transited once during the primary mission and will only transit once more in the first year of the extended mission. Based on only two transits we cannot infer the true orbital period, we know only that it is bounded. The maximum period is given by the separation of the two observed transits which is on the order of \( \sim 2 \) years. The minimum period is given by the constraint that only one transit is seen in each TESS observing run, approximately 10 days for a single sector of observations. Therefore we are left with a discrete set of period aliases [Cooke et al., 2019].

Period aliases are an important consideration when analysing non-continuous observations. This is a common problem for ground based observations due to the daytime and weather interruptions in the data; any signal short enough to fit within a day can be missed and it requires sufficient night time observations to rule out these potential periods. However, for longer period systems, this is a significant issue for space based observations as well [Becker et al., 2019; Dawson and Fabrycky, 2010]. Some attempts have been made to use period priors to prioritise aliases to search [Dholakia et al., 2019] but this is far from foolproof. To actually rule out aliases requires additional data in the form of photometry and/or spectroscopy. This paper looks at exactly what effect additional observations have on the number of period aliases of a given system and how the effect changes depending on the amount and quality of these observations. We explore both photometry and spectroscopy by simulating each method, and investigating how this reduces the set of period aliases for TESS discoveries. For a recent comparison of photometric and spectroscopic follow-up methods see Cooke and Pollacco [2020].

This paper is laid out in the following manner. Section 8.2 discusses the stellar and planetary population used as well as the simulated TESS, photometric and spectroscopic data. Section 8.3 details our attempts to use the data to identify and analyse the period aliases. Section 8.4 presents the results of our simulation and analysis while Section 8.5 outlines our conclusions.

8.2 Observations

8.2.1 Simulation population

To predict the number of planets for which TESS will observe two individual transits, we first create a planetary population. TESS observations of this population are then simulated. To do this we proceed as in Cooke et al. [2018] which we summarise below.

We use the TESS Input Catalogue Candidate Target List version 8 [Stassun et al., 2019] accessible via the Mikulski Archive for Space Telescopes (MAST\(^3\)) as our stellar

\(^3\)http://archive.stsci.edu/tess/tic_crtl.html
population, which includes Gaia DR2 parameters [Gaia Collaboration et al., 2018a]. We perform the following simple cuts on this population based on TESS-band magnitude ($3.0 \leq m_T \leq 17.0$) and effective stellar temperature ($2285 \leq T_{\text{eff}} \leq 10050$ K) leaving 4789372 targets in the southern ecliptic hemisphere.

We generate planets around these stars by drawing from period-radius bins under particular occurrence rates given by Dressing and Charbonneau [2015] for M-stars and Hsu et al. [2019] for FGK-stars. Note that this is an updated set of occurrence rates compared to the procedure laid out in Cooke et al. [2018]. In particular, Hsu et al. [2019] uses the final Kepler Data Release [DR25, Thompson et al., 2016] and the improved stellar parameters made available via GAIA DR2 [Gaia Collaboration et al., 2018a]. More specific transit parameters are then calculated using equations and distributions presented in Winn [2010b] and Barclay et al. [2018]. We include a uniform distribution in $\cos i$ for inclination used to calculate transit probability, but assume all planets around the same star share an inclination value.

8.2.2 TESS observations

To identify which systems would be observed to transit once in both the primary and extended TESS missions, we simulate TESS observations. Since it will be the first hemisphere to be re-observed in the extended mission, and due to our use of southern ecliptic hemisphere based follow-up observatories, we only simulate southern ecliptic hemisphere TESS targets. Our methods follow that described in Cooke et al. [2019]. We simulate observations only at times for which there are data points not affected by systematics from the TESS primary mission. This has the effect of producing a more realistic coverage fraction per sector and accounts for systematic effects such as momentum dumps. We assume that the primary mission coverage can be replicated during the extended mission, noting that in the extended mission Full Frame Images will be taken with a improved cadence of 10 minutes.

Based on our simulated TESS coverage we can now determine which planet transits are detectable with TESS. Using our planet parameters and a 5th order polynomial TESS noise approximation Stassun et al. [2018] we calculate a Signal-to-Noise ratio, $S/N$ for each planet. Those planets which exhibit a transit during a region of TESS observations that manage to achieve the required $S/N \geq 7.3$ are recorded as detectable. For more specific details regarding the simulated TESS observations see Cooke et al. [2019].

Once the simulation is complete we select all those planets with one detectable transit during the TESS primary mission and one detectable transit during the first year of the extended mission as our double transit sample.
8.2.3 Simulated photometry

We simulate additional photometric data using the Next Generation Transit Survey [NGTS, Wheatley et al., 2018]. For this additional photometry we follow the procedure being employed by NGTS [see, e.g., Gill et al., 2020a; Lendl et al., 2020b]. Observations are simulated from sunset to sunrise using a cadence of 12s. We use the NGTS noise model as a function of host magnitude shown in Figure 8.1 and compare this to the signal size of the planet, \((R_p/R_\star)^2\). Additionally, we allow the use of multiple NGTS telescopes if required to reach the necessary \(S/N\) level. To calculate the \(S/N\) level for multiple telescope we use a \(1/\sqrt{N}\) relation as has been shown to be a good approximation by Bryant et al. [2020a]. Where NGTS can achieve a \(S/N \geq 3.0\) we use the simulated photometry. Below this \(S/N\) threshold we assume that photometry is of insufficient quality to help constrain the period of the system and do not simulate follow-up photometry.

To best utilise photometric time we only simulate observations on nights when we predict a transit may occur (see Sect. 8.4.5 for details including accounting for transit time uncertainties). Since this paper explores those systems with transits in both the primary and extended TESS missions we can use the TESS transit times and coverage to create a set of discrete period aliases and we therefore only simulate NGTS observations on nights that at least one alias predicts will contain a transit. This is a realistic strategy that avoids wasted time that would occur by observing on nights which provide no period constraining information.

We can then run this strategy for a fixed number of nights, observing some subset of those. In the main body of this paper we consider an overall time-span of 50 nights. Appendix 8.A then includes the results for a range of different maximum time-spans (Table 8.1). As an additional measure, we include telescope downtime from instrumental and weather effects. This is included by assuming that a random 20% of nights are unavailable for observations (for more details see Cooke et al. 2020b).

8.2.4 Simulated spectroscopy

Spectroscopic data is simulated using the High-Accuracy Radial-velocity Planetary Search [HARPS, Mayor et al., 2003] on the ESO 3.6 m telescope and CORALIE [Queloz et al., 2000b] on the Euler 1.2 m telescope. We simulate an observing strategy of one 30 minute exposure every 7 days for each candidate, including a noise value based on the noise models of the instruments as a function of host magnitude shown in Figure 8.1. The amplitude of the radial velocity signal is a function of the true period (which is known), the stellar mass (which is known), and the planetary mass (which is unknown). Therefore we first
Figure 8.1: Instrumental precision as a function of target TESS magnitude for NGTS, HARPS and CORALIE (adapted from Figure 14 of Wheatley et al. 2018 and Figure 10 of Günther et al. 2017).

predict the planetary mass using the planetary radius and the MRExo mass-radius relation\(^4\) from Kanodia et al. [2019]. For each radius value we draw a random value from a mass distribution generated using MRExo.

As with photometry, this procedure can be run for any time-span, the difference being that now the number of data points is one seventh of the number of nights. In this paper we again consider an overall time-span of 50 nights using CORALIE with alternative results (including using HARPS) presented in Appendix 8.A. As with the simulated photometry in Section 8.2.3, we account for telescope downtime from instrumental and weather effects as well as scheduling conflicts. This is included by assuming that a random 20% of nights are unavailable for observations (for more details see Cooke et al. 2020b).

### 8.3 Alias determination

#### 8.3.1 Period aliases from TESS

Based on our simulated TESS observations in Years 1 and 3 (primary and extended mission) we select the subset of planets which exhibit one transit during each year. The separation of these transit times gives an upper limit to the period of the planet as well as the knowledge

\(^4\)https://github.com/shbhuk/mrexo
that the true period must be equal to this separation divided by an integer. Assuming that the minimum period that a planet could have and still only be observed once in a sector is \( \sim 10 \) days we can make a list of all possible period aliases. Some of these periods however can be ruled out by the TESS coverage we already have. We test each period alias by comparing with the TESS coverage. For each alias we create a list of all times at which the system would transit (including times before the Year 1 transit and after the Year 3 transit) and check if there are any TESS observations within half a transit duration of any of these. If there are TESS observations that overlap with any transit time (excluding the two discovery transits) we can rule out that alias as one that would result in an additional transit which has not been seen. In this way we reduce the list of period aliases down to only those allowed based on the TESS photometry.

### 8.3.2 Period aliases using additional data

With the newly simulated photometry and spectroscopy we can now attempt to better constrain the true period of each monotransit and rule out more aliases. We do this first using the TESS and follow-up photometric data, then the TESS and spectroscopic data, then finally using the TESS, photometric and spectroscopic data.

To combine the NGTS data we proceed on a nightly basis. For each night we check if any of the possible period aliases predict a transit to occur. If so, we simulate observations. If no transit is seen it shows that the true period is not one of those that predicted a transit this night and we therefore remove all periods that predicted a transit that night from our list of aliases. If a transit is seen we can use this third transit to remove any aliases that do not predict this additional detection. This process repeats each night until the period is solved or we reach our observing limit.

Ruling out aliases using spectroscopic data requires a different approach. Firstly, we require a minimum of 5 radial velocity measurements before we begin to rule out aliases. Once we have reached this threshold, every time we simulate a new spectroscopic data point we fold the radial velocity data on each alias and plot in phase space. For each allowed period we model the radial velocity variation assuming a circular orbit and fitting for the amplitude, phase, and gamma velocity. We then calculate the residuals between our simulated measurements and this fit. If the average of the absolute residuals is greater than 1.5 times the noise threshold at the chosen magnitude we reject the alias. Otherwise the alias is still a valid option for the system. Since we require at least 5 data points the average of the residuals is unlikely to reach this threshold for the true period. In multiple runs the true period was never rejected. This process is repeated until the period is solved or we reach our observing limit.

To find the viable aliases using a combination of TESS photometry, photometric
follow-up, and spectroscopic follow-up, we simply examine the alias sets from TESS and photometry and from TESS and spectroscopy and allow only those aliases present in both sets. This is carried out on a nightly basis to determine the earliest point at which a system is solved.

8.4 Results and Discussion

8.4.1 TESS extended mission monotransits

Based on our simulated observations we expect 759 monotransits will be found in the TESS southern ecliptic primary mission. Of these, approximately 65% will transit again in the first year of the extended mission with most only transiting once more. The result is that slightly over 50% of the primary mission southern monotransits will transit once more in the first year of the extended mission, giving a total of 395 planets that will exhibit one detectable transit in both the primary and extended TESS missions for the southern ecliptic hemisphere. This number exceeds those presented in Cooke et al. [2018] and Cooke et al. [2019] by a factor of 2-3. This is due to the updated occurrence rate used in this study. The occurrence rate used here comes from Hsu et al. [2019] which uses a more up to date and larger sample based on the Kepler mission. Therefore these results will be more robust than similar studies based on outdated occurrence rates.

8.4.2 Number of aliases, $N_{\text{aliases}}$

For each southern ecliptic hemisphere planet that TESS will observe to transit once in each year we have produced a list of period aliases. From this list we determine the number of aliases, $N_{\text{aliases}}$, for each system based on different subsets of observational data. For each system we then calculate $N_{\text{aliases}}$ using four different combinations of observational data; TESS data only, TESS with additional photometry, TESS with additional spectroscopy and finally TESS with additional photometry and spectroscopy. We define the number of aliases found using TESS data alone as $N_{\text{aliases},i}$ and the updated value found by taking into account different combinations of follow-up data as $N_{\text{aliases},f}$.

Figure 8.2 shows the distribution of the number of aliases from TESS, $N_{\text{aliases},i}$, and the number of aliases from the three combinations of follow-up data, $N_{\text{aliases},f}$.

For each plot in Figure 8.2 we show the value of $N_{\text{aliases},i}$ or $N_{\text{aliases},f}$ for all unsolved systems when accounting for the four regimes of observational data. The solved systems (those for which the period is known, i.e. $N_{\text{aliases},f} = 1$) are not shown in the plots as they show large spikes at 1 overwhelming the rest of the distributions. The number of solved systems, and therefore the number of systems not shown, is in the title of each plot. The
Figure 8.2: Distributions of the number of period aliases from each combination of observational data. The four plots show $N_{\text{aliases},i}$ and the three $N_{\text{aliases},f}$ values for the four combinations of observations. For the right hand column plots (those including additional photometry) the data is split into those for which an additional transit is seen (red) and those for which no additional transit is seen (blue). Each plot tile denotes the data used and how many systems have been solved with this data (these systems are not shown in the plots to avoid overwhelming the histograms). Here we assume a time-span of 50 days and use CORALIE for spectroscopy.
Figure 8.3: Distribution of signal size for photometry and spectroscopy for all considered monotransits.

top left plot shows the distribution of the number of aliases per system using TESS data alone, \( N_{\text{aliases},i} \). This gives an average value of 38 aliases per system, however it should be noted that there is a double peak in the distribution. This corresponds to those systems which are observed in multiple TESS sectors. Systems with more TESS coverage have correspondingly fewer aliases leading to a secondary peak around 20. Systems with a single TESS sector per year have more aliases and peak around 40.

The top right and bottom plots show the distributions of \( N_{\text{aliases},f} \) based on three combinations of follow-up data. As can be seen photometry solves significantly more systems than spectroscopy (149 and 50 systems solved respectively) but combining the two methods leads to a number of solved systems greater than simply adding the uniquely solved systems together (207 solved systems). Figures including additional photometry, those in the right hand column, are split depending on whether the photometry caught additional (one or more) transits. It is seen that catching a transit leads to a large reduction in \( N_{\text{aliases},f} \) (see Gill et al. 2020c for an example of this in practice) and all of the systems solved from additional photometry alone (149) had at least one additional transit detected. Included in the plots are those systems for which the photometry is of insufficient \( S/N \). Transits may occur during observations of these systems but would not be detected with sufficient confidence. For these systems additional photometry cannot help to constrain the period. Additionally, some of the systems result in spectroscopic signals smaller than the respective noise level. These points are present in the plot but additional spectroscopy would not be able to rule out any periods. Figure 8.3 shows the distributions of both transit depth and radial velocity semi-amplitude for the monotransits considered. For both photometry and spectroscopy we require \( S/N \geq 3 \) to simulate observations.
As well as the number of solved systems proved by each combination of observational data these distributions show the number of aliases for those unsolved systems. For TESS photometry only, all 395 systems are unsolved and the average (mean) value of $N_{\text{aliases,i}}$ is $\sim 38$. When photometry is included we have 246 unsolved systems with an average $N_{\text{aliases,f}}$ value of $\sim 14$. Using spectroscopy instead gives 345 unsolved systems with an average $N_{\text{aliases,f}}$ value of $\sim 26$. Finally, including both photometric and spectroscopic data on top of the TESS photometry gives 188 unsolved systems with an average $N_{\text{aliases,f}}$ value of only $\sim 12$.

From Figure 8.2 we see that the number of systems solved when combining the follow-up photometry and spectroscopy is larger than the sum of the systems solved using each method individually. Further analysing this we find that of the systems solved by spectroscopy, 19, or 38%, are also solved using photometry. This absolute value is obviously then the same for the photometrically solved systems also solved by spectroscopy but in this case it is only 13% of the total.

Taking this analysis forward to the systems solved when combining the two methods we find that 27 of these (13%) are not solved by either photometry or spectroscopy individually. These 27 systems are therefore systems for which each method manages to rule out different subsets of aliases but is incapable of completely solving the period. Taking the common aliases left from both follow-up methods results in these systems being solved. It should also be noted that photometric and spectroscopic data provide complementary information about the radius and mass of the planet, another reason that a combination of follow-up data is highly desirable.

Figure 8.4 shows the distribution of solved systems with orbital period and planetary radius. The top panel shows the results for spectroscopy with systems solved being coloured by magnitude and unsolved systems being grey. Systems with insufficient $S/N$ are omitted from the plot. The middle panel shows the corresponding plot for photometry and the bottom panel shows the combination of the two methods. In this panel all systems are plotted regardless of either photometric or spectroscopic $S/N$ and systems that are only solved when the methods are combined are shown as diamonds.

From this plot we see that the majority of systems solved by spectroscopy have radii between 10 and $15R_\oplus$. This is as expected since predictions and observations suggest this radii range corresponds to the most massive planets which are the ones that spectroscopy is most sensitive to. The photometric planets are more spread in radius since they tend to be insensitive to radius (and mass) as long as the transit depth is above the sensitivity threshold. Both methods favour shorter periods since these are obviously easier to confirm within a given observing program. It also seems that the subset of planets solved only when both data sets are combined (diamonds in Figure 8.4) have a broader period distribution. Once
Figure 8.4: Period and radius distribution of solved systems. The top plot shows the distribution for spectroscopically solved systems, the middle is that for photometrically solved systems and the bottom shows the combination of both methods. Diamond points denote systems only solved when both methods are combined. The colour of each point denotes its TESS-band magnitude.
Figure 8.5: True period index, $I$, for all 395 TESS systems which show a single transit in each of the Year 1 and Year 3 data sets. A value of $I = 0$ means the true period is the smallest allowed alias. Again, this is unsurprising as the additional data help to confirm these more challenging systems.

### 8.4.3 True period index, $I$

Based on the increased likelihood of detecting short period planets (transit probability scales as $P^{-3/2}$) it might be natural to assume that, even for unsolved systems, the true period is most commonly the shortest of the period aliases. However, this is frequently untrue. Figure 8.5 shows the distribution of the true period index $I$ for all 395 systems that transit once in each of the TESS primary and extended missions. $I$ is the index of the true period in the list of all period aliases based on the TESS data alone. A value of $I = 0$ means that the true period is the shortest allowed alias.

Figure 8.5 shows that the average value of $I$ is $\sim 15$ and only 16 of 395 systems have $I = 0$. This strongly supports the need for the additional photometric and/or spectroscopic observations discussed in this paper. Even working up from the shortest periods (i.e. lowest $I$ values) it still requires an average of $\sim 14$ aliases to be ruled out before the true period is found. Using additional observations to reduce $N_{\text{aliases},i}$ and therefore reduce $I$ is vital. There have been some attempts at predicting the true period of a systems from a set of aliases using statistical methods [Dholakia et al., 2019] but, due to the statistical nature, the
reliability for individual system is poor thus an attempt to reduce $N_{\text{aliases},i}$ is still needed.

Another method for identifying the true period is to use the transit shape. Some attempts have been made to use a Bayesian inference based on the transit shape and an assumption of zero eccentricity to constrain orbital period to $\sim 10\%$ from a single K2 transit [Osborn et al., 2016]. Longer period systems however are more likely to have non-zero eccentricity [Halbwachs et al., 2005; Pont et al., 2011] therefore this method is not robust. Additionally, many of these systems will be discovered in the TESS 30 min cadence images [Cooke et al., 2018, 2019] where the transit shape is less well defined.

8.4.4 Improvements in number of aliases, $\Delta N_{\text{aliases}}$

As well as exploring the absolute number of aliases as a result of different amounts and methods of observations it is useful to explore how these values change. This is shown in Figure 8.6 which plots $\Delta N_{\text{aliases}}$ where $\Delta N_{\text{aliases}}$ is defined as

$$\Delta N_{\text{aliases}} = \frac{N_{\text{aliases},f} - N_{\text{aliases},i}}{N_{\text{aliases},i} - 1}. \quad (8.1)$$

In this formalism a value of $\Delta N_{\text{aliases}} = 0$ means that $N_{\text{aliases},i}$ has not been improved by additional follow-up data whereas a value of $\Delta N_{\text{aliases}} = 1$ means all incorrect aliases have been ruled out and the period of the system has been solved. The value can be interpreted as the fraction of aliases that each method can rule out.

To better show the change in the number of allowed period aliases these plots show all 395 systems, including both solved and unsolved systems. This leads to 2 peaks in each plot at 0.0 and 1.0. The peak at 1.0 corresponds to the solved systems, i.e. $N_{\text{aliases},f} = 1$. The peak at 0.0 corresponds to those systems where $N_{\text{aliases},f} = N_{\text{aliases},i}$. For photometry these are the systems for which NGTS cannot reach sufficient $S/N$ meaning the data cannot be used to detect transits and for spectroscopy these systems are those for which the spectrograph noise (CORALIE in these plots) is comparable to, or greater than, the RV signal of the planets meaning the period cannot be seen under the scatter. These peaks overwhelm the rest of the data and are truncated in the plots but are described fully below.

Once again we see a clear separation in the right hand column plots based on the detection of an additional transit. Those systems seen to transit again are almost all above 0.8 (meaning over 80% of the aliases are eliminated) with 149 being solved fully. For systems without an additional transit the fraction lies mainly between 0.1 and 0.7, the full distribution gives 297 systems ($\sim 75\%$) that have more than half their aliases eliminated. Ignoring the unsolvable systems (insufficient $S/N$) this is $\sim 76\%$.

Spectroscopy shows a slightly different distribution. The peak at 1.0 is much smaller than photometry (50 compared to 149) but the peak at 0.0 is larger (210 compared to 4).
Figure 8.6: Improvements in number of period aliases found when additional observational data is accounted for. Shown are distributions of $\Delta N_{\text{aliases}}$ for the three combinations of additional observations. The top left plot repeats the top left plot in Figure 8.2 to show the values of $N_{\text{aliases},i}$ by which the other three subplots are divided. A value of $\Delta N_{\text{aliases}} = 0$ means no improvement has been made whereas $\Delta N_{\text{aliases}} = 1$ means the system has been solved. As in Figure 8.2 the right hand column plots (those including additional photometry) are split into those systems for which an additional transit is seen (red) and those for which no additional transit is seen (blue). Here we assume a time-span of 50 days and use CORALIE for spectroscopy. The spikes at 0.0 and 1.0 overwhelm the rest of the data and are truncated here, the values are given in the text.
Exterior to the two peaks the spectroscopy distribution is then geared towards higher fractions with a larger clump of systems between 0.9 and 1.0. In fact, we see that almost all systems with sufficient $S/N$ have more than half of their aliases ruled out (93%). This is as expected since spectroscopy data continues to add information on each subsequent night whereas photometry data is less useful until catching an additional transit. The reason that comparatively few spectroscopy systems are able to go from only a few remaining aliases to completely solved ($N_{\text{aliases},f} = 1$) is that by this point only similar period aliases are left. This method of spectroscopic observations is efficient at ruling out significantly different periods but finds it very difficult to identify the correct period between two similar values. Photometry is less affected by this feature as an additional transit is, in general, very good as discriminating between two similar periods.

Once again, combining the two methods gives the best result with the overall improvement in fraction of aliases eliminated by spectroscopy being further improved by systems with additional transit from photometry. When using both methods only 1 out of 395 systems are not improved and 335 systems (~85%) have more than half of their period aliases eliminated.

### 8.4.5 Time taken vs nights observed

An additional aspect of this simulation that must be considered is that though our simulation is run for 50 days no system is observed for that many nights. Firstly, each system is only observed by each method until that method has managed to solve the system after which further observations are not simulated. Secondly, for spectroscopy, a system is only observed once every seven days, allowing for weather. This means that the number of nights for which a system is observed spectroscopically is approximately equal to $T \times 0.8$ where $T$ is either the time until the system is solved or the time-span of the simulation, whichever is smaller. For photometry a system is only observed when one of its period aliases predicts a transit will occur. This means that the actual number of nights observed can be a small fraction of the time over which the simulation is run. Additionally, we do not necessarily observe for entire nights at a time. For spectroscopy we use a noise model that assumes half-hour exposures so each simulated night of observing is only half an hour of spectroscopy time. For photometry we observe each night for the full length of any transit that is predicted to occur that night i.e. if only a partial transit is predicted we only simulate observing time during that window. As such the total number of telescope hours is different from the number of nights observed and we show the values here. Figure 8.7 shows plots of the number of hours that systems are observed for over the 50 days for which this simulation is run.

Additionally, for photometry we build in a buffer time based on the propagation of
Figure 8.7: Telescope hours utilised by each system for photometry and spectroscopy. Systems are coloured for solved (red) and unsolved (blue).

transit timing uncertainties from the TESS transits (as a bonus this buffer time will allow us to obtain out of transit baseline photometry, helping to improve transit characterisation). For each period alias the uncertainty is equal to $2\sigma_t/N$ where $\sigma_t$ is the uncertainty on the transit centre time of each of the TESS transits and $N$ is the integer value used to calculate each period alias. We choose $\sigma_t = 1\text{hr}$ for this calculation, twice the cadence for the TESS full Frame images which are the source of the majority of our targets. Propagating this uncertainty forward we find the uncertainty on successive transit times, $\sigma_T$, to be given by

$$\sigma_T = 2\sigma_t \left( \frac{1}{2} + \frac{n}{N} \right),$$  \hspace{1cm} (8.2)

where $n$ is the number of periods after the second TESS transit. For this we assume that our simulated observations begin approximately 100 days after the end of the TESS extended mission for the southern ecliptic hemisphere. This is likely an overestimation since we should be able to start following up targets from early TESS sectors while the extended mission is ongoing. For each estimated transit time for which we observe, our observation window is extended by this value to ensure transits are not missed due to inaccuracies in the TESS transit timing measurements.

In these figures we see that each system is observed for a significantly smaller number of hours than are available within the observing span of 50 nights. Each method is split into 2 categories: those which are solved ($N_{\text{aliases, f}} = 1$, red) and those which remain unsolved ($N_{\text{aliases, f}} > 1$, blue). Systems with insufficient $S/N$ are not shown. For photometry the average number of hours observed is 62.8 and 114.2 for the solved and unsolved systems respectively (ignoring those systems for which we have insufficient $S/N$ that are not observed at all). The use of multiple telescopes is accounted for when calculating tele-
scope hours. This highlights the efficacy of a targeted photometric campaign as opposed to a stare campaign, where we simply wait for an additional transit. For spectroscopy we see the number of hours observed is 2.8 and 3.1 for solved and unsolved systems respectively. Based on the strategy of one half-hour every seven days, combined with the weather constraints we expect an average unsolved value of approximately \( \sim 3.3 \) which matches what is seen.

Looking at these plots combined shows us that the majority of solved systems are solved with the first 5 hours of spectroscopy time or the first 80 hours of photometry time. Therefore, if attempting to most efficiently solve the largest number of systems, the ideal strategy may be to observe each system for 10 nights (assuming 0.5 hours per night for spectroscopy and an average of 8 hours per night for photometry) and if not solved move on to a different system. However, this neglects the fact that some systems are of greater scientific interest than others. Therefore, it actually requires a more careful consideration of individual system parameters to determine which systems may be worth a larger fraction of available telescope time.

### 8.4.6 Effect of increasing follow-up time

The plots and analysis so far have used a time-span of 50 days and CORALIE for spectroscopy. Table 8.1 shows how the \( N_{\text{aliases,f}} \) values change for different combinations of time-span and spectroscopy instrument. However, it is also interesting to look in more detail at how the number of solved systems increases as the time-span is allowed to increase. This is shown in Figure 8.8.

Figure 8.8 shows that the fraction of systems that can be solved using either photometry or spectroscopy is a function which will asymptote before reaching unity. The reason for this is as explained above, systems with sufficiently small transit or RV signals will become lost in the noise of their respective instrumental precisions and will thus never be solved.

Initially we see that the number of systems solved through photometry rises quite slowly. This is because, in the early stages of the follow-up campaign, the coverage is sparse enough that only catching an additional transit will solve a system fully. Also, because there has not been time for other aliases to be ruled out by non-detections, only systems for which the observed transits produce no aliases will solve a system. As the observations continue the rate of solving begins to increase. This is due to the fact that at this stage, we are beginning to see enough aliases ruled out through nights of non-detections that catching a third transit is more likely to remove all aliases that are left. Additionally, the likelihood of a fourth transit being seen increases which dramatically reduces the number of aliases.

Spectroscopically solved systems only begin to show up around 28 days since we
require at least 5 data points, taken once every 7 days (including a data point at time zero). As we see in Figure 8.7(b) there is a spike in solved systems upon reaching this threshold, however, due to the randomness of weather effects different systems reach this threshold on different nights with 28 simply being the earliest possible time. The slow increase from here on is then the result of the difficulty of using spectroscopy to fully solve a system. Were we to focus on systems for which spectroscopy can rule out 80% of aliases this peak would be much higher but it is challenging for spectroscopy to definitively decide between similar periods.

The number of systems solved by combining the two methods is always above those solved using either method alone as expected. However, as time increases the fractional addition in number of solved systems caused by this combination above those solved by photometry alone is reduced. As seen above, at 50 days the total number of systems is increased from 149 to 207, 39%, but by 100 days the increase is from 240 to 305, only 27%. This seems to be due to the fact that by this time, the number of systems solved by spectroscopy is increasing only very slowly, therefore spectroscopy is adding little to the number of overall solved systems whereas the systems solved via photometry are still increasing, albeit slower than they were originally, due to the shorter and easier systems
having been solved already.

It is also of interest to note how many transits are required for a photometric solution. It is seen that for the majority of systems (122/149) only one additional transit is required to fully solve the period. The remaining solved systems are done so with two additional transits. This matches what we see for photometry in Figure 8.8. The initial quick increase in solved systems is a result of those systems that can be solved by catching only one more transit. This then levels off as we approach the longer period systems which require a larger number of additional transits to solve.

8.5 Conclusions

Catching two widely separated transits of a long period transiting system is insufficient for identifying the true period of the orbit. We have taken the TESS primary and extended mission observations of the southern ecliptic hemisphere as an example and showed that there are approximately 395 exoplanets with a single transit detected in each year. Based on the TESS data alone we show this leaves an average of \( \sim 38 \) period aliases per system with the true alias being, on average, the 15\(^{th}\) smallest possibility. Building on this, we have simulated attempts to mitigate these aliases using additional photometric and spectroscopic data.

We find that the amount of additional photometric and/or spectroscopic observations taken can have a significant impact on both the number of fully resolved systems (defined here as systems with \( N_{\text{aliases}} = 1 \)) and the mean number of aliases for the unsolved systems. The results presented in Figures 8.2 though 8.8 and Table 8.1 can be summed up by the following points.

1. For all simulation lengths photometry fully solves the period of significantly more systems than spectroscopy; roughly three times more for a 50 day simulation (Figures 8.2 and 8.8).

2. Combining both methods results in the highest total number of solved systems. At 50 days 13\% of these were not solved by either method independently (Figure 8.2).

3. When selecting only those systems which reach the required \( S/N \) threshold for each method photometry rules out at least half of the period aliases for 76\% whereas the corresponding value for spectroscopy is 93\% (Figure 8.6).

4. Spectroscopically solved systems favour higher mass planets which leads to a peak between 10 and 15\( R_\oplus \) which corresponds to the largest mass planets. Additionally it is easier to solve shorter period systems (Figure 8.4).
5. Photometrically solved systems are insensitive to mass or radius assuming they exceed the S/N threshold. As for spectroscopy, shorter period systems are easier to solve (Figure 8.4).

6. Systems solved through a combination of the two methods tend to have a broader period distribution, extending to longer period systems (Figure 8.4).

7. For an observing period of 50 days the majority of systems solved by photometry are done so in under 80 hours of telescope time. All spectroscopically solved systems are solved in under 5 hours (Figure 8.7).

8. We suggest that, with the singular goal of confirming the period for the greatest number of systems, a 50 day follow-up campaign is preferred when accounting for the cost-benefit ratio. However, if attempting to characterise more scientifically interesting systems different campaign lengths may be more efficient.

9. Longer observing periods lead to more solved systems but both methods asymptote before solving all systems (Figure 8.8).

10. Using HARPS instead of CORALIE leads to an increase in the number of spectroscopically solved systems by a factor of 3-5 which decreases as the observing period increases (Table 8.1).

8.A Number of aliases
Table 8.1: $N_{\text{aliases,f}}$ results from 10 combinations of time-span and spectroscopic instrument. For each combination of observations we show the number of solved systems. $N_{\text{aliases,f}}$ then refers to the mean number of aliases for all unsolved systems. The bold row matches the data plotted in Figures 8.2, 8.6 and 8.7. Data for the TESS only simulation are not shown as they are unchanged by follow-up strategies.

<table>
<thead>
<tr>
<th>Time-span (days)</th>
<th>Spectroscopy</th>
<th>Solved (phot)</th>
<th>$N_{\text{aliases,f}}$ (phot)</th>
<th>Solved (spec)</th>
<th>$N_{\text{aliases,f}}$ (spec)</th>
<th>Solved (specphot)</th>
<th>$N_{\text{aliases,f}}$ (specphot)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>CORALIE</td>
<td>86</td>
<td>19.2</td>
<td>9</td>
<td>33.7</td>
<td>104</td>
<td>18.4</td>
</tr>
<tr>
<td>30</td>
<td>HARPS</td>
<td>90</td>
<td>18.5</td>
<td>41</td>
<td>31.9</td>
<td>129</td>
<td>16.6</td>
</tr>
<tr>
<td>40</td>
<td>CORALIE</td>
<td>118</td>
<td>16.8</td>
<td>21</td>
<td>29.4</td>
<td>156</td>
<td>15.8</td>
</tr>
<tr>
<td>40</td>
<td>HARPS</td>
<td>131</td>
<td>16.7</td>
<td>94</td>
<td>23.4</td>
<td>220</td>
<td>14.1</td>
</tr>
<tr>
<td><strong>50</strong></td>
<td><strong>CORALIE</strong></td>
<td><strong>149</strong></td>
<td><strong>14.2</strong></td>
<td><strong>50</strong></td>
<td><strong>25.7</strong></td>
<td><strong>207</strong></td>
<td><strong>11.6</strong></td>
</tr>
<tr>
<td>50</td>
<td>HARPS</td>
<td>150</td>
<td>13.7</td>
<td>191</td>
<td>17.9</td>
<td>298</td>
<td>8.8</td>
</tr>
<tr>
<td>60</td>
<td>CORALIE</td>
<td>170</td>
<td>11.9</td>
<td>57</td>
<td>22.5</td>
<td>242</td>
<td>9.6</td>
</tr>
<tr>
<td>60</td>
<td>HARPS</td>
<td>180</td>
<td>12.8</td>
<td>232</td>
<td>17.4</td>
<td>323</td>
<td>7.9</td>
</tr>
<tr>
<td>70</td>
<td>CORALIE</td>
<td>187</td>
<td>10.8</td>
<td>82</td>
<td>24.9</td>
<td>254</td>
<td>10.2</td>
</tr>
<tr>
<td>70</td>
<td>HARPS</td>
<td>198</td>
<td>10.9</td>
<td>243</td>
<td>17.5</td>
<td>340</td>
<td>7.5</td>
</tr>
</tbody>
</table>
Chapter 9

TESS monotransit search and NGTS follow-up

Note

The following section details the work of the NGTS monotransit working group in identifying and characterising TESS monotransits as of February 2021. The work in this chapter is the result of group effort of the whole NGTS working group. My specific contributions are detailed in the following sections. Section 9.2.2: adaptation and application of a monotransit identification tool to run on TESS data (the original code was written by Hugh Osborn). Section 9.2.3: The majority of the eyeballing stages were carried out by myself with the best candidates being discussed with the full group. Section 9.3.3: I contributed significant amounts of eyeballing of the final data products, highlighting the best targets for NGTS follow-up, additional eyeballers operated in parallel. Section 9.4: I adapted and ran the Namaste code on NGTS monotransit working group targets (the original code was written by Hugh Osborn). Sections 9.5, 9.6 and 9.7: I contributed as part of the discussion, planning and analysis of the NGTS follow-up efforts, including archival photometry and RV data. Section 9.8: I contributed written sections of a number of the publications resulting from the work described in this chapter.

9.1 NGTS monotransit working group

The NGTS monotransit working group is a sub-group within the NGTS consortium [Wheatley et al., 2018]. Its remit is to characterise the properties of planetary candidates that exhibit a single transit in TESS with a particular emphasis on identifying the true period of the system. At its inception, before the confirmation of the TESS extended mission, the mono-
transit working group was limited to systems that showed just a single transit. In its current form however, systems which transit once in the primary and once in the extended mission are included, since, crucially, two widely separated transits are insufficient for determining period (see Chapter 8). NGTS is particularly well suited for this task since it can reach a precision comparable to TESS but has significantly higher cadence (~ 12 s compared to 2 or 30 mins for the TESS primary mission). Another advantage is that NGTS is comprised of 12 separately operable telescope meaning it can follow-up multiple targets simultaneously. Additionally, the large number of telescopes means that it has the availability to observe a target continuously while waiting for a second transit which other observatories cannot afford to do. The working group targets monotontransits with radii around 1.5 R_J or less (to limit the inclusion of EBLMs, except where desirable) and can operate in two regimes, a blind survey or an alias targeted approach. The two regimes are discussed in sections 9.5.1 and 9.5.2.

9.2 Difference Imaging Analysis light curves

The first set of TESS FFI light curves used to detect monotransit candidates are the Difference Imaging Analysis (DIA) light curves produced by Oelkers and Stassun [2018]. The light curve generation and subsequent monotransit search is detailed below.

9.2.1 DIA light curve generation

The DIA TESS FFI light curves are described in Oelkers and Stassun [2018] and Oelkers and Stassun [2019]. These light curves are produced for the first 5 sectors of TESS data based on the public FFIs. A light curve is generated for every TIC star that falls into the TESS field of view in the first 5 sectors using a Difference Imaging Analysis (DIA) pipeline. A DIA approach is chosen in this case due to the large TESS pixels (21") and crowded fields which make simple aperture photometry difficult as it is non-trivial to remove contamination from nearby sources, even when using small aperture sizes [Alard and Lupton, 1998; Miller et al., 2008]. DIA builds on the idea that stars with no astrophysical variation should not change between frames whereas variable stars will. The process then involves taking a single frame as a reference and subtracting it from subsequent frames. Non-variable stars will subtract cleanly with the only difference being on the level of Poisson noise. Variable stars however will show residuals in the subtracted frame. Aperture photometry can then be used to extract the residuals from the difference image, creating a light curve that avoids contamination from surrounding targets. Before subtraction the reference image must be blurred to match the seeing conditions of the subsequent frame. This is usually done assuming a Gaussian PSF but this can result in poor subtractions if the true
Figure 9.1: Science frame and corresponding difference frame generated using the DIA pipeline. The majority of stars are subtracted cleanly leaving only Poisson noise. The red arrow highlights an area of correlated residuals identifying a variable star. Taken from Figure 3 of Oelkers and Stassun [2018].

PSF is non-Gaussian [Bramich, 2008; Oelkers et al., 2015]. Oelkers and Stassun [2018] use a Dirac $\delta$-function kernel to blur the reference image which works better for arbitrary PSFs.

The pipeline used here is a modified version of one first designed to run on data from the Chinese Small Telescope ARray [CSTAR, Zhou et al., 2010; Oelkers et al., 2015]. First, background models are created for each frame by sampling the image in $32 \times 32$ pixel sections and replacing bright pixels by the median value. The background images are then subtracted from each frame [Oelkers and Stassun, 2018]. All frames must then be carefully aligned so as to reduce residuals due to misaligned stars. TESS pointing precision is less than 1 pixel but spacecraft jitter must still be accounted for [Oelkers and Stassun, 2019]. This is done by using the world coordinate solutions (WCS) from the headers of the FFIs. The next step would generally be generating a master reference frame to subtract from the science frames to determine the residuals. In previous pipeline iterations this was a median combination of the highest SNR science frames. In the final pipeline however, a minimal loss of quality was found when using the first image rather than the median average image so this strategy was adopted [Oelkers and Stassun, 2019]. Since the TESS PSFs can vary significantly across a frame they can become quite non-Gaussian. Therefore the pipeline uses a Dirac $\delta$-function kernel based on the measured PSFs of the 500 brightest, isolated stars across each frame to blur the reference image to the appropriate degree. The reference is then subtracted from each science frame leaving the difference image [Oelkers and Stassun, 2018]. Figure 9.1 shows an example science frame and corresponding difference frame.

Finally, the pipeline conducts some measure of detrending on the output light curves.
The 100 closest, similar magnitude stars (±0.1 mag) were selected, their light curves were weighted and median averaged and this master trend was removed from each target light curve. This detrending was not carried out for the most contaminated targets. If the sum of the fluxes of all stars within 7 pixels was greater than 5 times the flux of the target (defined as a contamination ratio > 5) a target was not detrended [Oelkers and Stassun, 2019]. These light curves were then made available via https://filtergraph.com/tess ffi.

9.2.2 Monotransit identification

Every DIA TESS FFI light curve with a contamination ratio less than 5 was downloaded as plain text files ready to be searched for monotransit candidates. This led to approximately 1 million targets per sector. Before the search began two processes were applied to the data. First the data was sigma clipped to remove large outliers caused by instrumental effects. The sigma clipping package from astropy [Astropy Collaboration et al., 2013] was used which removes any data points further than three standard deviations from the median value. The next step was to apply a light curve flattening module in an attempt to remove any variability that the DIA light curve detrending had missed. This includes a normalisation step. Flattening the data first was found to greatly reduce the number of false positives and spurious detections. The flattening software used was the Lightkurve package [Lightkurve Collaboration et al., 2018], a light curve analysis tool designed for the Kepler, K2 and TESS missions. The flattening procedure used is based on the Savitzky-Golay filter approach [Savitzky and Golay, 1964]. Individual TESS orbits (identified by gaps in the photometric data) are treated separately since some TESS sectors are seen to show different variability between orbits.

The monotransit search then follows the procedure originally described in Osborn et al. [2016] which is laid out in section 2.1.2. The large number of light curves per sector, ~1 million, meant the procedure must be run in parallel with the input light curve files separated into 12 subsets run simultaneously. The outputs of the search were saved to text files listing the TIC ID of the source light curve, the epoch of the detection and the key detection parameters as discussed in section 2.1.2.

9.2.3 Candidate eyeballing

Once the list of monotransit events is generated they must be examined to determine the reliability of the signals and for the best candidates to be scheduled for further follow-up. Before looking at the event individually some cuts must be made to leave a more manageable number. The outcome from the automated search left on the order of 1 million events per sector. This list was the cut down based on three criteria. First, the depth. The
selected events must be deep enough to be observable by NGTS and to be believable in the TESS data. Additionally, events should not be too deep that the probability of them being planetary in origin is small. The chosen range for this iteration of the searching was between 0.5% and 10%. 10% would allow EBLMs to be selected but would ensure that most giant planets around dwarf stars were not missed. The second criteria was SNR which was defined as the ratio of signal depth to the standard deviation of the light curve flux post flattening. The threshold chosen here was SNR≥2.0. Finally, certain systematic epochs of each light curve must be masked. This was done on a sector-by-sector basis. A histogram of all detections as a function of epoch was created and then regions that generated peaks in this histogram were ignored. This would result in some real detections that happened to fall at systematic epochs being removed but this was preferable to the vast increase in false positives caused by keeping these regions. Generally these regions were the start and end of each orbit with occasional other regions corresponding to telescope issues such as the loss of pointing accuracy which may inject spurious signals into multiple light curves at the same epoch. After these cuts the number of events was on the order of 10,000, a removal of approximately 99% of detections. Applying the cuts in the order described here reduced the total number of events by factors of 0.5, 0.2 and 0.1 respectively.

The surviving events would then be converted into an easily viewed form which involved additional cuts. The first step was to access the TIC catalogue to determine the stellar host properties. To facilitate observations by NGTS a magnitude cut was made requiring $M_V < 13$. Additional cuts were made based on declination to ensure the target was visible from NGTS. Next would be a filter based on number of events per sector. It was found that large numbers of events in a single sector indicated a problem with the light curve and not a true signal. Since the goal of the search was monotransits anyway this was a useful cut. Light curves with more than 5 events were rejected. For each event the TIC catalogue values would be used to convert the depth of the event into an inferred planetary radius. Additionally, flags were made for events that corresponded to known TOIs or CTOIs, whether they had TESS light curves in other sectors and whether they had a corresponding 2 minute light curve. Finally, the light curves were plotted and separated based on inferred companion radius with a dividing value of $1.5R_{\text{Jup}}$.

The final light curve plots were then eyeballed and the best looking candidates sent forward for additionally vetting including checking the field for blends and looking for similar signals in surrounding stars. Corresponding light curves were made using other methods such as the dedicated eleonor package [Feinstein et al., 2019] and used as references. If they existed, the corresponding TOI/CTOI reports and 2 min light curves would be examined. Any other sector in which the host was observed had the relevant light curve examined as well. Finally the best events would be modelled and a prediction of period
would be made using Namaste (see section 9.4). The surviving candidates would then be scheduled for NGTS follow-up (see section 9.5).

9.3 NGTS light curves

As the TESS mission continued and the NGTS monotransit working group continued to search for monotransit candidates a new, more bespoke search tool was needed. Additionally, since the DIA light curves were only created for sectors 1 to 5 a new procedure for creating light curves from FFI images was needed. To solve these problems a novel tool was created to generate FFI light curves and search them for monotransit candidates. This method is briefly discussed in section 2.1.2 but is expanded upon here.

9.3.1 NGTS light curve generation

The TESS light curves are generated directly from the full frame images available from the Mikulski Archive for Space Telescopes\(^1\) (MAST). FFIs are available via MAST around 1 month after the end of TESS observations for a sector. From here, the calibrated images are downloaded and must be converted in to light curves for individual stars (the calibrated images have had the dark current and cosmic ray effects removed). Aperture photometry is used to create light curves for every star in the TESS input catalogue. 2.5 pixel apertures are the chosen standard based on PSF observations but other apertures are also used. The aperture positions are determined using the World Coordinate System (WCS) data in each image (accurate to 0.1 pixels or better) with a fixed annulus location between images for background subtraction. Light curves are then detrended using a set of Co-trending Basis Vectors (CBVs) similar to the method used on Kepler data [Kinemuchi et al., 2012]. These CBVs are individually created for each combination of sector, camera and CCD to ensure any differences between them are accounted for. The light curves are also flattened using a Savitzov-Goly filter [Savitzky and Golay, 1964] and normalised. The created light curves are presented in a front-end web interface allowing for much easier access compared to the equivalent DIA light curves. These light curves are updated as new TESS sectors are made public.

9.3.2 Monofind

After creating the relevant light curves they must be searched for monotransit events, this is done using a bespoke pipeline called monofind. In the first instance, all light curves down to a magnitude limit of 11 are searched for events. This search involved detecting

\(^1\)http://archive.stsci.edu/tess/bulk_downloads/bulk_downloads_fli-tp-lc-dv.html
consecutive points that are significantly lower than the rest of the light curve. In this case this relevant threshold is defined as 3 time the Mean Absolute Deviation (MAD) as opposed to using some function of the standard deviation of the data. The MAD is beneficial here as it is more robust against outliers, such as drop outs, which change the flux drastically, or repetitive drops, such as a short period planet. Regions of the light curve which show 3 or more consecutive points below this threshold are flagged as candidate events. An additional step is to locate the lowest point in each flagged region and ensure the points either side are also below the threshold. There is no limit to how many potential events can be found in a single light curve, only that they do not overlap. Based on the distribution of events with epoch bespoke sector masks are made, similar to those made in section 9.2.3, where epochs showing spikes in number of events are masked from the results. With the masks made the rest of the light curves down to a magnitude of 14 are searched using the same method. For each non-masked event a transit model is then fitted, using a free impact parameter and a power-2 limb darkening law [Maxted and Gill, 2019], in an attempt to provide realistic transit parameters and robustness against eclipsing binaries which are generally more v-shaped. The stellar data used for the fits are pulled from the TESS input catalogue (currently version 8 [Stassun et al., 2019]). The fit takes into account dilution values which are estimated empirically for each target. The resulting stellar data, as well as the transit fit parameters are stored in a database as well as being presented alongside the transit event light curves in the web interface. It should be noted that alongside the monofind algorithm a BLS search is also run to identify repeating signals. The results of this are presented alongside the monofind results but neither influences the other.

9.3.3 Monofind eyeballing

The eyeballing of the monofind candidates is done through the web interface. For each combination of sector/camera/CCD a web page is shown containing all the light curves flagged as candidates alongside the key stellar and fitted parameters. Selecting a specific candidate opens a more detailed page with all of the data regarding the candidate and its flagged event. When eyeballing, an SQL query can be used to select only those candidates which match certain criteria. Any parameter from the TIC, the fitting procedure, or from external data can be sorted on. External data includes information such as TOIs/CTOIs on the same target. Built into the interface are a number of tools to help determine the validity of the candidates. Firstly, candidates are given flags based on whether they were detected as monofind events, BLS events or both. Additionally, the light curves are analysed for asteroid signatures. Asteroids can present as transit signals and these must be identified. When an asteroid passes through the region of the image used to calculate the background subtraction (as part of the aperture photometry) it causes a peak in the flux. This means that
a larger flux value is subtracted from the target light curve resulting in a dip that looks like a transit. Events that coincide with peaks in the background flux are flagged as possible asteroid events. For some sectors a catalogue of known asteroids in TESS (the TSSYS catalogue [Pál et al., 2020], not available for all sectors) is used to check if known asteroids match up with events both spatially and temporally. Each candidate page also has a link to both the relevant WASP archive page (if it exists) and the corresponding page of the TESS Follow-up Observing Program (TFOP) working group [Collins et al., 2018; Akeson and Christiansen, 2019]. This allows an easy method to check if detected events have already been flagged as planets/EBLMs or TOIs/CTOIs. To try and account for blended signals events that are closely matched in physical space, as well as occurring at similar times, are flagged and displayed as a cluster of events. Viewing these together can reveal transits on multiple nearby stars which are then most likely evidence of a background eclipsing binary being blended with multiple foreground stars. Finally, plots of the centroiding information are shown alongside the candidate transit events. Centroiding is a procedure to measure the central location of the measured flux on the CCD. If the centroid of a target moves simultaneously with its flux decreasing the chances are good that the actual event is on a different host.

Since a large goal of the current iteration of the NGTS monotransit working group focuses on TESS objects with a transit in both the primary and extended missions each candidate page contains a link to the Web TESS Viewing tool\(^2\) (WTV). This is an online tool that check to see when a target has been, and will be, observed by TESS based on its coordinates. Recently another tool has been used for the same purpose [Pál, 2020]. With this information known, a focus can be made on candidates which have one sector of coverage in each of the primary and extended TESS missions. The best candidates are flagged by individual users and then discussed as a full group. Surviving candidates are then scheduled for follow-up.

### 9.4 Namaste

As discussed in section 1.2.1 it is possible to make inferences about the period of a transiting system based on a single transit. Since the NGTS monotransit working group may have to dedicate significant amounts of telescope to follow-up a target it is best to have some idea of the probable period of the system first. One tool used as part of the working group to do this is Namaste [Osborn et al., 2016]. Namaste avoids the problems of determining planet velocity from the orbital period by instead estimating the planetary velocity scaled to the stellar radius, \(v'\). \(v'\) can be estimated from the radii ratio, impact parameter and transit

\(^2\)https://heasarc.gsfc.nasa.gov/cgi-bin/tess/webtess/wtv.py
duration using equation 1 from Osborn et al. [2016], reproduced here;

\[ v' \equiv \frac{v_p}{R_*} = \frac{2}{T_D} \sqrt{(1 + R_p/R_*)^2 - b^2}. \] (9.1)

Radii ratio, impact parameter and transit duration can all be estimated from fitting the transit shape. Namaste then estimates posterior probability distributions using the fitting regimes of Ian Crossfield\(^3\) and the emcee package [Foreman-Mackey et al., 2013]. Quadratic limb darkening is used (the NGTS monotransit working group updates this to use TESS bandpass parameters) with Gaussian errors based on the host temperature probability distribution.

Assuming \( e = 0 \) the circular period, \( P_{\text{circ}} \), is given by equation 2 from Osborn et al. [2016];

\[ P_{\text{circ}} = \frac{8\pi^2 G \rho_*}{3} \frac{R_p}{v'^3} = \frac{2\pi}{R_* v'^3}. \] (9.2)

Eccentric periods can be accounted for but the fitted eccentricity value turns out to be degenerate with period thus Namaste generally only considers \( e = 0 \) fits. Since longer periods are less likely due to transit geometry concerns (see section 1.2.1) they are given a reduced probability of occurring by enforcing a linear prior on transit velocity \( (P(v) \propto v) \). Additionally, the photometric coverage can be used to determine a lower limit on period and thus an upper limit on velocity. An example of the Namaste output probability distributions is shown in Figure 9.2 for Kepler-51d [Masuda, 2014]. For a more complete and detailed description of the Namaste pipeline see Osborn et al. [2016] and Osborn [2017].

Generally, because on the uncertainties from Namaste and the time elapsed since the original transit, this technique cannot be used to target a specific additional transit. However, this technique was used to preferentially rank monotransit candidates that had shorter periods for NGTS follow-up (particularly when operating in the blind survey mode) since longer period systems require more telescope time before catching additional transits.

### 9.5 NGTS follow-up

After promising monotransit candidates are identified they are scheduled for NGTS follow-up. This follow-up operates in two regimes depending on the amount of data available on each target. If the only available data is a single TESS transit then the method is the blind survey mode. If, however, additional data is known, such as a second TESS epoch, or archival transits, then specific transit times may be targeted, reducing the telescope time.

\(^3\)https://www.lpl.arizona.edu/ianc/python/
Figure 9.2: The output posterior distribution of Namaste for Kepler-51d. The period estimation is within 10% of the true value. Taken from Figure 3 of Osborn et al. [2016].
required (as discussed in Chapter 8).

9.5.1 Blind survey

The blind survey is employed when the only observational data on a target is a single TESS transit. Aside from potential lower limits on period from the TESS coverage and a possible Namaste prediction (see section 9.4) there is no knowledge of the period. Additionally, the follow-up may be happening a significant time after the TESS transit meaning period probabilities with large errors, such as those from Namaste, cannot help predict a future transit time. In these scenarios the only way to determine the period through photometry is to catch another transit which requires observing the target until a transit is seen. This method is obviously time consuming since the telescope cannot switch to other targets in case a transit is missed. Additionally, the target can only be observed during night so transits during the day are missed. NGTS is capable of this since it has multiple telescopes therefore, using one to observe in a blind survey leaves 11 that can work on different science cases. Where possible this strategy is preferably used when Namaste or other methods suggest a period that is not too long since longer periods increase the time until additional transits are caught.

This strategy has been used for monotransit candidates detected in the year 1 TESS data before TESS returned in the extended mission or in the cases that the extended mission will not reobserve a target. The method was successfully used to catch a second transit of the 61.8d eclipsing binary result presented in Gill et al. [2020b]. In the case of NGTS these targets are followed up by centering a telescope on the target in question. Due to NGTS’s large field of view however, these fields encompass many stars and therefore have a secondary purpose as general survey fields as part of the independent NGTS effort of planet discovery [Wheatley et al., 2018].

9.5.2 Alias chasing

Alias chasing is used when some prediction of the time of future transits is known. These predictions rely on additional observational data which can come from a range of sources. Having a second transit epoch, additional to the TESS discovery transit, allows for the determination of a discrete set of period aliases. These aliases then translate to a set of future potential transit times which can be individually targeted, saving telescope time when compared to a blind survey. Additional transits can come from the TESS extended mission [e.g. Gill et al., 2020c], NGTS follow-up in the blind survey mode [e.g. Gill et al., 2020b,c], or from archival photometry [e.g. Gill et al., 2020a]. The photometric coverage can then be used to reduce the potential aliases further. Estimates of future transit times can also come
from RV measurements which can be modelled to give a set of period aliases [e.g. Lendl et al., 2020b].

Once future transit times are predicted the NGTS strategy is to only observe at these times. Depending on the uncertainty on the transit times, or how much of a transit is predicted to occur during night, the strategy may involve observing for the full transit night, or even observing for a few nights around the time of predicted transit to ensure the event is not missed. NGTS can also employ multiple telescopes on a single target. These leads to enhanced precision when all telescope light curves are combined [Bryant et al., 2020a]. However, due to telescope time concerns, this is obviously only applicable during targeted observations, not the blind survey mode.

### 9.6 RV observations

Radial velocity data are a useful compliment to photometric observations of monotransit candidates. In the first instance radial velocity points can be used to identify some of the key sources of false positives for monotransits, namely eclipsing binaries. Taking two radial velocity points can help to determine if there is a large scale signal incompatible with the detected companion being of planetary mass. Using equation 1.7 shows that a Jupiter mass planet on a 10 day orbit will induce a radial velocity signal of $\sim 100 \, \text{ms}^{-1}$ on a solar mass host. If the difference in consecutive radial velocity points is significantly larger than this the implication is that the companion must be much more massive than Jupiter and therefore more likely to be a stellar object than a planetary one. Obviously this approach is not foolproof as there are other effects that may change the radial velocity such as a linear drift and large amplitude signals may be measured at points in their sinusoids where the measured values are not too different (this is especially hard to account for when a period is unknown). That being said however, this is a good first step. This process requires a spectrograph of sufficient precision to separate between the two regimes but one that is not too oversubscribed as to be unavailable for recon spectra. The NGTS monotransit working group often employs CORALIE recon spectra for this purpose. CORALIE has a precision of a few ms$^{-1}$ making it ideal for these sort of observations. Radial velocity data can also help rule out false positives in the special case of double-lined spectroscopic binaries. These are binary systems in which both stars are sufficiently bright as to have measurable spectra. Depending on the orientation of the stars in their orbits a spectrograph may measure a spectrum from each component. When this happens it is a confirmation that the companion is of a stellar, rather than planetary, nature.

Beyond its use as a tool for detecting false positives radial velocity instruments can be used for the characterisation of true monotransiting planets. Sufficient radial velocity
coverage can allow the period of a planet to be measured and, when combined with photometric data, this technique becomes even more powerful. Additionally, radial velocity data can be obtained even on nights which are of insufficient quality for photometry. This process is displayed in two recent papers. Firstly in Lendl et al. [2020b] where a monotransit candidate discovered in TESS data was observed using CORALIE. CORALIE’s first observations revealed a 10 kms$^{-1}$ drift over 9 days, confirming the system as an eclipsing binary. Further data, totalling 13 points in all, allowed for the determination of a 33.9 day period which was later confirmed with additional observed transits from NGTS and Euler-Cam [Lendl et al., 2012]. In a separate paper, Gill et al. [2020c], RV data were used to constrain the period of a monotransiting planet from TESS. Three initial CORALIE points were taken and confirmed no variations above 80 ms$^{-1}$ suggesting a planetary mass companion. An additional 9 HARPS points and 6 FEROS points, when combined with the TESS and NGTS transits, allowed for the true period of 35.46 days to be determined.

9.7 Archival photometry

Since many TESS targets are fairly bright a large proportion of them have some degree of previous photometric data. Depending on the quality of this data it may be possible to locate previous transits once a system is known to host a transiting planet. These events may have been missed when the original data was taken due to low SNR but may stand out more clearly when a template of the signal is generated from the higher quality TESS photometry. The archival data employed in the NGTS monotransit working group is usually that taken by WASP since the group has good access and experience with the data and the all-sky nature of the WASP survey means it is likely to have observed a large fraction of the TESS targets.

To search the WASP data a transit template is generated based on the single TESS transit. The model is generated using an MCMC with a fixed 30 day period (based on the length of a single TESS sector) and $R_*/a$, $R_p/R_*$ and $b$ are allowed to vary. The best fitting parameters are then used to generate the transit template. This template is then placed at every time step in the WASP data set and a $\chi^2$ value is found at each point based on the agreement between data and model as a function of transit midpoint. Defining a reference value, $\chi^2_{\text{ref}}$, based on the weighted mean of the full WASP data set, the $\chi^2$ is converted to a log likelihood value using the relationship

$$\Delta \log L = -\frac{\chi^2 - \chi^2_{\text{ref}}}{2}. \quad (9.3)$$

For full details of this formalism see Gill et al. [2020a]. Any points above a certain
threshold are flagged as potential precovery events. Once verified, these events allow additional constraints on period to be made. The technique was used effectively in Gill et al. [2020a] to constrain the period to one of 3 aliases. Figure 9.3 shows the precovery efforts for this system.

9.8 Results

The work produced by this candidate identification and follow-up effort has resulted in multiple publications, some of which are discussed here.

9.8.1 EBLMs

The first published results were the discovery of low mass eclipsing binaries in the FFI light curves. When looking at single transits, EBLMs can look very similar to planetary transits. This was a particular concern during the candidate identification in the DIA light curves because at this point the method was rather unpolished and would identify vastly more false positives than legitimate planet signals. These signals could be identified as EBLMs using RV follow-up as discussed above but, initially these were still observed by NGTS both as a proof of concept of the monotransit follow-up method and because, by definition, these EBLMs are long period and thus of scientific interest. The first success from this project was a 38 day F/M low mass eclipsing binary [Gill et al., 2020a]. This target was first identified from a single transit in the DIA light curves before archival WASP data was used to constrain the period to one of three values. NGTS was then used to target the potential transit times and caught a transit, confirming the period. CORALIE RVs were used to determine the mass.
This publication was the result of targeted NGTS observations but the blind survey mode was also successful in making discoveries. Gill et al. [2020b] report the discovery of another binary system, this time an M5V dwarf eclipsing a sun-like star on a 61.8 day period. Initially identified as a candidate from a single DIA light curve transit this target was monitored in a blind NGTS survey since no sufficient quality archival photometry was available. After catching a second epoch with NGTS a subset of alias were identified and chased with the confirmation coming from a transit caught by LCO [Brown et al., 2013].

A third success of the NGTS monotransit team came in the form of TOI-222, a 34 day period eclipsing binary [Lendl et al., 2020b]. This was slightly different as it was first identified as a candidate by the TESS team and then observed using RV data. CORALIE RVs managed to identify a set of aliases which NGTS was able to target to help confirm the true period.

9.8.2 NGTS-11

The first exoplanet to be discovered after being identified from a TESS single transit was NGTS-11b [Gill et al., 2020c]. NGTS-11b was initially identified as a candidate from the DIA light curves of TESS sector 3. Following this it was observed in a blind survey by NGTS until a second transit event was caught. The true period was then known to be a factor of the separation between the two transits. The coverage provided by the TESS data, as well as the NGTS blind survey data, were able to rule out multiple aliases leaving 13 possibilities. As soon as the second transit had been observed NGTS-11 was put forward for a CORALIE RV campaign which collected 3 data points. These points ruled out a stellar companion and motivated 15 more data points, 9 with HARPS and 6 with FEROS [Kaufer et al., 1999]. Joint fits of the photometric and spectroscopic data, along with the known aliases, identified a best fitting orbital solution of 35.46 days with marginal eccentricity. NGTS-11b is a Saturn like planet with a radius of 0.817 R_J and a mass of 0.344 M_J. Figure 9.4 show the TESS and NGTS transits, both binned to 30 mins.

NGTS-11 was re-observed as part of the TESS extended mission in sector 30 and a second transit of NGTS-11b was seen, confirming the 35.46d period. Additionally, the sector 30 light curve identified 2 additional transits of a second planetary body on a 12.8d period. Reanalysis of the TESS sector 3 and NGTS data sets, along with the CORALIE, HARPS and FEROS RVs support the presence of a second planet of approximately 0.42 R_J and 0.12 M_J. Further analysis of the three-body system is currently underway.
Figure 9.4: Transits of NGTS-11b from TESS and NGTS. Both transits are binned to 30 mins but NGTS observed at a cadence of $\sim 13$ s. Taken from Figure 1 of Gill et al. [2020c].

9.9 Summary

This chapter has outlined some of the efforts of the NGTS monotransit working group; the subsection of the NGTS consortium concerned with the follow-up of monotransit candidates. This working group was conceived partially in response to favourable results from some of the simulations discussed in previous sections, specifically Chapters 4, 5 and 8. This chapter describes some of the approaches that this working group has used to identify these candidates in the TESS data and to follow them up. The initial search method was described in Section 9.2 with the current version of the candidate search described in Section 9.3. This chapter also highlights the two techniques used by NGTS to follow-up monotransit candidates, the blind survey and alias chasing modes, discussed in Sections 9.5.1 and 9.5.2 respectively.

NGTS photometry is not the only recourse for observing these targets and also discussed here are some of the other methods used including radial velocity data and archival photometry in sections 9.6 and 9.7. Finally, a list is given of some of the current successes of the working group in terms of published results with an emphasis on the first exoplanet from a TESS single transit, NGTS-11b. Multiple other results are currently in preparation.
Chapter 10

Conclusions

10.1 Thesis summary

This thesis has laid out a discussion of the simulation, detection and characterisation of monotransiting exoplanets discovered in the TESS photometric data. Monotransits are important to the advancement of our astronomical knowledge of exoplanets as well as to the potential for the discovery of non-terrestrial life in the future.

10.1.1 The importance of monotransits

Monotransits require hard work to properly utilise. They are hard to identify, many of the traditional methods are not applicable, and they are hard to characterise, potentially requiring large amounts of valuable observing time. Despite this however, they are extremely valuable to astronomy. Firstly, and perhaps most interestingly, monotransits are one of the few ways to build up a statistically significant population of long period planets. Observing long period planets otherwise requires long baselines, such as those provided by Kepler, but because of this, only very small areas of sky can be observed. To find long period systems across the whole sky one must identify and characterise planets from single transit events. Reaching these longer periods, and thus greater separations from their hosts, can then help to explore many scientifically interesting avenues. For example, reaching the habitable zones of sun-like stars requires probing planets with periods of hundreds of days. Exploring planets in this region will not only give potential insights into our own history and future but may even allow for the first detections of life around other stars. A current downside to this hope is that currently the monotransits that are most easily detectable are gas giants, although, it has been theorised that life could exist on moons orbiting these giant planets. Attempting to reach smaller radius planets using monotransits is challenging and requires a high degree of confidence in even very shallow signals. There are a number of
ways this can be achieved however. Firstly, targeting smaller radius hosts means that small radius planets produce relatively larger transit signals which are then higher signal-to-noise. Secondly, targeting hosts which are known to have low levels of activity means that planetary signals are less likely to become lost in the noise or written off as aspects of stellar noise. Longer monitoring of these stars will also allow for better characterisation of this activity and therefore make it easier to detect changes caused by transit features. Finally, more precise photometry will help to improve signal-to-noise of transit events, increasing the confidence in shallow signals.

Planets on long periods have other characteristics that are scientifically interesting. Planets further from their hosts are less influenced by these hosts. Because of this their evolution pathways may be significantly different. Planets at these periods are less likely to be tidally locked to their hosts, providing insights into atmospheric processes more similar to our own. Additionally, longer period systems are more likely to be eccentric, having not been circularised by stellar interactions (at least, long period exoplanets seem to have a wider range of eccentricities). Eccentricity can influence habitability but exploring the exact nature of this relationship can be aided by a larger population of long period, eccentric planets. The larger planet-host separation and reduced irradiation means that gas giants are less likely to become inflated. This will provide a larger sample of planets more similar to our own solar system, which is potentially key to the discovery of life. Probing these long period systems will also allow for the examination of planetary formation beyond the ice line. This type of formation is discussed in section 1.3.1 but only occurs at long periods.

It should also be noted the interest in longer period systems is not only limited to planets. The majority of stars are in binary systems but the majority of well characterised eclipsing binaries are on short periods. Short period binaries interact much more closely than longer period systems and may experience common envelope evolution where mass is transferred or shared between the binary components. Long period systems are less likely to undergo this process and thus may have significantly different evolutionary pathways. A statistically significant population of long period eclipsing binaries will allow us to better explore this region of parameter space. This will allow for a better grasp of the stellar population and evolution and may also lead to the discovery of longer period circumbinary planets.

10.1.2 TESS monotransit simulations and period aliases

A large part of this thesis concerns the number and detectability of the monotransits that will be present in the TESS data. Chapters 4 and 5 in particular lay out the simulations designed to determine how many detectable monotransits are expected from the TESS primary and extended missions with Chapter 8 updating the simulations and extending the
procedure to determining period aliases using follow-up techniques. The specific details across the simulations used vary slightly, depending on the paper. For example, through the thesis improvements and updates are made to the underlying assumptions as well as the TESS mission specifications and SNR thresholds but using the most up to date and reliable assumptions we expect in the range of 750 detectable monotransits from the southern ecliptic hemisphere primary TESS mission alone. One of the key aspects of this population is that monotransits outnumber multitransits for periods $\geq 30$ days. This is perhaps the key figure since monotransits are significantly harder to characterise than multitransits. Thus they are only worth expending effort on when the comparable population of multitransits is insufficient, such as the case for long period systems.

Extrapolating through to the extended mission we see that, depending on the specific set of assumptions made, 50-60% of these will transit once more in the extended mission producing a simulated population of around 400 double transiting exoplanets. This paper proves however, that two transits separated by 2 years are insufficient for period characterisation since this still leaves, on average, almost 40 potential period aliases (depending on the amount of TESS coverage). Simulations of both photometric and spectroscopic follow-up show that both methods are able to reduce the number of possible aliases and fully determine the monotransit period but to constrain the largest number of systems requires the two methods be used in parallel. A follow-up campaign of 50 days is sufficient to determine the period of approximately half of the doubly transiting systems with longer campaigns constraining more systems. When used separately photometry is capable of fully determining the period of more systems than spectroscopy but requires more telescope time to do so.

10.1.3 Monotransit identification and follow-up

The remainder of this thesis focuses on follow-up for the TESS monotransits, including identifying the candidates within the TESS data. Chapter 6 examines the use of a single photometric mission, CHEOPS, for the follow-up of TESS monotransits. This analysis shows that, assuming a baseline efficiency of 40%, CHEOPS is capable of observing up to 90% of the TESS primary mission monotransits at an average of 60 days per year. The majority of these would then have an additional transit caught by CHEOPS. Unlike NGTS however, CHEOPS cannot be employed in a stare campaign so would require a period estimate and predicted epoch before observations. Chapter 7 then looks at both photometric and spectroscopic follow-up in a more generalised way. Using a set of underlying assumptions this chapter shows that photometry and spectroscopy are each preferred as follow-up methods for different regions of exoplanet parameter space, namely planetary radius, stellar radius and period. Looking at two aspects of follow-up, SNR and time elapsed, it is seen
that the ratio of photometric to spectroscopic SNR goes like $R_p/P^{1/3}$ and that photometry is faster for short periods (<10 days) and smaller radius planets whereas spectroscopy is faster for longer period systems and larger radius planets.

Finally, Chapter 9, discusses the NGTS monotransit working group and efforts to identify, follow-up and characterise TESS monotransits. This chapter lays out two methods of identifying candidates, through the DIA light curves, and through a bespoke set of light curves as well as two methods of NGTS follow-up. These two methods are the blind survey mode, where the target is monitored until an additional transit is seen, and the alias chasing mode, where specific period aliases are targeted. Obviously the alias chasing mode is preferred as it is more efficient but requires a second transit in either the TESS extended mission or from some other source. The chapter also discusses some of the complimentary follow-up methods used by this group including archival photometric data and RV observations, primarily with CORALIE. The chapter ends with the details of the current successes of the group including a number of fully characterised long-period EBLMs and NGTS-11b, the first exoplanet discovered after identification as a TESS monotransit candidate.

### 10.2 Future work

The simulations and results presented in this thesis are by no means the end of the science to be done with TESS monotransits. A number of improvements and expansions to this work can be made and some are mentioned here.

#### 10.2.1 TESS simulation validation

A key piece of follow-up work is the validation of the monotransit simulations that have been presented in this thesis. The validation is already underway through the identification and confirmation of monotransit events in the TESS light curves. To best complete this validation the TESS data must all be searched in a consistent way and the realistic candidates must be confirmed. This involves the full eyeballing and follow-up methods discussed in Chapter 9 being carried out on, ideally, the full TESS data set. When this has been done the total number and distribution of monotransits can be compared with that predicted from these simulations and the simulations can then be validated or rejected. In the case that the found population does not match the simulated one changes will have to be made regarding the underlying assumptions. For example, perhaps the assumed underlying populations of long period systems are inaccurate, or perhaps the assumed detection methods were too lax or too stringent. These validations are vital since they may act to change our current knowledge of exoplanet populations and thus the way in which future surveys are designed and managed. For example, TESS is currently completing its first extended mission but
may survive past this threshold. Any future extensions will want to be optimised to produce the greatest scientific output which can only be done if simulations, based on correct assumptions, are carried out first.

Current work in the identification and confirmation of TESS monotransits is, as of yet, too immature to fully validate or reject the simulations presented in this thesis but current projections are favourable. In the TESS sectors searched so far the number of monotransit candidates is comparable to the number expected based on these simulations. Time is then required to complete these searches across all sectors and to then robustly confirm the nature of the candidates.

10.2.2 Statistical study of monotransits

With a large enough population of long-period monotransits it will be possible to carry out a statistical analysis of this planetary population. To be able to draw statistically significant conclusions about this population requires a large, unbiased sample of planets. This then requires completing the consistent monotransit search and following-up the candidates to get a confirmed population of long period systems. From this population one will be able to make inferences about the nature and distribution of planets at these longer periods. One of the key issues with the underlying occurrence rates used in the simulations in this thesis is that populations of long-period planets must be inferred or extrapolated from short period regimes where a more complete view of the populations is available. A larger, more complete sample of long-period systems will mean that these extrapolations are unnecessary, leading to more accurate long-period planet demographics.

Comparisons of this new population to a range of planetary formation physical models will then allow for the comparison of different theories leading to a more in-depth understanding of long-period planet formation and evolution. This is a key point as the pathways experienced by close-in planets may not be easily extended to the long-period regime. Having a statistically robust population to compare with simulations is vital for the determination of the validity of these theories. Specific physical processes that may be informed include the formation and evolution of planets beyond the snow line, the migration and circularisation of far-out planets and the determination of a more accurate value for the occurrence rate of Earth-like planets, $\eta_{\oplus}$, which requires knowledge of planets with periods of hundreds of days. It should be noted that it is often planetary density that is used or required for comparison with various physical models of planetary formation and structure. Therefore, it is necessary to have RV measurements of monotransiting planets, as well as photometric ones, in an effort to determine both radius and mass. This is discussed in Chapters 7 and 8.
10.2.3 More bespoke analysis of SpecPhot

Properly utilising follow-up time is very important when it comes to achieving the maximal science output for the available photometric and spectroscopic resources. To that end the simulation presented in Chapter 7 was carried out but this could be expanded upon. Firstly, as described above, repeating the analysis once a larger, more robust, population has been analysed would ensure that more accurate underlying assumptions are utilised leading to more realistic conclusions. An additional improvement would be to focus on more bespoke science cases. Because of the general nature of the problem approach in Chapter 7 the results, by necessity, must be fairly coarse. Were the simulation to focus on a narrower parameter range, for example, only planets with periods between 50 and 100 days, or only terrestrial planets, the results would allow for a more detailed exploration of the comparisons between the two methods. A greater range of options could also be explored when choosing the strategy employed by each method. The paper as it is presented uses only a single observing strategy but comparing multiple approaches may reveal a more efficient strategy. For example, instead of simple taking an RV point every \( n \) days it may actually be more efficient to observe in more logarithmic time increments. Or perhaps observing more or less frequently would change the required telescope time to accurately follow-up a system. Simulating the follow-up of a system using both methods concurrently may also reveal a more optimal strategy than using each method independently. Finally, the simulation presented puts no limits on available telescope time and uses only one photometric survey and two spectroscopic ones. In reality, particular groups will have access to different amounts of time on a whole range of instruments. Creating a bespoke simulation, or one that is able to be easily adapted by each interested party, that takes into account any chosen instruments and any limited amount of observing time may be more useful to individuals instead of the broad overview that was our goal for the simulation as it stands. A final enhancement that could be made is to develop the real time nature of the simulation. That is, to re-run the simulation as true observations are made with a view to adapting and optimising the ongoing observational strategy.

10.2.4 Improvements in NGTS monotransits working group

The pipeline for creating TESS FFI light curves, searching them for monotransit candidates and confirming these candidates within the NGTS monotransit working group is evolving and being updated constantly. Thus, there are improvements that can be made and some particular ones are discussed here. A key aspect of the procedure is simply streamlining and speeding up the entire work flow. Reducing the time between the availability of the FFIs and the generation of our candidate lists will mean that we can start to observe targets
closer in time to the transits visible in TESS. This is important from the point of view of
the science, wanting to produce scientifically interesting results quickly, but also from a
telescope efficiency point of view. Especially in cases where two transits are seen, every
orbit after an observed transit reduces the precision of our predictions of additional events.
Acting quickly means less telescope time is require to combat inaccuracies caused by ex-
trapolating future transit times. An additional improvement is to use other FFI light curves
as comparisons. As the TESS mission has continued other teams have produced some sets
of FFI light curves similar to those produced and used in this chapter. Comparing these
will both help to catch strange events, such as errors in detrending but also help to improve
confidence in candidates if they are seen in multiple, independently created light curves.
This is already being carried out on an object-by-object basis but could be optimised to be
run automatically for all targets. Finally, the way the monotransit search is currently run
(as well as the concurrent BLS search) is to run on each sector of data individually. Sec-
tors are only looked at together during the eyeballing stages. Ideally, when a new sector of
data comes it for a target that already has coverage it should have the monofind and BLS
algorithms run on it individually and in combination with all previous sectors of data. This
would make it easier and quicker to identify targets with multiple events across different
sectors and even between the primary and extended missions.

10.3 Final remarks

The study of exoplanets is one of the ways in which we endeavour to learn more about the
universe in which we find ourselves. The discovery of planets around other stars has had
a profound effect on science and the way in which we see ourselves as part of something
more. This exploration is really only beginning and a vital next step is exploring Earth ana-
logues and monotransits are one way of doing that. To reach out to the long period regimes
required, in a statistically robust and informative way, we need to explore long-period plan-
etes across the sky and this can be achieved by devoting time to monotransiting planets
observed by TESS. These systems are hard to detect, and require dedicated time and effort
to confirm, follow-up and characterise but the results presented in this thesis have shown
that we are capable of the task. Though the first steps in this field have been taken, and
key discoveries have been made, there is certainly the potential for further improvements.
These could include the continued improvement of photometric precision, the development
of more bespoke and sophisticated simulations, or a change in strategy, such as targeting
particularly favourable potential hosts, or even the construction of dedicated observing fa-
cilities. Further work in this exciting area will lead to a more nuanced understanding of the
universe and maybe even to the detection of extra-solar life.
Chapter 11

Full author publication list

11.1 First author


2. Cooke, Pollacco, and Bayliss [2019], An examination of the effect of the TESS extended mission on southern hemisphere monotransits, *Astronomy & Astrophysics*


11.2 Co-author


5. West, Gillen, Bayliss, Burleigh, Delrez, Günther, Hodgkin, Jackman, Jenkins, King, McCormac, Nielsen, Raynard, Smith, Soto, Turner, Wheatley, Almleaky, Armstrong, Belardi, Bouchy, Briegal, Burdanov, Cabrera, Casewell, Chaushev, Chazelas, Chote,


197


19. Bayliss, Wheatley, West, Pollacco, Anderson, Armstrong, Bryant, Cegla, Cooke, Gänsicke, Gill, Jackman, Loudon, McCormac, Acton, Burleigh, Casewell, Goad,


Bibliography


Néstor Espinoza, Rafael Brahm, Thomas Henning, Andrés Jordan, Caroline Dorn, Felipe Rojas, Paula Sarkis, Diana Kossakowski, Martin
Courtney D. Dressing and David Charbonneau. The Occurrence of Potentially Habitable Planets Orbiting M Dwarfs Estimated from the


M. Cecconi, H. M. Cegla, R. Cloutier, A. Collier Cameron, R. Cosentino, A. Ghedina, M. Gonzalez, R. D. Haywood, D. W. Latham,
M. Lodi, M. Lopez-Morales, J. Maldonado, L. Malavolta, G. Micela, E. Molinari, A. Mortier, H. Perez Ventura, M. Painamonti, E. Poretti,
2020.

Jason Eastman. EXOFASTv2: Generalized publication-quality exoplanet modeling code, October 2017.


Philipp Eigmüller, Alexander Chaushman, Edward Gillen, Alexis Smith, Louise D. Nielsen, Oliver Turner, Szilard Csizmadia, Barry Smalley,
Daniel Bayliss, Claudia Belardi, François Bouchy, Matthew R. Burleigh, Juan Cabrera, Sarah L. Casewell, Bruno Chazelas, Benjamin
F. Cooke, Anders Eriksson, Boris T. Gansicke, Maximilian N. Günther, Michael R. Goad, Andrew Grange, James A. G. Jackman,
James S. Jenkins, James McCrorc, Maximiliano Moyano, Don Pollacco, Katja Poppenhaeger, Didier Queloz, Lian Raynard, Heike
Rauer, Stéphane Udry, Simon. R. Walker, Christopher A. Watson, Richard G. West, and Peter J. Wheatley. NGTS-5b: a highly inflated

Néstor Espinoza, Rafael Brahm, Thomas Henning, Andrs Jordán, Caroline Dorn, Felip Rojas, Paula Sarkis, Dian Kossakowski, Martin
Schlecker, Matías R. Díez, James S. Jenkins, Claudia Aguileraga, Gomez, Jon M. Jenkins, Joseph D. Twicken, Karen A. Collins, Jack
Lissauer, David J. Armstrong, Vardan Anidbyekyan, David Barrado, Susana C. C. Barros, Matthew Batmanghelich, Daniel Bayliss, François Bouchy,
Edward M. Bryant, Benjamin F. Cooke, Olivier D. S. Demangeon, Xavier Dumusque, Pedro Figueira, Helen Giles, Jorge Lillo-Box,
Christophe Lovis, Louise D. Nielsen, Francesco Pepe, Don Pollacco, Nuno C. Santos, Sergio G. Sousa, Stéphane Udry, Peter J. Wheatley,
Oliver Turner, Maxime Marmier, Damien Ségransan, George Ricker, David Latham, Sara Seager, Joshua N. Winn, John F. Kielkopf,
Rhodes Hart, Geoff Wingham, Eric L. N. Jensen, Krzysztof G. Hełminiak, A. Tokovinin, C. Briceño, Carl Ziegler, Nicholas M. Law,
Andrew W. Mann, Tatsu Daylan, John P. Doty, Natalia Guerrero, Patricia Boyd, Ian Crossfeld, Robert L. Morris, Christopher E. Henn.
and Aaron Dean Chacon. HD 213885b: a transiting 1-d-period super-Earth with an Earth-like composition around a bright (V
Brown, E. Bryant, J. Cabrera, W. D. Cochran, S. Csizmadia, H. Deeg, O. Demangeon, M. Deleuil, X. Dumusque, P. Eigmüller, M. Endl,
S. Udry, and J. N. Winn. HD 213885b: a transiting 1-d-period super-Earth with an Earth-like composition around a bright (V = 7.9) star

M. Esposito, D. J. Armstrong, D. Gandolfi, V. Adibekyan, M. Fridlund, N. C. Santos, J. H. Livingstone, E. Delgado Mena, L. Fossati, J. Lillo-
S. Udry, and J. N. Winn. HD 213885b: a transiting 1-d-period super-Earth with an Earth-like composition around a bright (V = 7.9) star

Thomas M. Evans, David K. Sing, Tiffany Kataaria, Jayesh Goyal, Nikolay Nikolov, Hannah R. Wakefield, Drake Deming, Mark S. Marley,
David S. Amundsen, Gilda E. Ballester, Joanna K. Barstow, Lotfi Ben-Jafal, Vincent Bourrier, Lars A. Buchhave, Ofir Cohen, David
Ehrenreich, Antonio Garcia Muñoz, Gregory W. Henry, Heather Knutson, Panayotis Lavvas, Alain Lecavelier Des Etangs, Nikole K. Lewis,
Mercedes López-Morales, Avi M. Mandell, Jorg Sanz-Forcada, Pascal Tremblin, and Roxana Lupu. An ultrahot gas-giant exoplanet with


211


Thomas Little Heath and Aristarchus of Samos. *Aristarchus of Samos, the ancient Copernicus; a history of Greek astronomy to Aristarchus, together With Aristarchus’s Treatise on the sizes and distances of the sun and moon*. Clarendon press, 1913.


218


Senior authors correspond to corresponding author. This work was supported in part by the National Aeronautics and Space Administration (NASA), the European Research Council (ERC) through the H2020 program, the French space agency (CNES), and the French National Research Agency (ANR).


