Library Declaration and Deposit Agreement

1. **STUDENT DETAILS**

   James Kirk

   u1456865

2. **THESIS DEPOSIT**

   2.1 I understand that under my registration at the University, I am required to deposit my thesis with the University in BOTH hard copy and in digital format. The digital version should normally be saved as a single pdf file.

   2.2 The hard copy will be housed in the University Library. The digital version will be deposited in the Universitys Institutional Repository (WRAP). Unless otherwise indicated (see 2.3 below) this will be made openly accessible on the Internet and will be supplied to the British Library to be made available online via its Electronic Theses Online Service (EThOS) service. [At present, theses submitted for a Masters degree by Research (MA, MSc, LLM, MS or MMedSci) are not being deposited in WRAP and not being made available via EthOS. This may change in future.]

   2.3 In exceptional circumstances, the Chair of the Board of Graduate Studies may grant permission for an embargo to be placed on public access to the hard copy thesis for a limited period. It is also possible to apply separately for an embargo on the digital version. (Further information is available in the Guide to Examinations for Higher Degrees by Research.)

   2.4 (a) Hard Copy I hereby deposit a hard copy of my thesis in the University Library to be made publicly available to readers immediately.
   I agree that my thesis may be photocopied.

   (b) Digital Copy I hereby deposit a digital copy of my thesis to be held in WRAP and made available via EThOS.
   My thesis can be made publicly available online.

3. **GRANTING OF NON-EXCLUSIVE RIGHTS**

   Whether I deposit my Work personally or through an assistant or other agent, I agree to the following: Rights granted to the University of Warwick and the British Library and the user of the thesis through this agreement are non-exclusive. I retain all rights in the thesis in its present version or future versions. I agree that the institutional repository administrators and the British Library or their agents may, without changing content, digitise and migrate the thesis to any medium or format for the purpose of future preservation and accessibility.

4. **DECLARATIONS**

(a) I DECLARE THAT:

- I am the author and owner of the copyright in the thesis and/or I have the authority of the authors and owners of the copyright in the thesis to make this agreement. Reproduction of any part of this thesis for teaching or in academic or other forms of publication is subject to the normal limitations on the use of copyrighted materials and to the proper and full acknowledgement of its source.
- The digital version of the thesis I am supplying is the same version as the final, hardbound copy submitted in completion of my degree, once any minor corrections have been completed.
- I have exercised reasonable care to ensure that the thesis is original, and does not to the best of my knowledge break any UK law or other Intellectual Property Right, or contain any confidential material.
- I understand that, through the medium of the Internet, files will be available to automated agents, and may be searched and copied by, for example, text mining and plagiarism detection software.

(b) IF I HAVE AGREED (in Section 2 above) TO MAKE MY THESIS PUBLICLY AVAILABLE DIGITALLY, I ALSO DECLARE THAT:

- I grant the University of Warwick and the British Library a licence to make available on the Internet the thesis in digitised format through the Institutional Repository and through the British Library via the EThOS service.
- If my thesis does include any substantial subsidiary material owned by third-party copyright holders, I have sought and obtained permission to include it in any version of my thesis available in digital format and that this permission encompasses the rights that I have granted to the University of Warwick and to the British Library.

5. **LEGAL INFRINGEMENTS**

I understand that neither the University of Warwick nor the British Library have any obligation to take legal action on behalf of myself, or other rights holders, in the event of infringement of intellectual property rights, breach of contract or of any other right, in the thesis.

Please sign this agreement and return it to the Graduate School Office when you submit your thesis.

Student's signature: ....Date: ....
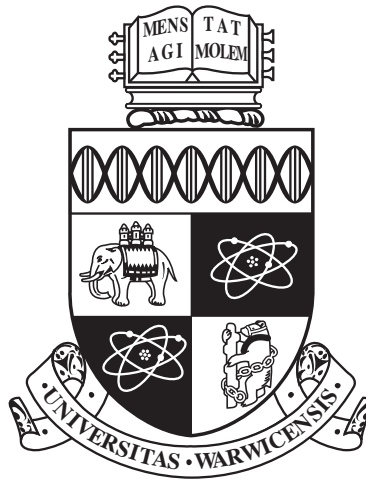
Optical Transmission Spectroscopy of Hot Jupiter Atmospheres

by

James Kirk

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

April 2018

THE UNIVERSITY OF

WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

Firstly, I would like to thank my supervisor Prof. Peter Wheatley for his excellent guidance and advice over the last four years, in addition to his friendship. I am very grateful to him for offering me the opportunity to work in such an exciting field.

I am also immensely grateful to my family and Alex for their endless support, love and encouragement, which I can always count on. My regular trips to Poole and Boston, with diversions to London and Australia, have been a lot of fun. Additionally, I would like to thank my parents for christening me 'James', little did they know how excellent the name 'James Kirk' would be as an ice breaker at conferences and that it would inspire the design of all my presentation slides and posters.

I would also like to thank everyone in the astronomy group at Warwick for making it such an enjoyable place to work and for always being on hand whenever I had a question about data reduction, coding and interpretation. My thanks also go to Ian Skillen at the Isaac Newton Group of Telescopes for all his help in gathering and interpreting the ACAM data which have formed a large part of this thesis. I also thank all my friends at Warwick who have helped to make these years immensely enjoyable and unforgettable.

Finally, it seems appropriate to begin my thesis with the following (slightly altered) quote:

Space: the final frontier. These are the voyages of the PhD student James Kirk. His four-year mission: to explore strange new worlds, to seek out opacity sources in hot Jupiter atmospheres, to boldly go where no one has gone before.

Declarations

I declare that the work in this thesis is my own except where stated otherwise, and was carried out at the University of Warwick, during the period October 2014 to April 2018, under the supervision of Prof. Peter Wheatley. The research reported here has not been submitted, either wholly or in part, in this or any other academic institution for admission to a higher degree.

The work in chapter 4 is based on the paper 'Transmission spectroscopy of the inflated exoplanet WASP-52b, and evidence for a bright region on the stellar surface' published in *Monthly Notices of the Royal Astronomical Society* (MNRAS), 463, 2922 (2016).

The work in chapter 5 is based on the paper 'Rayleigh scattering in the transmission spectrum of HAT-P-18b' published in MNRAS, 468, 3907 (2017).

The work in chapter 6 is based on the paper 'LRG-BEASTS III: Ground-based transmission spectrum of the gas giant orbiting the cool dwarf WASP-80' published in MNRAS, 474, 876 (2018).

Other contributions based on this thesis are:

- 'Characterising exoplanet atmospheres as part of the LRG-BEASTS survey'. Oral presentation at ESO Vitacura, Santiago, Chile, March 2018.

- 'Characterising exoplanet atmospheres as part of the LRG-BEASTS survey'. Poster presentation at UK Exoplanet Community Meeting, University of Oxford, March 2018.

- 'Characterising exoplanet atmospheres as part of the LRG-BEASTS survey'.

Oral presentation at the MIT Kavli Institute for Astrophysics and Space Research, Cambridge, MA, January 2018.

- 'Characterising exoplanet atmospheres as part of the LRG-BEASTS survey'. Oral presentation at the 231st Meeting of the AAS, Washington DC, January 2018.

- 'LRG-BEASTS: The Low Resolution Ground-Based Exoplanet Atmosphere Survey using Transmission Spectroscopy'. Poster presentation at the Transiting Exoplanets Conference, Keele University, July 2017.

- 'Rayleigh scattering in the atmosphere of HAT-P-18b'. Poster presentation at UK Exoplanet Community Meeting, University of St. Andrews, March 2017.

- 'Transmission Spectroscopy using ULTRACAM'. Oral presentation at UK Exoplanet Community Meeting, University of Exeter, March 2016.

- 'Transmission Spectroscopy of WASP-52b'. Oral presentation at National Astronomy Meeting, Llandudno, July 2015.

- 'Transmission Spectroscopy of WASP-52b'. Poster presentation at UK Exoplanet Community Meeting, University of Warwick, April 2015.

Abstract

The field of exoplanet atmospheres has advanced rapidly in the 16 years since the detection of sodium in the atmosphere of HD 209458b. Giant planets on short period orbits ('hot Jupiters') have been key to this progression, due to their relatively characterisable atmospheres. Optical transmission spectroscopy revealed the first detection of an exoplanet atmosphere and has continued to play a vital role in atmospheric characterisation. In this thesis I present optical transmission spectroscopy of three hot Jupiters and place these in context within the LRG-BEASTS transmission spectroscopy survey.

My ground-based transmission spectrum of WASP-52b revealed the presence of clouds in the planet's atmosphere and evidence for stellar faculae on the host. This study demonstrated that HST precision could be achieved from the ground.

I also present ground-based transmission spectroscopy of HAT-P-18b. My precise transmission spectrum was well fitted with a Rayleigh scattering slope at the equilibrium temperature of the planet. The absence of broad sodium and potassium absorption led me to conclude that a high altitude haze was present.

Finally I present a transmission spectrum of WASP-80b. I found that a haze was likely present in this planet's atmosphere and my transmission spectrum was inconsistent with a previously claimed detection of pressure-broadened potassium.

My results are in agreement with the emerging correlation that cooler planets are less likely to have clear atmospheres. However, this correlation is still tentative. The LRG-BEASTS survey will test this relation and help improve our understanding of the underlying physics driving the formation of clouds and hazes in exoplanet atmospheres.

Abbreviations

ADU Analogue-to-Digital Unit

BIC Bayesian Information Criterion

CCD Charge-Coupled Device

FWHM Full Width at Half Maximum

GP Gaussian Process

GTC Gran Telescopio Canarias

HST Hubble Space Telescope

LRG-BEASTS Low Resolution Ground Based Exoplanet Atmosphere Survey using Transmission Spectroscopy

MCMC Markov chain Monte Carlo

ppm Parts per million

PSF Point Spread Function

RV Radial Velocity

VLT Very Large Telescope

WFC3 Wide Field Camera 3

WHT William Herschel Telescope

AU Astronomical Unit, 1.496×10^8 km

$M_\oplus$ Earth Mass, 5.972×10^{24} kg

$R_\oplus$ Earth Radius, 6378 km

M_J Jupiter Mass, 1.898×10^{27} kg

R_J Jupiter Radius, 71492 km

$M_\odot$ Solar Mass, 1.988×10^{30} kg

$R_\odot$ Solar Radius, 695700 km

Chapter 1

Introduction

In the 26 years since the first detection of an extra-solar planet (exoplanet), the field of exoplanets has taken off, with more than 3,700 confirmed or validated detections to date (NASA Exoplanet Archive[1], 16/02/2018). This boom has led to studies of their formation, evolution, environment, atmospheres and destruction. As the field is so rich, it is beyond the scope of this introduction to go into all stages of an exoplanet's life. Instead, I give a broad overview, beginning with a brief history of exoplanet detection in section 1.1 before moving on to the methods of detecting exoplanets in section 1.2. Following this, I give an overview of the population of known exoplanets in section 1.3 and outline the formation and evolution of exoplanets in section 1.4. Since this thesis is primarily concerned with the study of exoplanet atmospheres, I go into more detail regarding this field in section 1.5.

1.1   A brief history of exoplanet detection

The first detection of an exoplanet came in 1992 (Wolszczan & Frail, 1992) when two planets were found to be orbiting the pulsar PSR B1257+12. This detection was made via the observation of differences in the time of arrival of the pulses of the millisecond pulsar, which could be fit by Keplerian orbits of planetary mass objects. A follow up study later found the presence of a third orbiting body which remains to this day the lowest-mass exoplanet discovered (NASA Exoplanet Archive, 14/12/2017) with a mass twice that of the moon (Wolszczan, 1994).

Before this discovery, there were tentative detections of exoplanets using radial velocity measurements (see section 1.2.2), such as the planet γ Cepheid b in 1988 (Campbell et al., 1988), however this was not confirmed until 2003 (Hatzes

---

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

et al., 2003) and only after the initial signal had been re-evaluated as likely being due to stellar activity (Walker et al., 1992). In 1989, radial velocity measurements revealed the presence of a body orbiting the solar-type star HD 114762 however it was thought to be an orbiting brown dwarf due to its minimum mass of $11\,M_J$ (Latham et al., 1989).

The first bona fide detection of an exoplanet orbiting a Sun-like star was made in 1995, that of 51 Pegasi b (Mayor & Queloz, 1995). This discovery combined the star's measured $v\sin i$ with the rotational velocity (v) inferred from the star's chromospheric activity to conclude that the planet's maximum mass was twice that of Jupiter. For this reason, the authors were able to conclusively rule out the possibility that the orbiting body was a brown dwarf which previous studies were unable to do. 51 Pegasi b also made the community aware of the existence of *hot Jupiters* which had not been predicted by models of giant planet formation (Mayor & Queloz, 1995).

Following this discovery in 1995, a steady flow of exoplanets detected with the radial velocity method ensued (Fig. 1.1). The next major milestone in exoplanet detection came in 1999 with the first detection of the transit of an exoplanet (Charbonneau et al., 2000; Henry et al., 2000). This detection was of the planet HD 209458 b, whose host star had previously been seen to show radial velocity variations indicating the presence of a planet (Henry et al., 2000).

Since the detection of the transit of HD 209458 b, dedicated transit surveys from the ground, such as WASP (Pollacco et al., 2006), HAT (Bakos et al., 2004), KELT (Pepper et al., 2007), NGTS (Wheatley et al., 2017a); and space through CoRoT (Baglin et al., 2006), Kepler (Borucki et al., 2010) and now K2 (Howell et al., 2014), have been very successful, with the transit method now accounting for the majority of detected exoplanets (Fig. 1.1). However, a caveat of these detections is that many of the Kepler planets were 'validated' not 'confirmed'. Owing to the relatively faint stars Kepler observed, these planets could not be confirmed in the conventional way (by obtaining masses through radial velocities) and instead were validated through robust statistical tests that ruled out false positives (e.g. Morton 2012, 2015).

1.2 Methods of detecting exoplanets

1.2.1 Transits

As mentioned in the previous section, the transit method of detection accounts for the vast majority of discovered exoplanets, with nearly 80% of all known exoplanets
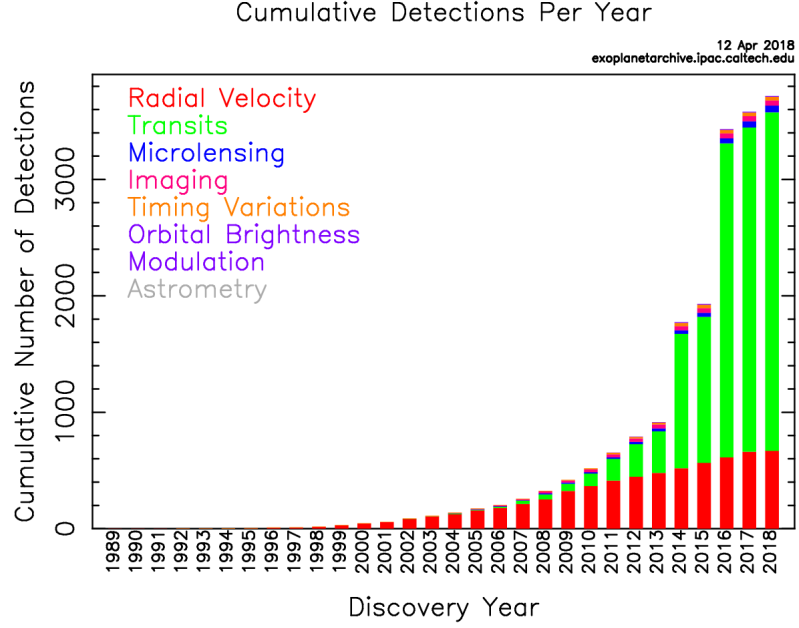
Figure 1.1: Plot of the cumulative detections of exoplanets per year and colour-coded by detection method. Figure from the NASA Exoplanet Archive, accessed 20/04/18.

discovered with this technique (NASA Exoplanet Archive, 20/12/2017).

The transit probability

Transiting exoplanets are often on short period orbits as they have the highest probability of detection. The reason for this can be understood when considering the configuration depicted in Fig. 1.2. Following the working of Haswell (2010), a transit can only occur when the disc of the planet and the disc of the star overlap as seen from the observer's viewpoint. Referring to Fig. 1.2, the sky projected distance between the centres of the planet and star at conjunction is defined by the impact parameter b. The impact parameter is given by

$$b = \frac{a}{R_*} \cos i \tag{1.1}$$

where a is the semi-major axis of the exoplanet's orbit, R_* is the radius of the star and i is the inclination of the orbital plane. For a transit to be observed, the orbital inclination must satisfy the following condition

$$a \cos i \leq R_P + R_* \tag{1.2}$$

3
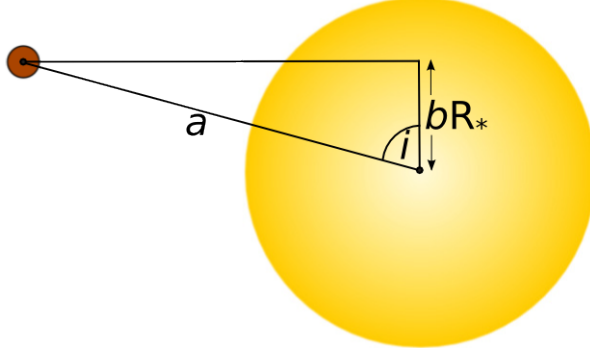
Figure 1.2: Figure showing the trigonometry of the impact parameter b given the orbital inclination i and semi-major axis a for a transiting exoplanet. Figure adapted from https://www.paulanthonywilson.com/, accessed 20/12/17.

where R_P is the radius of the planet. To find the geometric transit probability, we must integrate this over all possible angles of the stellar inclination, which gives

$$\text{geometric transit probability} = \frac{R_P + R_*}{a} \approx \frac{R_*}{a} \tag{1.3}$$

assuming $R_* >> R_P$. Equation 1.3 therefore reveals that transits are most likely to occur in systems with large host stars and small separations. Assuming a star with $1\,\mathrm{R_\odot}$, the probabilities for transits given semi-major axes of 0.1, 1 and $10\,\mathrm{AU}$ are $\sim 5\%$, 0.5% and 0.05%. This means that most transiting planets are on short period orbits where they are easier to find. Fig. 1.3 shows the distribution of semi-major axes of all exoplanets discovered with the transit technique, which demonstrates that nearly all have semi-major axes less than $1\,\mathrm{AU}$, with the peak in the distribution at around $0.05\,\mathrm{AU}$.

Transit depth

Assuming the stellar disc is of uniform brightness, the transit depth, δ, is equal to the fraction of the stellar disc eclipsed by the planet, which is given by

$$\delta = \frac{\Delta F}{F} = \frac{L_*^{OT} - L_*^{IT}}{L_*^{OT}} = \frac{4\pi R_*^2 \sigma_{SB} T_*^4 - 4\pi(R_*^2 - R_P^2)\sigma_{SB} T_*^4}{4\pi R_*^2 \sigma_{SB} T_*^4} = \frac{R_P^2}{R_*^2} \tag{1.4}$$

where F is the flux of the star, L_*^{OT} and L_*^{IT} is the luminosity out-of-transit and in-transit, R_* is the radius of the star, R_P is the radius of the planet, T_* is the temperature of the star and σ_{SB} is the Stefan-Boltzmann constant. In reality,
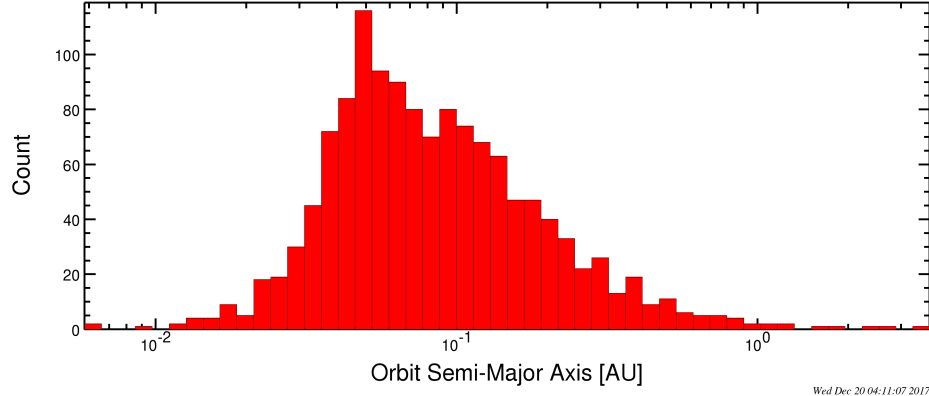
Figure 1.3: Histogram of exoplanets discovered with the transit technique and binned by their semi-major axes. Plot from NASA Exoplanet Archive, accessed 20/12/2017.
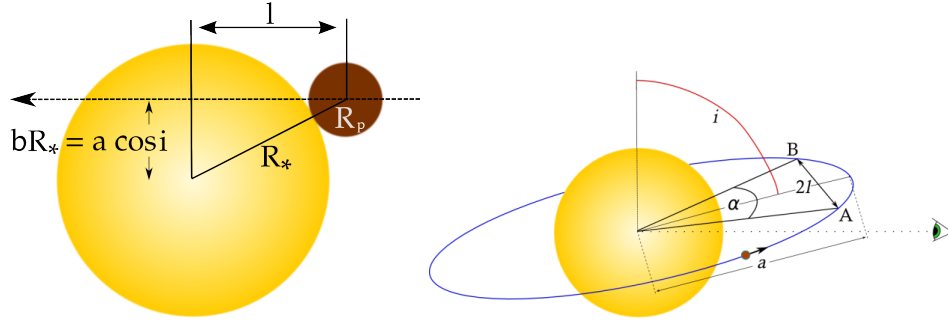
Figure 1.4: Plots useful in the calculation of the transit duration. Figures from https://www.paulanthonywilson.com/, accessed 20/12/17.

the above is not strictly true as the star is not of uniform brightness and instead displays limb darkening (which is discussed later in this section).

Equation 1.4 shows the transit depth is simply given by the radius ratio of planet to star squared, when the planet is in a non-grazing orbit (i.e. the full disc of the planet transits across the star). Transits can therefore reveal the radius of an exoplanet.

Transit duration

Following Haswell (2010) and using the depictions in Fig. 1.4 we can calculate the duration of an exoplanet's transit.

From Fig. 1.4 (left-hand panel), the projected length l is given by Pythago-

ras's theorem as

$$l = \sqrt{(R_* + R_P)^2 - a^2 \cos^2 i} \tag{1.5}$$

and from Fig. 1.4 (right-hand panel), we can see that

$$\sin\left(\frac{\alpha}{2}\right) = \frac{l}{a}. \tag{1.6}$$

The duration of the transit is given by the time the planet takes to travel along the length of the arc between the first and fourth contact points of the transit (given by A and B). For a circular orbit this is given by

$$T_{dur} = P\frac{\alpha}{2\pi} = P\frac{\arcsin(l/a)}{\pi} = \frac{P}{\pi}\arcsin\left(\frac{\sqrt{(R_* + R_P)^2 - a^2 \cos^2 i}}{a}\right) \tag{1.7}$$

where P is the orbital period of the planet and α is in radians. Assuming $R_* \gg R_P$ and $a \gg R_*$, equation 1.7 simplifies to

$$T_{dur} \approx \frac{P}{\pi}\left[\left(\frac{R_*}{a}\right)^2 - \cos^2 i\right]^{1/2}. \tag{1.8}$$

This can be combined with Kepler's third law, which is given by

$$\frac{a^3}{P^2} = \frac{G(M_* + M_P)}{4\pi^2} \tag{1.9}$$

to derive the orbital inclination if we know the stellar mass and radius. This then gives us the latitude on the star of the planet's transit chord, which is needed for accurate modelling of the limb darkening (as discussed later in this section).

Other parameters from the transit light curve

Seager & Mallén-Ornelas (2003) showed that the density of the host star, ρ_*, can be derived directly from the transit light curve in the case of a circular planetary orbit through

$$\frac{\rho_*}{\rho_\odot} = \frac{32}{G\pi}P\frac{\Delta F^{3/4}}{(t_T^2 - t_F^2)^{3/2}} \tag{1.10}$$

where $\rho_\odot$ is the Solar density, G is the gravitational constant, P is the orbital period in days, t_T is the transit duration (1st to 4th contact points) in days, t_F is

the duration between the 2nd and 3rd contact points in days and ΔF is the transit depth.

Additionally, Southworth et al. (2007) showed that the planetary surface gravity, g_P, could be derived from the transit light curve and semi-amplitude of the radial velocity (K_*, discussed in section 1.2.2) through

$$g_P = \frac{2\pi}{P} \frac{(1 - e^2)^{1/2} K_*}{r_P^2 \sin i} \qquad (1.11)$$

where e is the eccentricity of the planet's orbit, i is the orbital inclination and r_P is the planet's radius divided by the semi-major axis (R_P/a).

Transit timing and duration variations

Transit timing variations (TTVs) and transit duration variations (TDVs) are two methods that can be used to find and characterise planets.

TTVs are observed as a deviation from a linear ephemeris, i.e. the transit occurs before or after when it is predicted to from a linear trend. These variations can be caused by gravitational interactions with an additional body in the system. The amplitude of the TTVs is proportional to the mass ratio of the planets and so the masses of the planets can be determined (e.g. Holman et al. 2010; Agol & Fabrycky 2017), as has been done, for instance, for the seven planets in the TRAPPIST-1 system (Gillon et al., 2017; Delrez et al., 2018). TTVs can also be used to detect additional planets in a system (e.g. Agol et al. 2005) such as the detection of Kepler-19c (Ballard et al., 2011).

TDVs can be caused by changes in the inclination of a planet's orbit, due to a perturbing body, which changes the length of the transit chord and hence the duration of the transit (Agol & Fabrycky, 2017). An example of which has been observed in the Kepler-108 system, where TDVs revealed the two Saturn-mass planets have a high mutual inclination hinting at a chaotic past (Mills & Fabrycky, 2017).

Limb darkening

When modelling exoplanet transits we need to take into account the limb darkening of the star which causes the disc of the star to have a non-uniform brightness distribution. It is important to use an accurate description of the stellar limb darkening in order to recover accurate parameters from an exoplanet's transit.

Limb darkening arises because the observer's line of sight passes through the stellar atmosphere at different angles. At the limb, the line of sight reaches an

optical depth of unity at a shallower depth than it does at the centre of the disc. We therefore see less deeply into the star at the limb and we see a slightly lower temperature than at the centre of the stellar disc. This causes the limb to be darker than the centre of the disc. The effect is greater in the blue than the red as it is approximately the ratio of two black bodies of slightly different temperatures. In addition to this wavelength dependence, the limb darkening is also dependent on the properties of the star.

Limb darkening is a complex property of the opacity sources of the stellar atmosphere but can be adequately parameterised by relatively simple laws. To describe the limb darkening profile of a star, one often uses a linear (one parameter), quadratic (two parameter) or non-linear (four parameter) law. These laws are given in Claret (2000) as

Linear

$$\frac{I(\mu)}{I(1)} = 1 - u(1 - \mu) \tag{1.12}$$

Quadratic

$$\frac{I(\mu)}{I(1)} = 1 - u_1(1 - \mu) - u_2(1 - \mu)^2 \tag{1.13}$$

Non-linear

$$\frac{I(\mu)}{I(1)} = 1 - u_1(1 - \mu^{1/2}) - u_2(1 - \mu) - u_3(1 - \mu^{3/2}) - u_4(1 - \mu^2) \tag{1.14}$$

where $\mu = \cos\theta$, with θ being the angle between the line of sight and the stellar surface normal, and u_n are the *limb darkening coefficients.* Limb darkening profiles generated using these three laws are shown in Fig. 1.5. Often one will use tabulated values for the limb darkening coefficients which have been calculated from a grid of stellar atmosphere models (e.g. those of Claret 2000). Recently, Parviainen & Aigrain (2015) introduced the *Limb Darkening Tool Kit* (LDTK), a PYTHON module that uses PHOENIX stellar atmosphere models (Husser et al., 2013) to compute limb darkening coefficients for user-defined bandpasses. This module also propagates the errors in the stellar parameters through to the limb darkening coefficients and has thus been very useful in the analyses presented in this thesis. The results presented in Chapters 5 and 6 both made use of LDTK.

The linear limb darkening law is now infrequently used in the literature due to its inability to accurately model the true limb darkening profile of the star (e.g. Southworth 2008; Espinoza & Jordán 2016). Higher order limb darkening laws, however, suffer from degeneracies between the limb darkening coefficients so one limb darkening coefficient is often fixed at a theoretical value while the other
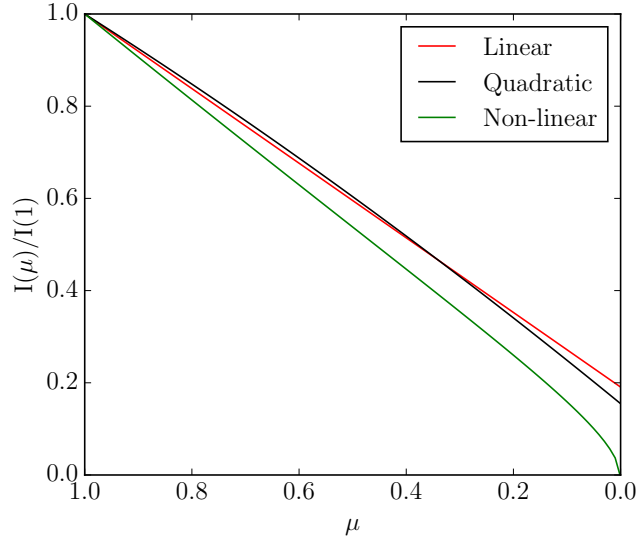
Figure 1.5: A plot of limb darkening profiles using the linear (red), quadratic (black) and non-linear (green) laws. The intensity is shown as a function of μ. The limb darkening coefficients were calculated using LDTK for an example star with an effective temperature of 5500 K, log g of 4.5 and metallicity, [Fe/H], of 0.25 for a passband between 5000 and 5100 Å.

is fitted for (e.g. Southworth 2008; Mallonn et al. 2015a). However, Espinoza & Jordán (2016) argue that fixing limb darkening coefficients is bad practice as it can lead to inaccurate transit parameters, in part due to our limited knowledge of the true intensity profiles of stars. To try and overcome this, there have been efforts made to recover more accurate limb darkening profiles using 3D model atmospheres rather than the standard 1D atmospheres (e.g. Hayek et al. 2012; Magic & Asplund 2014).

Instead of fixing the limb darkening coefficients, Espinoza & Jordán (2015) argue that they should be free parameters when fitting transit light curves. Similar to Espinoza & Jordán (2015), I have found that model-predicted values do not always accurately reflect the data, and so I treat the limb darkening on a case by case basis in Chapters 4, 5 and 6.

Kipping (2013) introduced a new parametrisation of the two coefficient laws which allows for more efficient sampling of the joint posterior distribution function of u_1 and u_2 while reducing the mutual correlation between the two limb darkening coefficients, shortening the time of convergence. In the case of the quadratic law (equation 1.13), Kipping (2013) advocated the use of $q_1 = (u_1 + u_2)^2$ and $q_2 = u_1/(2u_1 + 2u_2)$.

Espinoza & Jordán (2016) compared the biases in the derived transit parameters using different limb darkening laws with coefficients derived using ATLAS9 stellar models (Kurucz, 1993). In their study, the authors found that the biases introduced by different laws only deviate significantly from one another when considering light curves with noise of $\lesssim 100\,$ppm.

Due to the noise level of the light curves in this thesis, the non-linear limb darkening is not necessary and would add additional free parameters. Instead I opt to use the quadratic law and consider the effects of fixing and fitting the limb darkening coefficients in Chapters 4, 5 and 6.

1.2.2 Radial velocity

Since all bodies in a planetary system orbit a common centre of mass, the barycentre, the change in velocity of a star due to its reflex orbit around the barycentre can be measured through observing the Doppler shift in the emitted light from the star. This can be used to detect the presence of an orbiting exoplanet whose mass can be determined from the radial velocity semi-amplitude, K_*, using the equation (e.g. Perryman 2011)

$$K_* = \left(\frac{2\pi G}{P}\right)^{\frac{1}{3}} \frac{M_P \sin i}{(M_* + M_P)^{\frac{2}{3}}} \frac{1}{(1 - e^2)^{\frac{1}{2}}} \qquad (1.15)$$

where G is the gravitational constant, P is the orbital period, M_P is the mass of the planet, i is the orbital inclination, M_* is the mass of the star and e is the eccentricity of the orbit.

This method has been very successful and has yielded the discovery of $>$ 400 planets to date (Fig. 1.1). These detections include planets at large orbital separations, such as the 1.9 M_J planet HD 222155b orbiting its host star with a semi-major axis of 5.1 AU (Boisse et al., 2012).

With precise masses from radial velocity measurements and radii from transit measurements, the bulk density of a planet can be calculated. The masses and radii can also be plotted against models of different planetary compositions, such as those of Fortney et al. (2007) and Zeng & Sasselov (2013), to constrain the composition of the planet. Examples of such composition plots and mass-radius relations are presented in section 1.3.2.
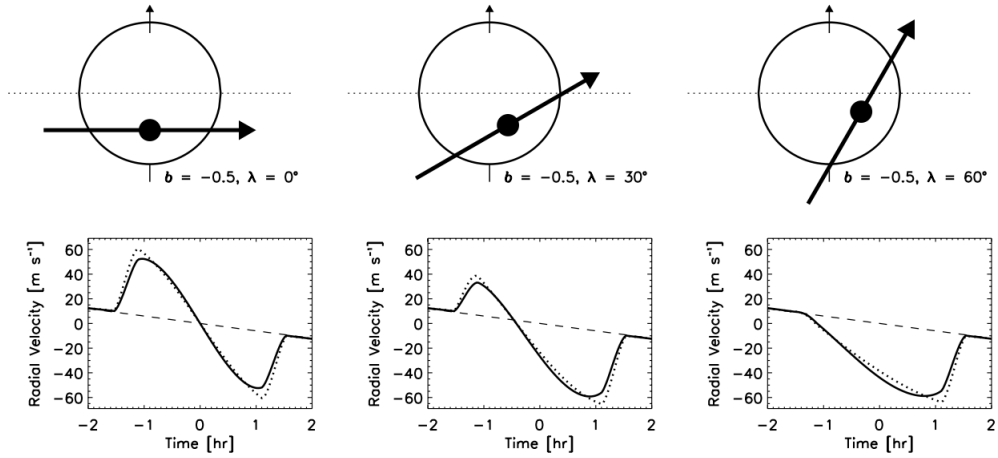
Figure 1.6: Schematic showing exoplanets with three different obliquities, λ. In each case the star is rotating anticlockwise so that the left-hand limb is blue shifted and the right-hand limb is red shifted. The bottom row shows the associated RM effects for each of the obliquities considered. Figure from Gaudi & Winn (2007).

Rossiter-McLaughlin Effect

The Rossiter-McLaughlin effect (Rossiter 1924, McLaughlin 1924) can be observed in the radial velocity measurements of a star during an exoplanet transit. Due to rotation of the star (providing the star is not viewed pole on), part of the star will be rotating towards the observer and part of the star will be rotating away from the observer causing a rotational Doppler shift. As a planet transits across the star, it will block part of this Doppler shifted light, causing a deviation in the measured radial velocity compared to the true velocity (Fig. 1.6). Using this technique it is possible to determine the projected spin-orbit alignment of the system, that is the angle between the spin axis of the star and the sky-projected orbital plane of the planet. This quantity, also known as the obliquity, is a useful quantity to know when characterising an exoplanetary system, since it has possible implications for the planet's migration history (as discussed in section 1.4).

The first detection of the Rossiter-McLaughlin effect due to an exoplanet was made by Queloz et al. (2000) through observing the spectra of HD 209458 using the ELODIE spectrograph.

The Rossiter-McLaughlin effect is not the only method to calculate the obliquities of star-planet systems however. Asteroseismology can be used to measure the inclination of the stellar rotation axis of solar like oscillators (Gizon & Solanki, 2003) which has been used to measure the obliquities of a handful of systems (e.g., Chaplin et al. 2013; Lund et al. 2014). Another method to measure the obliquity of a system

11

is through observations of star spot crossings in consecutive transits, as has been done for e.g. WASP-52b (Mancini et al., 2017) and Qatar-2b (Močnik et al., 2017).

An in-depth review of the current state of Rossiter-McLaughlin measurements is presented by Triaud (2017).

1.2.3 Direct imaging

Direct imaging is one of the more difficult methods of detection and as such fewer exoplanets have been detected via this method, with 44 detections to date (NASA Exoplanet Archive, 13/02/2018), although only 28 of these have masses $\leq 13\,\mathrm{M_J}$, the usual definition of the transition between giant planets and low mass brown dwarfs (e.g. Bodenheimer et al. 2013).

Exoplanets are very difficult to resolve in the optical because of the overwhelmingly bright light of the host star. There has, however, been success with this method in the infrared, particularly for young planets which are hot from their formation and can be detected from their own infrared emission. This method yielded the detection, in 2004, of a $5\,\mathrm{M_J}$ mass planet orbiting the brown dwarf 2MASSW J 1207334-393254 in the near-infrared using the NACO instrument on the Very Large Telescope (VLT; Chauvin et al., 2004). Another notable discovery using this technique was of the four planets orbiting the star HR 8799 (Marois et al., 2008, 2010).

Direct imaging is biased towards exoplanets with large orbital semi-major axes, which adds to the diversity of known planets as the transit and radial velocity methods are insensitive to such planets (see sections 1.2.1 and 1.2.2). Dedicated high-contrast imaging instruments such as the Spectro-Polarimetric High-contrast Exoplanet Research instrument on the VLT (SPHERE; Beuzit et al., 2008) and the Gemini Planet Imager (GPI; Macintosh et al., 2008) have revealed only a couple of detections to date, those of 51 Eridani b with GPI (Macintosh et al., 2015) and HIP 65426 b with SPHERE (Chauvin et al., 2017). It is hoped that new high-contrast imaging instruments on the upcoming generation of extremely large telescopes (ELTs) will increase the number of directly imaged planets (Baudoz, 2017), although current studies suggest that the number of stars with giant planets on wide orbits is low, with occurrence rates of $\sim 1\%$ (e.g. Bowler, 2016; Bowler & Nielsen, 2018).

1.2.4 Gravitational Microlensing

Gravitational microlensing occurs when a lensing object amplifies the light of a more distant background source along the same line of sight. It was first predicted by Einstein as a result of his theory of general relativity but he believed the effect would be too rare to be observed (Einstein, 1936). However, this effect is observable, with the first detection of microlensing by a stellar object coming in 1989 (Irwin et al., 1989). In 2003, the first detection of a planet using microlensing was made, that of the $1.5\,M_J$ planet OGLE-2003-BLG-235L b with a semi-major axis of $\sim 3\,AU$ (Bond et al., 2004). The planet's signal was seen as a short-duration deviation, caused by a low mass binary companion, in the longer term amplification of the source by the primary (Bond et al., 2004).

Since that first detection, microlensing has found a total of 35 exoplanets (NASA Exoplanet Archive, 14/02/2018) thanks to surveys such as OGLE (Udalski, 2003) and MOA (Bond et al., 2001), which detect 1000s of candidate microlensing events each year[2]. Due to the dependence on favourable and improbable orientations of objects, these are rare events. In order to increase the chances of observing a favourable orientation, it is necessary to observe a region with a high background of stars. As a result, observing the Galactic bulge offers the best opportunity to detect these events. This method has a peak sensitivity to objects with orbital separations of $\sim 1\,AU$, which is a distinct advantage to this method and could be the best opportunity to measure the statistics of planets at $\sim 1\,AU$ separations. Microlensing is also sensitive to low mass planets, such as the $< 2\,M_\oplus$ planet orbiting a late M dwarf host with a separation of $\sim 0.8\,AU$ (Gould et al., 2014).

The recently begun KMTNet microlensing survey (Kim et al., 2016) is expected to discover ~ 10 planets per year with masses $0.1 \leq M_P/M_\oplus \leq 5$ (Henderson et al., 2014). Future space missions such as NASA's WFIRST (Spergel et al., 2015) and ESA's Euclid (Laureijs et al., 2011) are predicted to find 1000s more planets via microlensing (Penny et al., 2013; McDonald et al., 2014; Spergel et al., 2015).

1.2.5 Astrometry

The astrometric method of exoplanet detection relies on very precise determination of a star's position. Due to the same reflex motion as detected by the radial velocity method, the change in the position of a star can be detected to indirectly reveal the presence of a planetary companion. To date, a single discovery has been made with astrometry, although the companion has a mass of $28\,M_J$ and is perhaps more

---

[2]e.g. from http://ogle.astrouw.edu.pl/ogle4/ews/ews.html, accessed 05/07/2018.

likely to be a brown dwarf (Sahlmann et al., 2013). Many more exoplanets should be detected via this method in the coming years, however, with Gaia (Perryman et al., 2001) predicted to detect thousands of exoplanets (Perryman et al., 2014) owing to its superlative 0.02 mas resolution (Gaia Collaboration et al., 2018).

1.3   The population of observed planets

Using the combination of methods outlined in section 1.2 we now know of 3,717 exoplanets (NASA Exoplanet Archive, 12/04/2018). Although shaped by strong selection effects, we can begin to consider the overall nature of the population of exoplanets.

The distributions of exoplanets with measured masses[3] and radii are shown in Fig. 1.7 and Fig. 1.8 respectively. From these figures, we can identify a few regions of interest and define some key terms.

Firstly, we can see there is a population of giant planets with masses around $1\,\mathrm{M_J}$ and radii of $\gtrsim 10\,\mathrm{R_\oplus}$ on short period orbits. These planets are known as *hot Jupiters*. The precise definition of a hot Jupiter varies across the literature but typically they are planets with masses $\gtrsim 0.3\,\mathrm{M_J}$ and periods $< 10\,\mathrm{d}$. The hot Jupiters occupy a fairly dense region in these plots owing to their relatively easy to detect transits, which have a comparatively high probability of occurring (see section 1.2.1), and masses through radial velocities (see section 1.2.2) even though they are inherently rare compared to other known classes of exoplanet (as discussed in section 1.3.1).

In addition to the hot Jupiters, there is a population of Jupiter mass planets at longer periods of $\sim 1000\,\mathrm{d}$. These planets have primarily been detected via the radial velocity technique as the transit method is less sensitive to such periods (see section 1.2.1). This is the reason why these planets are more numerous on the mass-period plot (Fig. 1.7) than the radius-period plot (Fig. 1.8) as few have measured radii.

At masses greater than $10\,\mathrm{M_J}$ Fig. 1.7 is relatively sparse. This region between $13\,\mathrm{M_J} < M < 80\,\mathrm{M_J}$ is called the 'Brown Dwarf Desert' (e.g. Marcy & Butler, 2000; Grether & Lineweaver, 2006; Triaud, 2017). It has been suggested that this desert could be due to the migration of brown dwarfs through an evolving protoplanetary disc to the point that they are destroyed through a merger with the primary (Armitage & Bonnell, 2002). In this scenario, the brown dwarf evolves contempo-

---

[3]Here measured masses are not just from radial velocities but also include other methods to measure the mass such as through transit timing variations (see section 1.2.1).
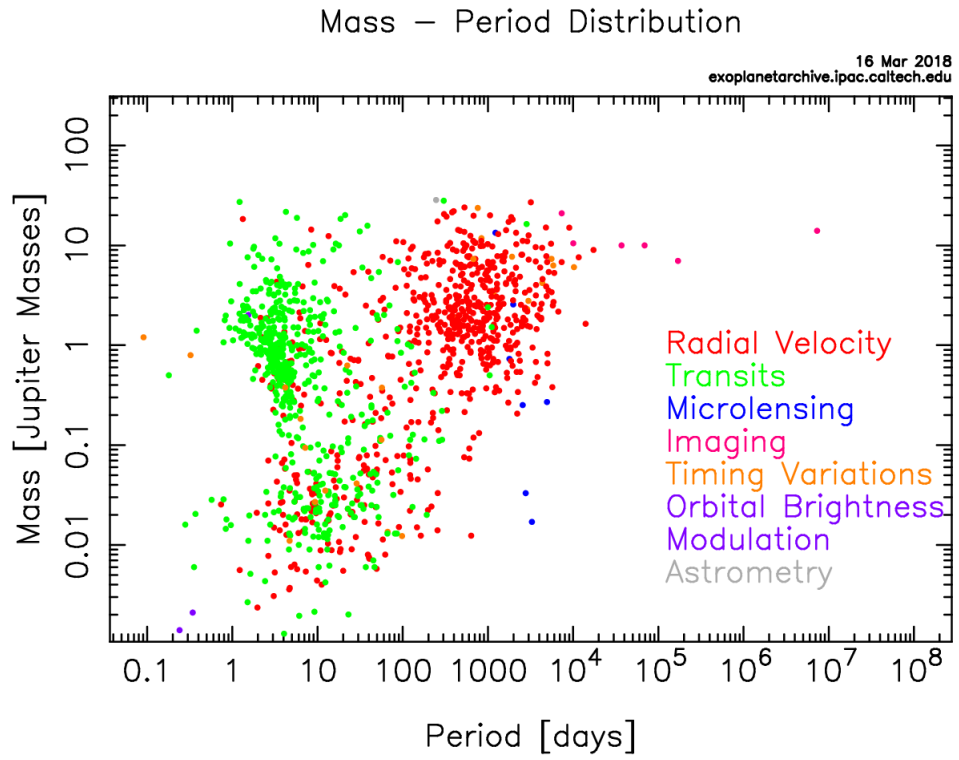
Figure 1.7: Plot of exoplanets with measured masses against their orbital periods, and coloured by the method of detection. Plot from NASA Exoplanet Archive, 24/03/2018.
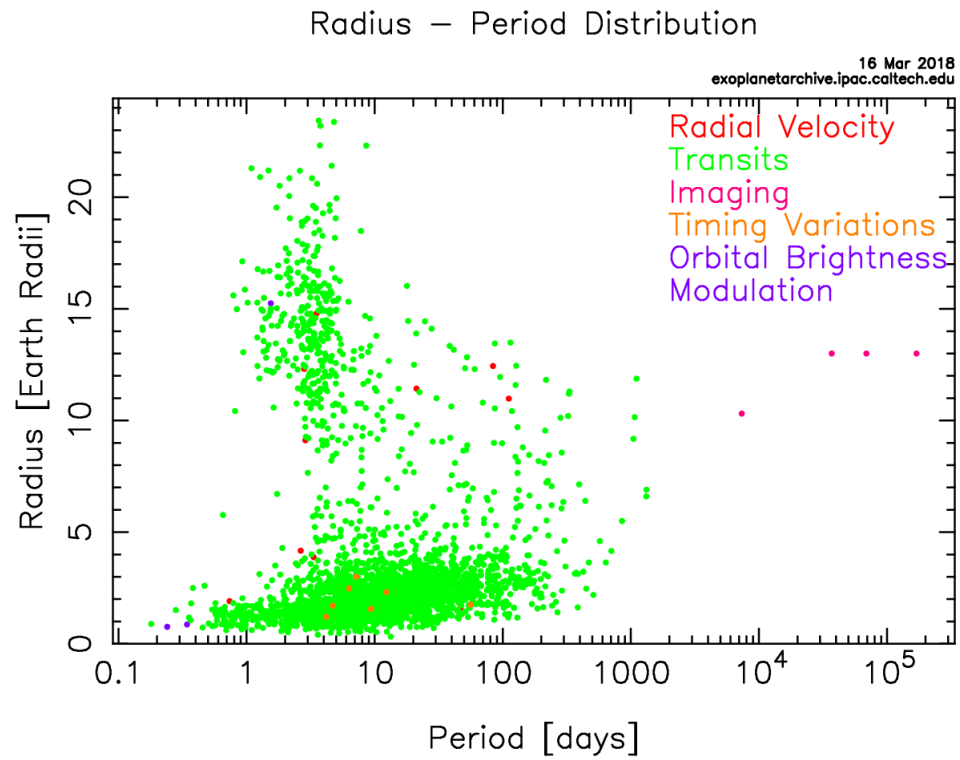
Figure 1.8: Plot of exoplanets with measured radii against their orbital periods, and coloured by the method of detection. Plot from NASA Exoplanet Archive, 24/03/2018.

raneously with the primary, while giant planets would avoid this fate by evolving later on near the dispersal of the disc, and thus do not drive migration through to merger (Armitage & Bonnell, 2002).

Moving to smaller planets, we can see from Fig. 1.8 that there is a vast number of planets with radii $\lesssim 5\,R_\oplus$, which have predominantly been detected with Kepler. These planets tend to orbit relatively faint host stars and so it is difficult to measure their masses, hence there are far fewer planets appearing in the mass-period plot (Fig. 1.7).

These smaller planets can be subdivided into *Neptunes* $(2 < R/R_\oplus < 6)$, *super-Earths* $(1.25 < R/R_\oplus < 2)$ and *Earths* $(R/R_\oplus < 1.25)$, where these definitions are from Borucki et al. (2011).

From Fig. 1.7 we can see that the Neptunes mostly cluster around periods of $10\,\mathrm{d}$, these are the so-called *warm Neptunes*. There is a lack of such planets on short period orbits. This is called the 'Neptune Desert' (e.g. Mazeh et al., 2016). When looking more closely at the radius distribution of smaller planets, we see there is a gap here too (Fulton et al., 2017). These features are believed to be the result of atmospheric evaporation and are discussed in more detail in section 1.4.2.

1.3.1 Occurrence rates

Exoplanets are common, with most Sun-like stars harbouring them (Winn, 2018). Thanks to the transit surveys mentioned in section 1.1 and the radial velocity surveys introduced in section 1.2.2, we are able to define the occurrence rates for exoplanetary systems based on the size of the planet and the spectral type of the star.

The occurrence rates of giant planets with periods less than 1000 days are shown in Fig. 1.9. This figure indicates that about 10% of solar-type stars host such a planet.

In this thesis, we are primarily concerned with hot Jupiters, as their atmospheres are amenable to characterisation (as discussed in section 1.5.4). Ground-based surveys, in particular, have discovered many hot Jupiters as they are most easily detected via the transit method (see section 1.2.1). However, as we can see in Fig. 1.9, hot Jupiters are rare, with occurrence rates of only $0.5 - 1.0\%$ (e.g. Mayor et al., 2011; Howard et al., 2012; Santerne et al., 2016).

At larger orbital separations, Rowan et al. (2016) showed that Jupiter analogs[4] are also rare, with an occurrence rate of 3%.

Smaller planets are more common. Howard et al. (2012) found that 13% of

---

[4]Here defined as planets with masses between 0.3 and $3\,\mathrm{M_J}$ and semi-major axes between 3 and $6\,\mathrm{AU}$.
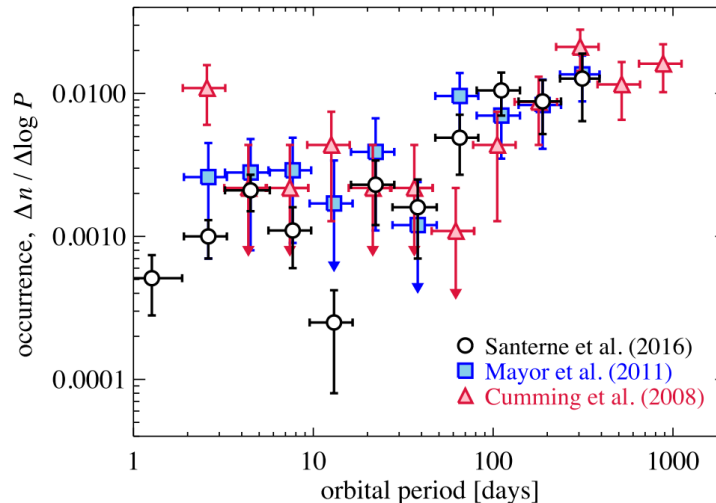
Figure 1.9: The occurrence rates of giant planets from RV surveys (Cumming et al., 2008; Mayor et al., 2011) and Kepler (Santerne et al., 2016). These occurrence rates are split into period bins with $\Delta \log P = 0.23$. Figure from Winn (2018).

G and K stars from the Kepler survey had a planet with a radius between 2 and $4\,\mathrm{R}_\oplus$ on an orbital period of less than 50 days. Interestingly, these authors also found that the occurrence of these planets was seven times higher for cool stars (3600-4100 K) than hot stars (6600-7100 K). Furthermore, in a study of the occurrence rate of planets around M stars using Kepler data, Dressing & Charbonneau (2015) found that on average there are 2.5 planets per M dwarf with radii $1 - 4\,\mathrm{R}_\oplus$. Excitingly for the prospects of detecting life in the Galaxy, these authors also found an occurrence rate of 0.16 Earth-sized ($1 < R < 1.5\,\mathrm{R}_\oplus$) planets in the habitable zone per M dwarf. This implies the nearest potentially habitable non-transiting and transiting planets are 2.6 and 10.6 pc away, respectively. Since this study, a number of remarkable habitable zone planet detections have been made. These include the detection of Proxima Centauri b (Anglada-Escudé et al., 2016), a non-transiting Earth-sized planet around the Sun's nearest stellar neighbour; Ross 128 b, a non-transiting exo-Earth around a quiet M dwarf at 3.4 pc; and the TRAPPIST-1 system of seven near-resonant transiting Earth sized planets, any one of which could be potentially habitable (Gillon et al., 2016, 2017). However, several factors will impact the potential habitability of these planets, such as potentially large day-night temperature contrasts resulting from tidal locking; stellar activity and XUV irradiation, which is thought to have significantly altered the TRAPPIST-1 planets' primary and secondary atmospheres (Wheatley et al., 2017b).

A good review of exoplanet occurrence rates is provided by Winn (2018).

1.3.2 Mass-radius relation

With radii resulting from transits (section 1.2.1) and masses from RVs (section 1.2.2) and TTVs (section 1.2.1) the bulk density of an exoplanet can be calculated. By comparing the measured mass and radius to mass-radius relations from models of differing compositions, we can estimate the likely composition of the planet from pure iron to pure hydrogen, with various combinations in between (e.g. Fortney et al., 2007; Seager et al., 2007; Zeng & Sasselov, 2013).

Weiss & Marcy (2014) studied the masses and radii of a sample of 65 planets with radii $< 4\,\mathrm{R}_\oplus$ and periods < 100 days and found that density increases with planetary radius up to $1.5\,\mathrm{R}_\oplus$ before dropping off towards larger radii. The authors interpreted the lower density planets as being comprised of rocky cores beneath large amounts of volatiles. However, they also noted the large scatter between radius and density.

Fig. 1.10 shows a recent mass-density plot and indicates how the density can reveal information about the bulk composition of the planet. However, these estimates of the composition tend to be degenerate as the same bulk density can be obtained with different combinations of these components. This is not a direct measurement and the only way currently to directly measure the internal composition of exoplanets is during the post-main sequence lifetime of the star (as discussed in section 1.4.3). This figure also shows there is a population of giant planets that have densities much lower than predicted by hydrogen and helium compositions, which is discussed in the next section.

1.3.3 The radius anomaly of giant planets

The population of hot Jupiters displays a wide range of radii, which are often much larger than predicted by models of planets with H/He envelopes (e.g. Charbonneau et al., 2000; Laughlin, 2018). This in turn causes them to have very low densities. Indeed, examples such as WASP-54b (Faedi et al., 2013), with a mean density of $0.187\,\mathrm{g\,cm}^{-3}$, and WASP-127b, with a mean density of just $0.09\,\mathrm{g\,cm}^{-3}$ (Lam et al., 2017), are far less dense than Jupiter (whose mean density is $1.33\,\mathrm{g\,cm}^{-3}$). These 'inflated' hot Jupiters are particularly interesting for atmospheric studies as they are expected to have large scale heights (as discussed in section 1.5.1).

This range of radii is shown as a function of the equilibrium temperature of the planet (which is dependent on insolation and is discussed in section 1.5.1) in Fig. 1.11 (left-hand panel). From this figure, it is clear that the radii of giant planets are correlated with the equilibrium temperature, such as was found in the Kepler

19
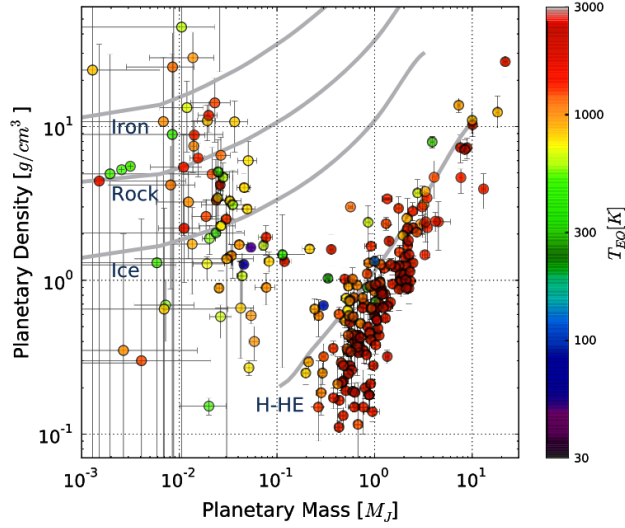
Figure 1.10: A plot of density against mass for exoplanets with measured masses and radii, and coloured by equilibrium temperature (see section 1.5.1). This plot extends from terrestrial planets to giant planets and has several different composition models overlaid. Figure from Laughlin (2018).

giant planets by Demory & Seager (2011). However, when considering the difference between the observed and predicted radii, it is clear that the models under-predict the radii of hotter planets and over-predict the radii of cooler planets, although there is scatter in this relation (Fig. 1.11, right-hand panel). This is known as the 'radius anomaly' of giant planets (e.g. Burrows et al., 2007).

Theories seeking to explain the radius anomaly broadly fall into two categories; those which require an internal heat source to inflate the planet and those that require the contraction of the planet to be slowed post-formation.

In the first category, theories invoking tidal heating by orbital circularisation have been suggested (e.g. Bodenheimer et al., 2001; Ibgui et al., 2010). However, Leconte et al. (2010) showed that this alone cannot explain the radii of the observed inflated hot Jupiters that have long since circularised. Another method invokes heating due to strong weather patterns driven by the conversion of stellar flux into kinetic energy in a planet's atmosphere (Guillot & Showman, 2002; Showman & Guillot, 2002), although this is would lead to a weaker dependence on temperature than observed (Laughlin et al., 2011). Batygin & Stevenson (2010) suggested that Ohmic dissipation could be responsible for this internal heat. In their model, zonal winds interact with the planet's magnetic field which induces an electric field. This drives currents into the planet's interior, which dissipate Ohmically causing resistive
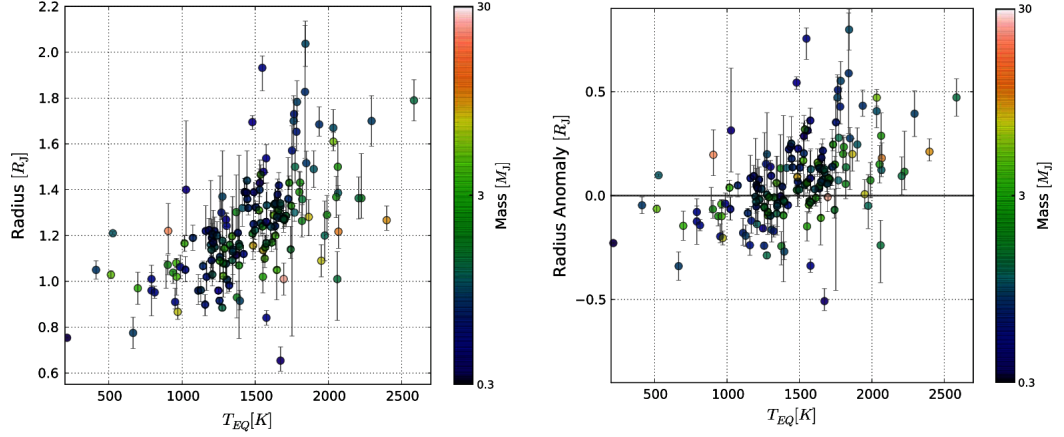
Figure 1.11: Left-hand panel: the radius versus equilibrium temperature for the population of giant planets and colour-coded by mass. Giant planets are defined here as planets with masses $> 100\,\mathrm{M_\oplus}$. Right-hand panel: the radius anomaly of the giant planets (measured as the difference between the actual radius and model-predicted radius) against planetary equilibrium temperature. Figure adapted from Laughlin (2018).

heating.

Theories that fall into the second category include enhanced opacities, either through clouds or increased metallicities, which would slow down the loss of heat (Burrows et al., 2007), or through gradients in composition that inhibit large scale convection and slow the rate of cooling (Chabrier & Baraffe, 2007). A difficulty these theories face, however, is that they don't depend on incident flux and therefore do not lead to the observed correlation with equilibrium temperature.

The radius anomaly of giant planets remains an open question. Recently, however, Thorngren & Fortney (2017) re-analysed the correlation between inflation and incident flux, finding the efficiency of inflation peaks at equilibrium temperatures of around $1500\,\mathrm{K}$ before falling off towards $2500\,\mathrm{K}$. They found that the Ohmic dissipation model of Batygin & Stevenson (2010) is the only inflation mechanism that predicts such a trend. In a recent review of the radius anomaly of giant planets, Laughlin (2018) also supports the theory of Batygin & Stevenson (2010) due to its ability to reproduce the observed correlation with equilibrium temperature.

1.4 Formation and Evolution of Planets

1.4.1 Formation

The idea that planets form from circumstellar material surrounding a recently collapsed star was first postulated in 1755 (Kant, 1755). This notion led to the development of the *nebular hypothesis*, which is now the commonly accepted formation mechanism for the solar system. In this theory, the Sun formed 4.567 billion years ago[5] (Connelly et al., 2012) with the gas giants forming in the first 10 million years prior to the dispersal of the circumsolar disc, and the terrestrial planets forming approximately 100 million years afterwards (e.g. Montmerle et al., 2006).

In the core accretion scenario of planet formation, the planets formed via low velocity collisions between particulate grains, growing into pre-planetary embryos (e.g. Lissauer, 1993). The continued collision and accumulation of these embryos formed planetesimals and subsequently cores. Within the snow line (the location in the disc at which volatiles such as H_2O can condense into solid ices, $\sim 4\,\mathrm{AU}$; Montmerle et al., 2006), this formed rocky cores. Beyond this line, however, ices contributed to the formation of much larger cores. Once these cores reached a critical mass of $\sim 10\,\mathrm{M_\oplus}$, they began runaway gas accretion and formed the gas giants (e.g. Pollack et al., 1996).

Theories of planet formation have been tested by the discovery of classes of exoplanet for which there is no equivalent in the solar system, such as warm Neptunes, super-Earths and, in particular, hot Jupiters.

The origin of hot Jupiters remains debated, with three possible scenarios leading to such planets: disc-migration, in-situ formation, or eccentric migration through tidal interactions (Dawson & Johnson, 2018).

Gravitational instability has been suggested as an alternative formation mechanism to core-accretion for giant planets. In this model, regions of the protoplanetary disc which are cool or dense enough will collapse to form self-gravitating clumps that could form the cores of giant planets (e.g. Boss, 1997; Durisen et al., 2007). However, it is now believed that gravitational instability cannot occur at the locations of hot Jupiters owing to the fact that the protoplanetary disc would be too hot (Rafikov, 2005). Furthermore, the positive correlation between host star metallicity and giant planet occurrence (Santos et al., 2003; Fischer & Valenti, 2005) is interpreted as evidence for core accretion (e.g. Winn, 2018). This is because a giant core will grow more rapidly in a more metal-rich disc. However, gravitational instability remains a viable formation mechanism for the long-period directly-imaged exoplan-

---

[5]This age being calculated from the radioactive dating of meteorites.

ets presented in section 1.2.3, where the timescales for core accretion are problematically long (e.g. Durisen et al., 2007; Dodson-Robinson et al., 2009; Vorobyov & Elbakyan, 2018).

While core-accretion has not been ruled out as a mechanism for in-situ formation, it is difficult for the core to grow adequately before the dissipation of the gas disc (Dawson & Johnson, 2018). Instead, it is far easier to form giant planets beyond the snow line as cores can grow quickly. For this reason, there has been much effort in theories involving the migration of giant planets to short orbital periods following their formation beyond the snow line.

In the theory of disc migration, a giant planet migrates before the dispersal of the protoplanetary gas disc by exchanging angular momentum with the disc (e.g. Goldreich & Tremaine, 1980; Lin et al., 1996; Baruteau et al., 2014). This is a smooth migration which results in little change in eccentricity. Recently, Nelson et al. (2017) found that disc migration is necessary to account for the semi-major axes of 15% of hot Jupiters discovered by the WASP (Pollacco et al., 2006) and HATNet (Bakos et al., 2004) surveys. Disc migration does, however, require a fine tuning of the disc's initial parameters (e.g. Dawson & Johnson, 2018).

Eccentric migration can occur due to interactions between the planet and another body, such as another planet (e.g. Rasio & Ford, 1996), star (e.g. Wu & Murray, 2003) or even the disc (e.g. Terquem & Ajmia, 2010). In this scenario, the perturbing body drives the eccentricity of the planet up, possibly through Kozai oscillations (Kozai, 1962; Wu & Murray, 2003). Once the planet is on a sufficiently eccentric orbit, it loses orbital energy through tidal interactions with the host star, causing the planet's orbit to shrink and circularise (e.g. Dawson & Johnson, 2018). This method could play an important role in the presence of hot Jupiters, with 85% of the WASP and HATNet hot Jupiters potentially undergoing this process (Nelson et al., 2017).

Initially it was thought that the migration history of these planets could be recovered by measuring the projected spin-orbit alignment (see section 1.2.2; e.g. Dawson, 2014 and references therein). If the planet experienced a smooth migration through a protoplanetary disc it was expected to have a small spin-orbit misalignment, whereas a planet which underwent a violent period of migration, possibly due to interactions with a third body, should show a larger spin-orbit misalignment. However, Dawson & Johnson (2018) point out that there are other ways to change a planet's obliquity which are unrelated to its migration history, such as through stellar tides raised by the planet (e.g. Winn et al., 2010) and through misalignment of the disc through interactions with a binary partner (e.g. Batygin,
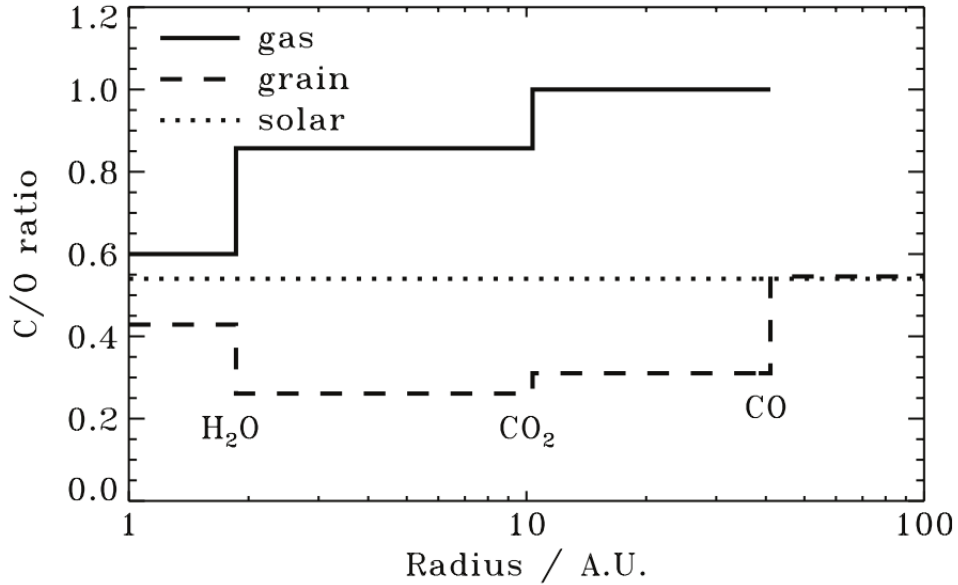
Figure 1.12: The C/O ratios of gas and dust in a protoplanetary disc around a Sun-like star. The H_2O, CO_2 and CO snow lines are labelled. Figure from Öberg et al. (2011).

2012). This complicates the use of obliquities to infer formation scenarios.

Measuring the C/O ratio in an exoplanet's atmosphere has been suggested as a route to calculating where in the disc a planet formed (Fig. 1.12; Öberg et al., 2011). Fig. 1.12 suggests that planets forming within the H_2O snow line should have near solar C/O ratios whereas planets forming further out in the disc should be more carbon-rich (Öberg et al., 2011). Beyond the H_2O snow line there is less oxygen in the gas, which is instead in ices, increasing the relative carbon abundance of the gas accreted by the planet. However, measurements of the atmospheric C/O ratio are hampered by the effects of clouds and hazes (e.g. Sing et al., 2016 and which are discussed in more detail in section 1.5.3). The C/O ratio can also be influenced by inhomogeneities in the C/O ratio that are inherent to the disc, in addition to the time-dependent chemical properties of the disc at the location of a planet and settling within a planet's atmosphere (e.g. Madhusudhan et al., 2016 and references therein).

1.4.2 Evaporation

As alluded to in section 1.3, evaporation of exoplanet atmospheres is thought to have played a major role in shaping the population of planets we observe.

Evaporation of close-in exoplanets occurs due to high energy extreme ultra-violet and X-ray photons that are incident on the planet's upper atmosphere (e.g. Lammer et al., 2003). This radiation heats the atmosphere, accelerating hydrogen to velocities greater than the escape velocity. As this hydrogen escapes, it can drag heavier elements away from the planet through hydrodynamic escape. This effect was first seen in observations of the stellar Lyman α line at UV wavelengths during the transit of HD 209458b by Vidal-Madjar et al. (2003). Since then, this effect has been observed in other exoplanets including HD 189733b (Lecavelier Des Etangs et al., 2010) and the warm Neptune GJ 436b, which was observed to have a transit depth of 56.5% in the UV (Ehrenreich et al., 2015).

As noted in section 1.3, there is a lack of Neptune-mass planets on short period orbits. This 'Neptune Desert' can be more clearly seen in Fig. 1.13. The origins of this desert have been linked to atmospheric evaporation. In this scenario, while hot Jupiters are evaporating, they are not evaporating quickly enough to remove their envelopes entirely, while Neptune mass planets could have their H/He envelopes completely stripped (e.g. Owen & Jackson, 2012). Planets falling initially within this desert would therefore quickly migrate to lower masses, beneath the lower boundary of the desert.

Evaporation has also been suggested as a cause for the bimodal radius distribution of small planets that was recently discovered in Kepler planets, with few having radii between $1.5 - 2.0\,R_{\oplus}$ (Fig. 1.14; Fulton et al., 2017). This 'evaporation valley' was predicted before its discovery, with two populations of small planets postulated; those that were completely stripped of their H/He envelopes and those which retained an envelope of $\sim 1\%$ in mass (Owen & Wu, 2013; Lopez & Fortney, 2013).

1.4.3 Post-Main Sequence Evolution

Following a host star's main sequence lifetime, the majority of stars will evolve into white dwarfs. Owing to the extreme surface gravities of white dwarfs, their atmospheres are expected to be composed of pure hydrogen or helium. However, 30% of white dwarfs show heavy elements in their spectra, implying they are actively accreting this material as it would otherwise quickly sink below the surface (e.g. Zuckerman & Young, 2017). This material is presumed to come from disintegrating planetesimals destroyed during the post-main sequence evolution of the host, with recent observational evidence for such planetesimals coming via transits of material orbiting the white dwarf WD 1145+017 (Vanderburg et al., 2015). The abundance of this material informs us about the bulk compositions (e.g. Hollands et al., 2018)
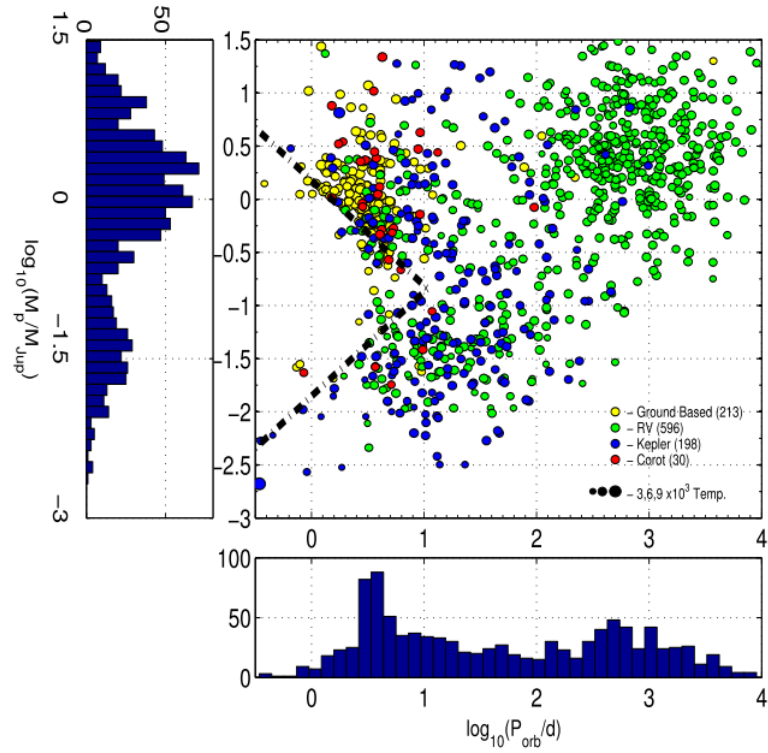
Figure 1.13: A plot of the masses and periods of exoplanets with measured masses. The dashed lines indicate the extent of the 'Neptune Desert' with the colours of the points indicating the detection method and their size proportional to the host star's temperature. Figure from Mazeh et al. (2016).
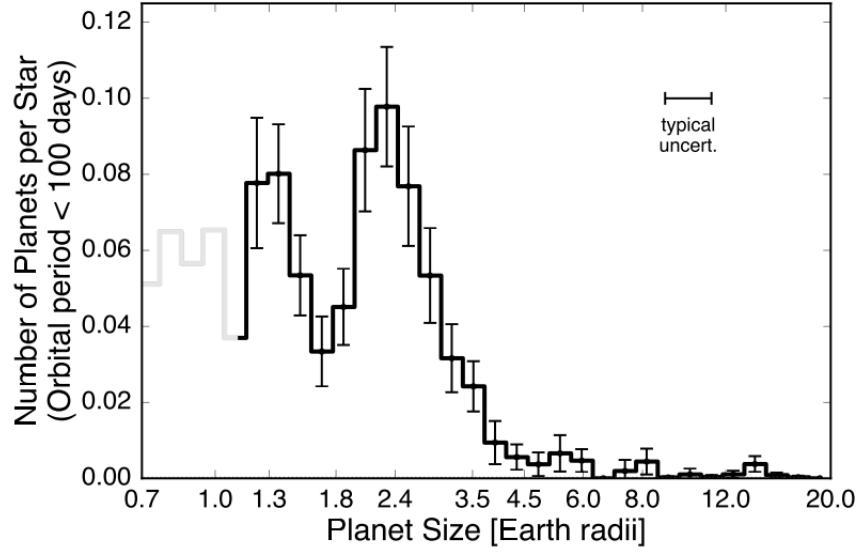
Figure 1.14: The radius gap of close-in small planets detected by Kepler. The distribution in radius shows small planets with radii between $1.5 - 2.0\,\mathrm{R}_{\oplus}$ are rare and has been interpreted as evidence for photo-evaporation. Figure from Fulton et al. (2017).

and C/O ratios of the planetesimals (e.g. Wilson et al., 2016). Polluted white dwarfs currently represent the only way to directly measure the bulk compositions of exoplanets.

1.5 Exoplanet atmospheres

1.5.1 Basic properties

The equilibrium temperature

When considering exoplanet atmospheres, it is useful to define the exoplanet's *equilibrium temperature* (T_{eq}) which, along with an exoplanet's surface gravity and atmospheric metallicity, plays a key role in determining the main atmospheric characteristics.

The equilibrium temperature is an estimate of the effective temperature of the planet without any internal luminosity and is the temperature of an isothermal planet once it has reached equilibrium with radiation from its host star (Seager, 2010). The effective temperature is the temperature of a blackbody at the distance of the planet with the same total flux as the planet. The equilibrium temperature can be derived by equating the power radiated by the planet with the power incident

on the planet.

The power incident on the planet's surface, P_{in}, is given by

$$P_{in} = (1 - A_B)F_*\pi R_P^2 \tag{1.16}$$

where A_B is the Bond albedo, F_* is the flux of the host star at the location of the planet and R_P is the radius of the planet. The flux of the star is given by

$$F_* = \frac{L_*}{4\pi a^2} \tag{1.17}$$

where L_* is the star's luminosity and a is the planet's semi-major axis. L_* is given by the Stefan-Boltzmann law as

$$L_* = 4\pi R_*^2 \sigma_{SB} T_*^4 \tag{1.18}$$

where σ_{SB} is the Stefan-Boltzmann constant, R_* is the radius of the star and T_* is its temperature.

Under the assumption that the heat is perfectly redistributed around the planet, the outputted power of the planet is given by its luminosity, L_P, as

$$L_P = 4\pi R_P^2 \sigma_{SB} T_P^4 \tag{1.19}$$

where T_P is the planet's temperature.

As we are interested in the case that $P_{in} = L_P$, such that $T_P \equiv T_{eq}$, we can equate equation 1.16 to equation 1.19. Substituting in equation 1.17 and 1.18 and re-arranging gives

$$T_{eq} = (1 - A_B)^{1/4}T_* \frac{1}{\sqrt{2}}\left(\frac{R_*}{a}\right)^{1/2} \tag{1.20}$$

The temperature of an exoplanet has important effects for it's atmospheric components and circulation. As we shall see in section 1.5.3, there have been suggestions that the temperature can affect whether an exoplanet is cloudy or hazy (Heng, 2016; Stevenson, 2016; Crossfield & Kreidberg, 2017; Fu et al., 2017), while also determining what kind of clouds can form (e.g. Morley et al., 2015; Wakeford et al., 2017b). Recently, exoplanets with very high equilibrium temperatures ($\gtrsim 2500\,\mathrm{K}$) have revealed evidence for temperature inversions in their atmospheres (Haynes et al., 2015; Evans et al., 2017; Sheppard et al., 2017; Arcangeli et al., 2018).

Atmospheric scale height

The atmospheric scale height, H, is defined as the length scale over which the pressure drops by a factor of e and is derived through the equation for hydrostatic equilibrium and the ideal gas law. It is given by

$$H = \frac{kT}{g\mu} \qquad (1.21)$$

where k is Boltzmann's constant, T is the temperature of the planet, g is the acceleration due to gravity and μ is the mean molecular mass, which is often assumed to be 2.3 times the mass of a proton for a Jupiter-like composition of $\sim 85\%$ H_2 and $\sim 15\%$ He (Niemann et al., 1996).

The atmospheric scale height is a useful quantity to consider when selecting targets for atmospheric characterisation as larger scale heights should lead to larger absorption signals from the planet's atmosphere.

For this reason, studies are often focused on inflated hot Jupiters; planets with very low density due to their large radii and relatively low masses (see section 1.3.3). This causes them to have relatively large atmospheric signals. As a result of their low densities, low surface gravities, and high temperatures, these planets have large atmospheric scale heights. The current largest scale height of any exoplanet (as of 08/02/2017) is that of the hot Jupiter KELT-11b (Pepper et al., 2017), with an atmospheric scale height of 2400 km. This is far larger than any solar system planet's scale height. Saturn has the largest scale height of any solar system planet with a scale height of 59.5 km, while Jupiter has a scale height of 27 km and Earth just 8 km[6].

The outermost ~ 5 scale heights may account for up to 10% of the cross-sectional area of a hot Jupiter (Brown, 2001), where different atmospheric species can affect the observed *transmission spectrum*, which is described in the following section.

1.5.2 Methods of characterisation

Transiting exoplanets currently offer the best opportunity for atmospheric characterisation. Through observations of the transit, secondary eclipse and variations in the phase curve (Fig. 1.15), we can measure the composition and dynamics of an exoplanet's atmosphere.

Although I focus on the methods available to characterise transiting exoplanet atmospheres in this section, there do exist methods for detecting the atmo-

---

[6]All data from https://nssdc.gsfc.nasa.gov/planetary/factsheet/, accessed 08/02/2017
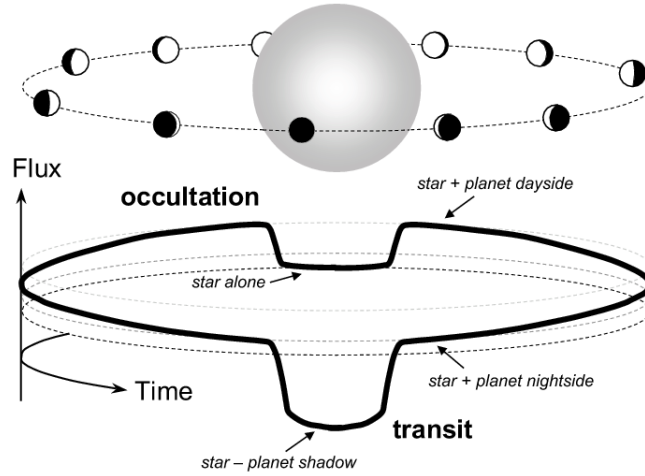
Figure 1.15: Schematic of a transiting exoplanet's phase curve showing the transit at inferior conjunction, the occultation/secondary eclipse at superior conjunction and the constant variation in the overall flux due to the planet's dayside rotating into and out of view. Figure from Winn (2010).

spheres of non-transiting exoplanets.

One technique for analysing the signal from the atmospheres of unresolved, non-transiting exoplanets is to use high resolution cross correlation to extract the planetary signal from the bright stellar host's signal. This technique was attempted by Collier Cameron et al. (1999) to detect the reflected starlight by the planetary companion to τ Boo. Subsequent studies by Charbonneau et al. (1999) and Leigh et al. (2003) could not confirm the earlier detection but instead placed upper limits on the albedo of the planet τ Boo b. Carbon monoxide was later detected in the dayside of this planet by Brogi et al. (2012) using VLT/CRIRES to extract the planet's Doppler shifted CO absorption profile, a method pioneered by Snellen et al. (2010) in their detection of CO in the atmosphere of HD 209458 b. This technique has also revealed detections of CO and H_2O in the non-transiting exoplanet HD 179949b (Brogi et al., 2014) and recently H_2O in the atmosphere of 51 Pegasi b (Birkby et al., 2017).

Transmission Spectroscopy

As the exoplanet transits across the face of its host star (Fig. 1.15), sources of opacity within the planet's upper atmosphere raise the altitude at which the atmosphere becomes optically thick. This causes a larger fraction of the stellar light to be blocked and therefore a deeper transit is observed. By observing an exoplanet's transit as a function of wavelength, a technique called *transmission spectroscopy*, we
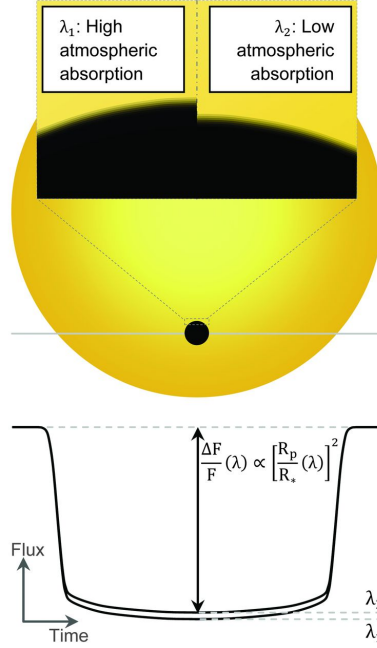
Figure 1.16: Diagram showing the principle of transmission spectroscopy. Sources of opacity raise the altitude at which the atmosphere becomes optically thick causing a variation in the transit depth with wavelength. Figure from de Wit & Seager (2013).

can detect the sources of opacity within the planet's upper atmosphere, at pressures of $\sim$ mbar to μbar (Crossfield, 2015). This is shown pictorially in Fig. 1.16.

Transmission spectroscopy using Hubble Space Telescope (HST)/Space Telescope Imaging Spectrograph (STIS) was used to detect the first spectral feature of any exoplanet, that of the narrow core of atomic sodium originating from the upper atmosphere of HD 209458b (Charbonneau et al., 2002). Since this first discovery, many exoplanets have been studied in transmission and I discuss the main findings of these studies in section 1.5.3.

Secondary eclipses and phase curves

As an exoplanet passes behind its star, and is obscured from view, we observe a secondary eclipse (Fig. 1.15). By observing the depth of the secondary eclipse, a technique called *occultation spectroscopy*, we can calculate the brightness temperature of the planet as a function of wavelength, which can reveal information about the planet's vertical temperature profile (e.g. Haswell, 2010). Secondary eclipse observations probe the deeper, hotter layers of an exoplanet's atmosphere and are sensitive to pressures of $\sim$ bar to mbar (Crossfield, 2015).

Secondary eclipse observations are often performed at infrared wavelengths where the contrast ratio between the stellar and planetary flux is smallest. As a result, the Spitzer Space Telescope made the first detections of exoplanet thermal emission, those of HD 209458b (Deming et al., 2005) and TrES-1b (Charbonneau et al., 2005). Since these early studies, a number of exoplanets have been studied at secondary eclipse both from space (e.g. Seager, 2010; Rostron et al., 2014; Nikolov et al., 2018) and the ground (e.g. López-Morales et al., 2010; Burton et al., 2012; Martioli et al., 2018). Recent highlights in this field include the first robust detection of a temperature inversion in an exoplanet atmosphere (that of the 2500 K WASP-121b; Evans et al., 2017) and a precise metallicity measurement of WASP-18b through considerations of the opacity from H^- ions (Arcangeli et al., 2018).

In addition to the transit and secondary eclipse, an exoplanet's phase curve can be used to characterise its atmosphere. As the planet's day side rotates into and out of view, it causes the phase curve to modulate on the planet's orbital period (see Fig. 1.15). This can reveal information about circulation within an exoplanet's atmosphere in the infrared and planetary albedo in the optical (e.g. Knutson et al., 2007; Crossfield, 2015).

Since hot Jupiters' rotations are expected to be tidally-locked to their orbit, they can develop large day-night temperature contrasts. Early models predicted that such planets should develop strong, super-rotating eastward winds at the equator, which shift the hottest point in the atmosphere away from the substellar point (Showman & Guillot, 2002). Indeed, the first measured phase curve of a planet, that of the hot Jupiter HD 189733b, showed an atmospheric hotspot shifted by 30 degrees to the east of the substellar point (Fig. 1.17; Knutson et al., 2007). Following this measurement, the first spectroscopic study of an exoplanet's phase curve was performed using HST/WFC3 measurements of the hot Jupiter WASP-43b (Stevenson et al., 2014).

A recent highlight from phase curve analysis includes the first detection of changing weather patterns in an exoplanet atmosphere. This was revealed through the continually changing location of the hotspot in the atmosphere of the hot Jupiter HAT-P-7b, which was seen in Kepler data (Armstrong et al., 2016).

1.5.3 Sources of opacity

Atoms and molecules

Early models of close-in extrasolar giant planets predicted that the absorption lines of the alkali metals of neutral sodium (Na I) and potassium (K I) would be strong
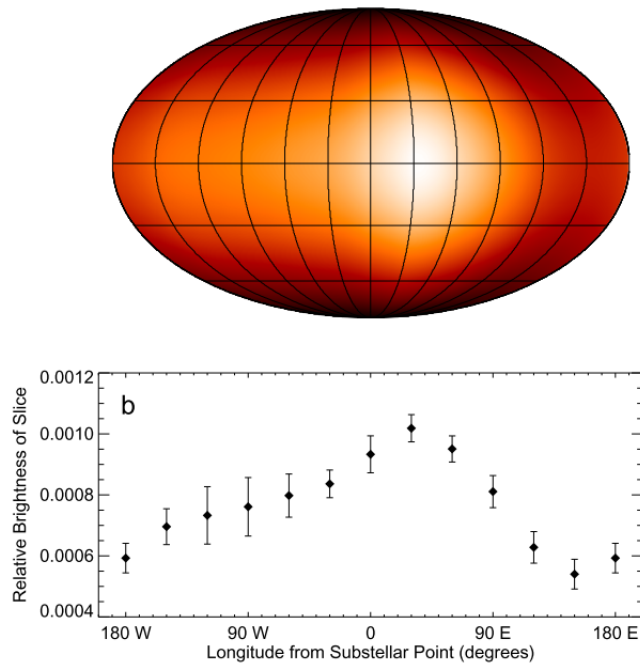
Figure 1.17: The brightness map of HD 189733b constructed from Spitzer measurements of the planet's phase curve and secondary eclipse. The bottom panel shows how the brightest point is offset by 30 degrees to the east of the substellar point. Figure from Knutson et al. (2007).
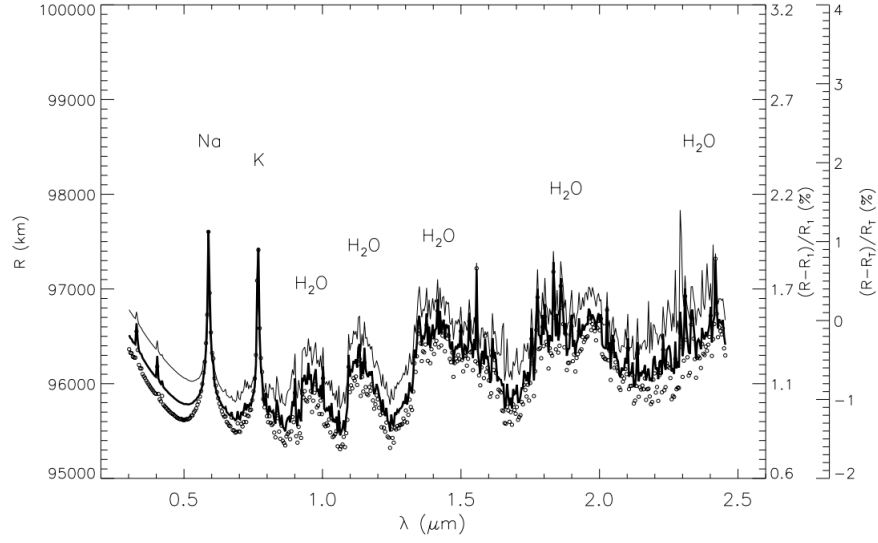
Figure 1.18: This figure shows a model transmission spectrum of a hot Jupiter for three different temperature-pressure profiles, from cooler (top line) to hotter (bottom line). The planet's apparent radius (R on the y-axis) shows strong variations associated with Na, K and H_2O absorption and Rayleigh scattering (shortwards of 0.5 microns). Figure from Hubbard et al. (2001).

absorbers in the optical (Seager & Sasselov, 2000; Sudarsky et al., 2000; Brown, 2001; Hubbard et al., 2001). This was based on models of cool dwarfs where other opacity sources, such as Ti, condense out of the atmosphere at the temperature of hot Jupiters, thereby removing them as opacity sources in the optical (e.g. Lodders, 1999). In the infrared, molecular opacity was predicted to be important, with water (H_2O), methane (CH_4) and carbon monoxide (CO) being the primary opacity sources. These studies also predicted the influence of Rayleigh scattering at short optical wavelengths. An example model absorption spectrum of such a planet is shown in Fig. 1.18.

These studies also predicted the effects of pressure broadening and clouds on the absorption line profiles. Pressure broadening is due to collisions between the emitter with other particles, which are hydrogen molecules in the case of hot Jupiters (Seager & Sasselov, 2000). These collisions disrupt the phase coherence of the emitted photons, giving rise to the pressure-broadened wings of sodium and potassium, which can be described by a Lorentzian profile (Seager, 2010). This effect can be seen in the line profiles of Fig. 1.18. However, as discussed later in this section, clouds can mask the wings of the alkali metals which originate at higher pressures than the line cores.

As noted in section 1.5.2, the first detection of an exoplanet atmosphere came

in 2002 and was of the narrow core of atomic sodium in the hot Jupiter HD 209458b (Charbonneau et al., 2002). However, the strength of the sodium feature detected was lower than predicted by atmospheric models (Seager & Sasselov, 2000). This led the authors to conclude that either clouds were present in the atmosphere, raising the altitude at which the atmosphere became optically thick and therefore reducing the amplitude of the feature, or that the abundance of sodium was inherently low.

Broad-band opacity sources were later detected in this planet's atmosphere in the form of the pressure-broadened wings of the sodium feature (Sing et al., 2008a) and a blueward slope in the transmission spectrum, which was interpreted as Rayleigh scattering by H_2 (Lecavelier Des Etangs et al., 2008b), and is discussed in more detail later in this section. The first ground-based detections of exoplanet atmospheres were also of the narrow sodium core in the atmospheres of HD 189733b (Redfield et al., 2008) and HD 209458b (Snellen et al., 2008).

Since these early studies, transmission spectroscopy has successfully revealed detections of many absorption features in exoplanet atmospheres. The alkali metals have been observed in many hot Jupiter atmospheres to date with the narrow cores often resolvable even in the presence of clouds as they originate at higher altitudes. Detections of sodium and potassium have been made both from space (e.g. Huitson et al., 2012; Nikolov et al., 2014; Sing et al., 2015, 2016; Fischer et al., 2016) and from the ground (e.g. Wood et al., 2011; Zhou & Bayliss, 2012; Wilson et al., 2015; Nikolov et al., 2016; Chen et al., 2017).

H_2O has been detected in a number of exoplanet atmospheres, primarily in hot Jupiters through HST/WFC3 observations of the $1.4\,\mu$m water bandhead (e.g. Huitson et al., 2013; Kreidberg et al., 2014b, 2015) although there have been detections of water in Neptunes too, such as in HAT-P-11b (Fraine et al., 2014) and HAT-P-26b (Wakeford et al., 2017a). Furthermore, some of these detections of water have been so robust that they have allowed the authors to constrain the metallicity of the exoplanets' atmospheres by using oxygen as a proxy for heavy element abundance (Kreidberg et al., 2014b; Fraine et al., 2014; Kreidberg et al., 2015; Wakeford et al., 2017a, 2018). This has allowed for comparisons between exoplanets' mass-metallicity relation and that of the solar system giant planets (Fig. 1.19).

Additionally, Na, CO and H_2O have been detected at high resolution in the atmospheres of HD 209458b and HD 189733b from the ground (e.g. Snellen et al., 2010; Louden & Wheatley, 2015; Wyttenbach et al., 2015; Brogi et al., 2016, 2018). Observations of the Doppler shifts in these lines have allowed the atmospheric wind speeds to be calculated, which have been found to be consistent with eastward jets (Snellen et al., 2010; Louden & Wheatley, 2015; Wyttenbach et al., 2015; Brogi
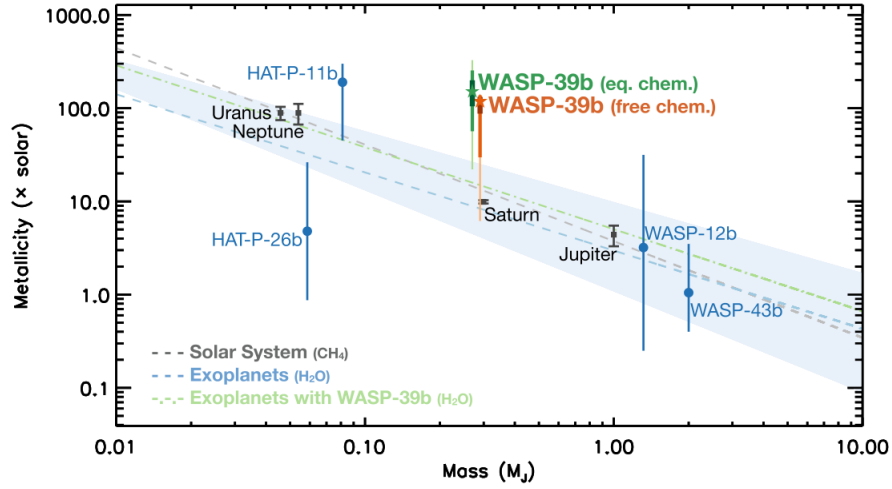
Figure 1.19: Plot of the mass-metallicity relation for exoplanets (as calculated through the abundance of water) compared to that of the solar system giant planets (as calculated through the abundance of methane). Figure from Wakeford et al. (2018).

et al., 2016).

Recently, there has also been the suggestion that absorption features in the transmission spectrum of HD 209458b could be due to ammonia (NH_3; MacDonald & Madhusudhan, 2017) although this molecule is not expected to be abundant at temperatures above 500 K (c.f. Fig. 1.24; Crossfield, 2015).

Titanium oxide (TiO) and vanadium oxide (VO) are important opacity sources in late-type stars and were predicted to cause temperature inversions in very hot Jupiters by trapping heat at relatively high altitudes. The temperature inversions would be apparent through the forcing of molecular bands into emission rather than absorption (Fortney et al., 2008). Due to the lack of detection of these molecules for several years, the role of TiO and VO in hot Jupiter atmospheres was called into question. However, more recent studies have revealed TiO in transmission (WASP-19b, Sedaghati et al., 2017b; WASP-33b, Nugroho et al., 2017) and possibly even VO in emission (WASP-121b, Evans et al., 2017). The presence of TiO and VO has also been inferred in the relatively cool WASP-127b (T_{eq} = 1400 K, Palle et al., 2017), although a more recent study by Chen et al. (2018) has called this result into question.

Rayleigh & Mie scattering

In addition to the relatively narrow features described in the previous section, scattering slopes extending across optical wavelengths have been detected in a number of exoplanets. These features were initially detected in the atmospheres of HD 189733b (Pont et al., 2008) and HD 209458b (Sing et al., 2008a), and interpreted as being caused by Rayleigh scattering by condensates and molecular hydrogen respectively (Lecavelier Des Etangs et al., 2008a,b). Since theses initial detections, scattering slopes have been detected in a number of exoplanets (e.g. Jordán et al., 2013; Sing et al., 2015; Kirk et al., 2017).

Rayleigh scattering (Rayleigh, 1899) describes the wavelength dependence of scattered light. It is only valid when the scattering particles are very small compared to the wavelength of the light, up to sizes approximately less than one tenth of the wavelength. The Rayleigh scattering cross-section is proportional to λ^{-4} and so is important for blue, optical wavelengths and gives rise to a bluewards slope in the transmission spectrum (Fig. 1.18). Rayleigh scattering is the reason that the sky is blue.

From the Rayleigh scattering slope in the transmission spectrum, Lecavelier Des Etangs et al. (2008a) showed that the atmospheric temperature can be calculated through

$$\alpha H = \frac{dR_P}{d\ln\lambda} \tag{1.22}$$

where H is the atmospheric scale height (and is dependent on T through equation 1.21), R_P is the planet's radius and $\alpha = -4$ for Rayleigh scattering. The logarithmic dependence on wavelength can be understood by first considering the optical depth, τ, of a planet's atmosphere at an altitude z, which is given by

$$\tau(\lambda, z) = n(z)\sigma(\lambda)x \tag{1.23}$$

where n is the number density of the absorbing species, σ is the cross section of the absorbing species and x is the path length along the line of sight grazing the planetary limb at the altitude z.

The number density at an altitude z is related to the number density at $z = 0$ through

$$n(z) = n_0 e^{-z/H} \tag{1.24}$$

Inserting equation 1.24 into equation 1.23 and re-arranging, we get

$$\ln \sigma(\lambda)^{-1} \propto -\frac{z}{H} \tag{1.25}$$

Since we are considering the slope in the transmission spectrum, we are interested in the change of altitude with wavelength. Assuming we know the variation of the cross section as a function of wavelength, equation 1.25 can be used to find

$$\frac{dz}{d\lambda} \propto H \frac{d\ln\sigma}{d\lambda} \tag{1.26}$$

Assuming a scaling relation of the form $\sigma = \sigma_0(\lambda/\lambda_0)^\alpha$ and given $dR_P = dz$, we recover the logarithmic dependence on λ given in equation 1.22.

While Rayleigh scattering is important for particles $< 1/10$ the wavelength of light, Mie scattering is important when the scattering particles are larger. Mie scattering is not strongly wavelength dependent and is valid when the scattering particles are homogeneous spheres (Seager, 2010). Mie scattering is the reason that clouds in Earth's atmosphere appear grey.

Clouds and hazes

As discussed earlier in this section, clouds were thought to be potentially present in the first detected exoplanet atmosphere (Charbonneau et al., 2002). Since then, clouds and hazes have been observed across the full range of exoplanetary parameter space, which can mute narrow opacity sources or sometimes mask them entirely (e.g. Line et al., 2013; Kreidberg et al., 2014a; Kirk et al., 2016). Indeed, they are so common that the pressure-broadened wings of both sodium and potassium have been observed in only a single hot Jupiter to date (WASP-39b; Fischer et al., 2016; Nikolov et al., 2016; Sing et al., 2016; Wakeford et al., 2018).

HD 189733b (Bouchy et al., 2005) is one of the most studied exoplanets to date and shows significant aerosols in its atmosphere (Pont et al., 2008; Sing et al., 2011a; Pont et al., 2013). In the case of this planet's atmosphere, the aerosols raise the altitude at which the atmosphere becomes optically thick, masking the pressure-broadened wings of sodium and potassium while at the same time giving rise to the Rayleigh slope that extends across visible and infrared wavelengths and over ~ 5 scale heights (Pont et al., 2008; Sing et al., 2011a; Pont et al., 2013). It is, however, still possible to resolve the narrow line core of the sodium feature which originates above the haze layer (e.g. Huitson et al., 2013; Louden & Wheatley, 2015) as Fig. 1.20 demonstrates.

HD 189733b has an equilibrium temperature of $\sim 1200\,\mathrm{K}$ (Southworth, 2011), however, aerosols have been observed in very hot exoplanets such as WASP-12b (Sing
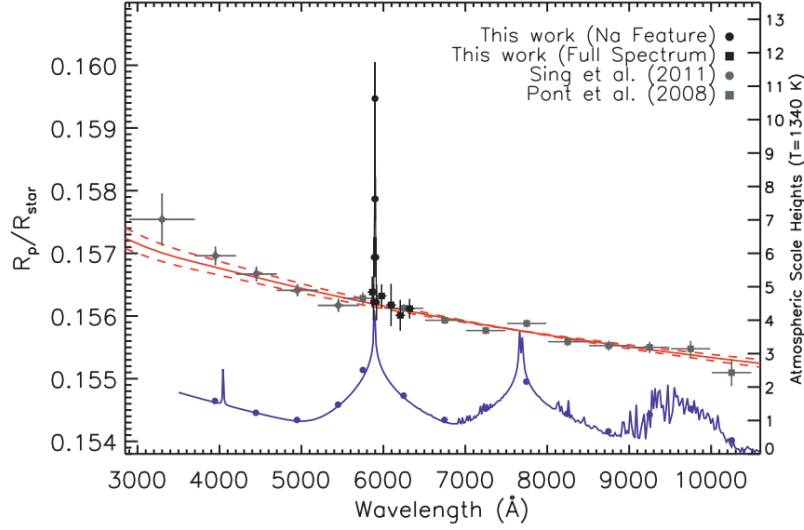
Figure 1.20: Transmission spectrum of the hot Jupiter HD 189733b (black points). In this case, a haze gives rise to a Rayleigh scattering slope (red line) which masks the pressure-broadened wings of sodium and potassium (blue line) although the sodium line core is still detectable. Figure from Huitson et al. (2013).

et al., 2013) ($T_{eq} = 2580$ K; Southworth, 2011) and relatively-cool exoplanets, such as GJ 1214b (Kreidberg et al., 2014a) ($T_{eq} = 547$ K; Southworth, 2011). GJ 1214b is perhaps the most precise transmission spectrum of any exoplanet to date owing to the 15 transits of the planet observed with HST producing errors in the transit depth of < 30 ppm (Kreidberg et al., 2014a). In the case of GJ 1214b, the transmission spectrum was completely flat (Fig. 1.21) and the authors interpreted this as a grey cloud layer.

Aside from GJ 1214b, clouds have produced flat transmission spectra in a number of other exoplanets (e.g. Gibson et al., 2013a; Mallonn et al., 2015b; Kirk et al., 2016; Louden et al., 2017). The prevalence of clouds and hazes was further highlighted in a sample of 10 hot Jupiters studied with HST and Spitzer (Fig. 1.22; Sing et al., 2016). In this paper, the authors found a continuum from cloudy to clear atmospheres with no correlation between the presence of aerosols with equilibrium temperature or surface gravity. They did, however, see a correlation between the amplitude of the water feature in the infrared and the presence of aerosols, which suggested the weaker spectral features were not due to primordial water depletion.

In early theoretical efforts, clouds and hazes were predicted to have important implications for the emergent spectra of exoplanets. Clouds are condensates and occur when a species' condensation curve intersects with a planet's temperature-
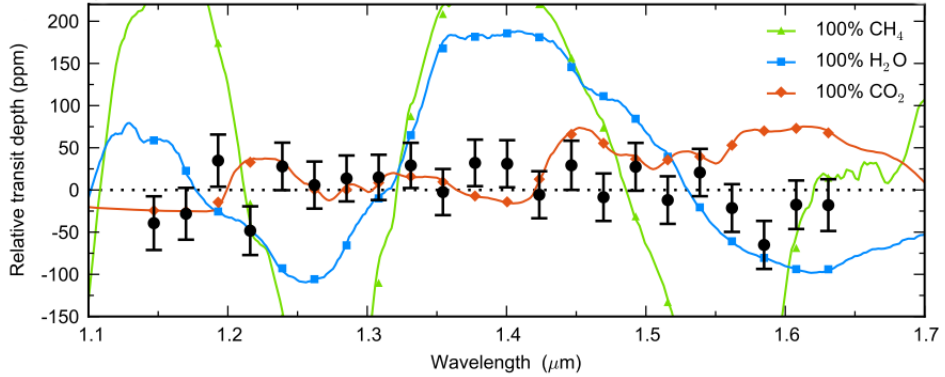
Figure 1.21: HST transmission spectrum of the super Earth GJ 1214b (Kreidberg et al., 2014a). Models of pure H_2O, CH_4 and CO_2 are significantly ruled out, indicating the presence of clouds in the planet's atmosphere.

pressure profile. In hot Jupiter atmospheres, these cloud particles may be liquid iron droplets or solid silicate particles (Seager, 2010).

A diagram of the expected cloud compositions for varying temperatures of giant exoplanet atmosphere is shown in Fig. 1.23. This figure demonstrates the effects of temperature on the transition from a Jupiter-like atmosphere to the atmospheres of T (800-1200 K), L (1200-2000 K) and M (2000-2500 K) dwarfs (Lodders, 2004). A similar sequence is predicted for giant exoplanet atmospheres (Seager, 2010; Marley et al., 2013). As with much of hot Jupiter atmospheric science, the theory is based on the atmospheres of brown dwarfs which occupy a similar temperature range and whose study pre-dates exoplanet science, however, they do differ in surface gravity and do not experience the extreme insolation of hot Jupiters.

Initially it was believed that clouds were unlikely to form in exoplanets with high equilibrium temperatures as they were thought to be too hot for condensation to occur. However, the detections of clouds across a wide range of temperatures has motivated studies into species that condense at higher temperatures. Indeed, Wakeford et al. (2017b) produced condensation curves for aluminium and titanium bearing molecules that could explain the presence of clouds in the hottest exoplanet atmospheres.

Unlike clouds, hazes are photochemically produced in the upper atmosphere and are predicted to be more important at lower temperatures of $\lesssim 1000\,K$ (Liang et al., 2004; Zahnle et al., 2009a; Morley et al., 2013; Fortney et al., 2013). The reason for this is that at temperatures above this threshold, the atmosphere of an exoplanet is driven towards chemical equilibrium such that CO becomes the dominant carbon-bearing species over CH_4 (e.g. Fortney et al., 2013; Morley et al., 2015 and as shown
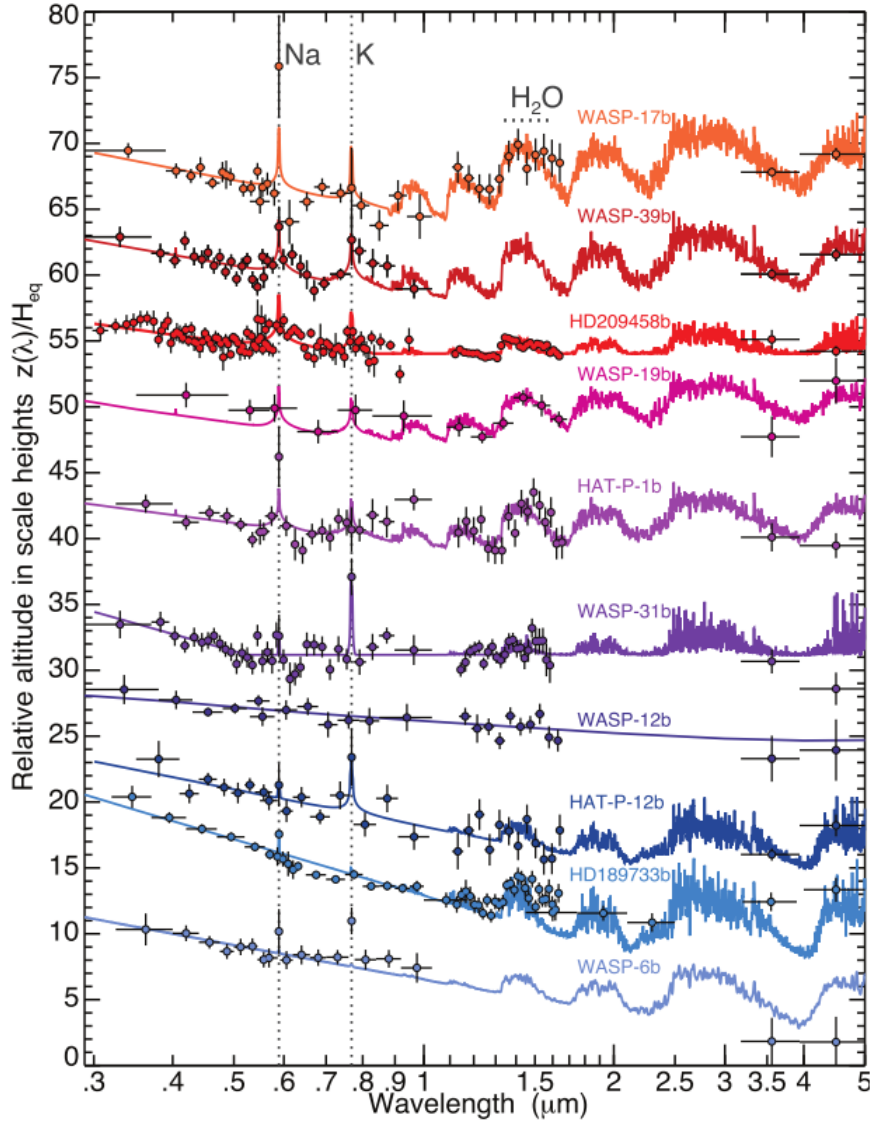
Figure 1.22: Results of the HST and Spitzer transmission spectroscopy survey of hot Jupiters (Sing et al., 2016). This survey found a continuum from clear to cloudy atmospheres (top to bottom on figure).
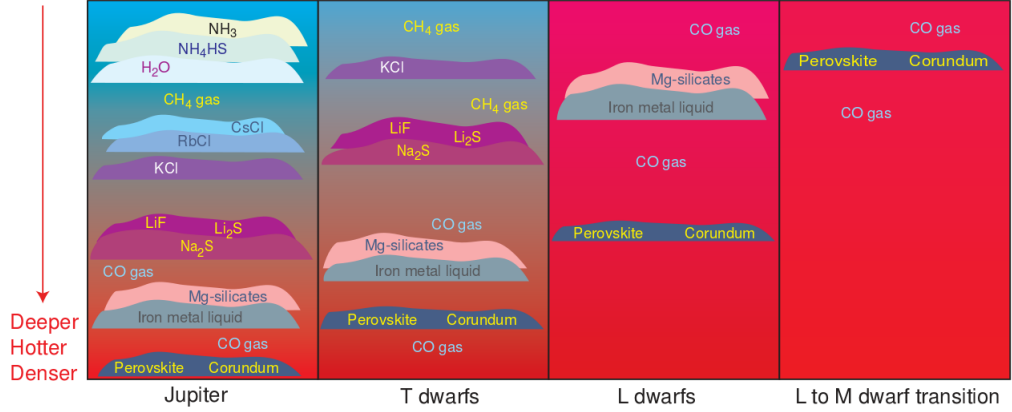
Figure 1.23: Diagram of cloud layers from Jupiter to brown dwarfs. As the atmospheres get hotter towards the right-hand panels, the upper cloud layers are stripped off. Figure from Lodders (2004).

in Fig. 1.23 and Fig. 1.24). This is similar to what happens at the L-T transition in brown dwarf atmospheres (Fig. 1.23). The photolysis of CH_4 forms hydrocarbons, which are precursors to haze and soot, that are optically thick and obscure spectral features within an exoplanet's atmosphere (Morley et al., 2013; Fortney et al., 2013; Morley et al., 2015). At temperatures $\gtrsim 1000\,K$ the low abundance of CH_4 should reduce the impact of hazes. Indeed, this has been used to explain the differences between the transmission spectra of HD 189733b (cooler and hazy) and HD 209458b (hotter and less hazy) (Lavvas & Koskinen, 2017). While hazes formed from sulphur may exist at higher temperatures ($1200 < T < 2000\,K$, Zahnle et al., 2009b), these are not optically thick in the visible light regime and so are unlikely to be important when considering the optical transmission spectra in this thesis.

As a result of the prevalence of clouds and hazes in exoplanet atmospheres, a number of recent studies have looked into whether their presence correlates with planetary parameters such as equilibrium temperature and surface gravity.

Stevenson (2016) studied a sample of 14 exoplanets ranging from super-Earths to Jupiters that had published HST/WFC3 observations. This study used the amplitude of the $1.4\,\mu m$ water feature as a proxy for cloudiness and found that planets with temperatures $> 700\,K$ and logg > 2.8 were more likely to be cloud free.

In more recent analyses of the results of Sing et al. (2016), Heng (2016) found tentative evidence for a correlation between temperature and clouds, with hotter planets more likely to be cloud free in this sample of hot Jupiters. In this case, Heng (2016) used the difference between the transit depth at the core and wing of the sodium and potassium lines to define the cloudiness of the planet. The
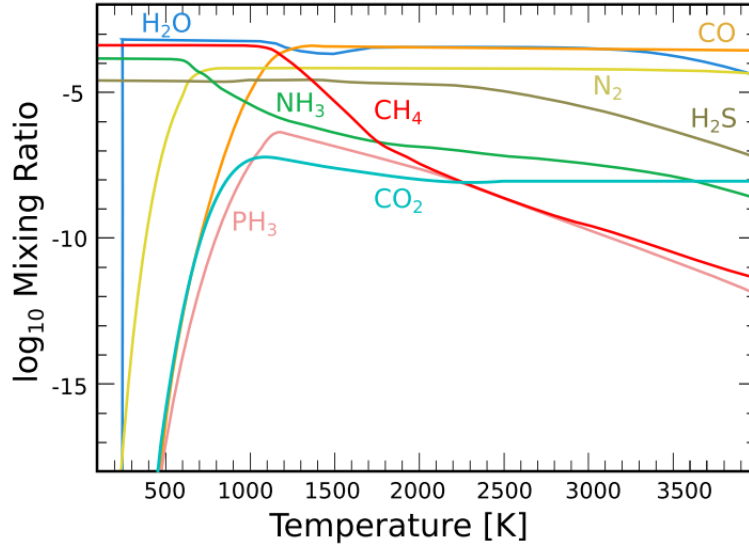
Figure 1.24: A plot of the molecular equilibrium abundances against temperature for a solar metallicity atmosphere at 1 bar. This demonstrates how CO becomes the dominant carbon-bearing molecule at temperatures $\gtrsim 1000\,\mathrm{K}$. Figure from Crossfield (2015).

author could not claim a significant correlation, however, due to the small sample size. Furthermore, in comparisons with the results of Stevenson (2016), Heng (2016) found some disagreement with the cloudiness of certain planets, leading the author to speculate whether the presence of clouds in the infrared did not necessarily imply clouds in the visible.

Crossfield & Kreidberg (2017) looked at a sample of 6 warm Neptunes ($2 < \mathrm{R}/\mathrm{R}_\oplus < 6$) with published HST/WFC3 observations and again used the amplitude of the water feature as a proxy for cloudiness. These authors also found tentative evidence for a temperature dependence, with planets $< 850\,\mathrm{K}$ more likely to be hazy. However, this was again based on a small sample.

Fu et al. (2017) updated the study of Stevenson (2016) by using the 34 planet sample of Tsiaras et al. (2017) with HST/WFC3 transmission spectroscopy observations. Using the amplitude of the water feature as a proxy for cloudiness, these authors also saw a correlation with temperature from 500-2500 K, after binning their sample in intervals of 150K (Fig. 1.25). However, these authors suggest that their correlation needs to be tested as their sample had few hot planets and a degeneracy between temperature and mass. As a result, they propose more hot, and preferably low mass, exoplanets should be targeted.

Such correlations can only be tested with a larger sample size which has been
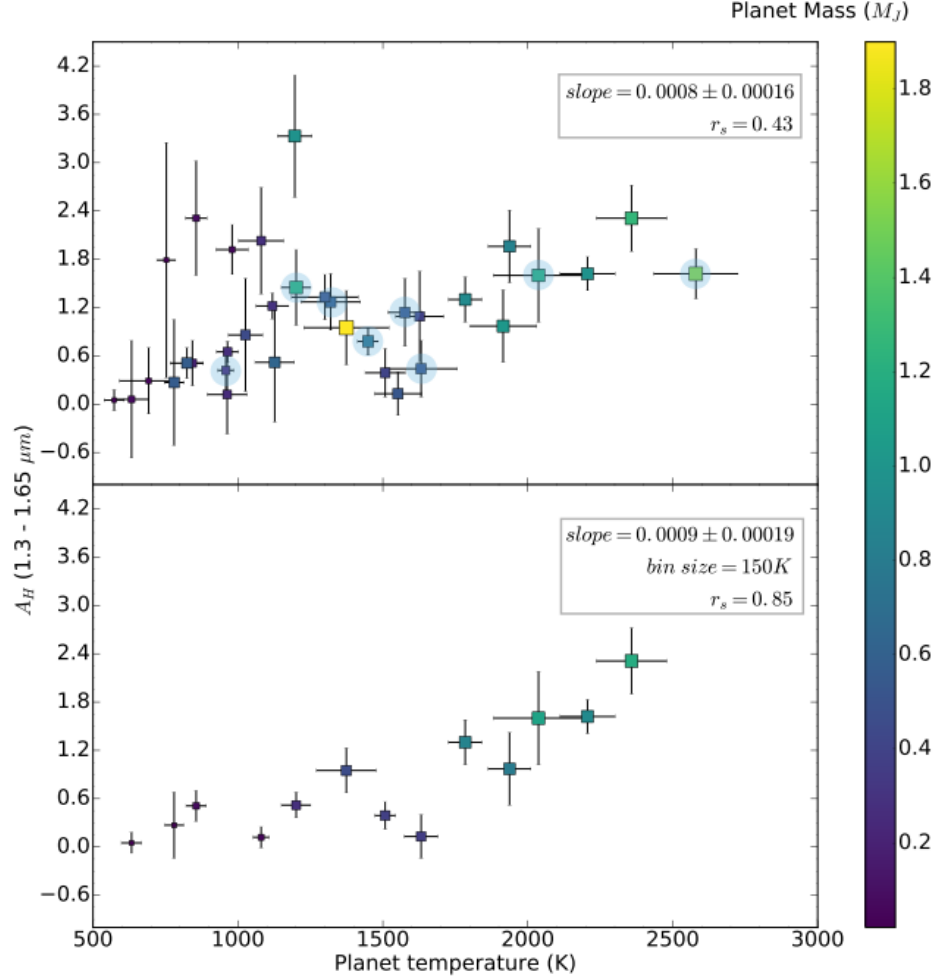
Figure 1.25: The correlation between temperature and water absorption observed in 34 exoplanets with HST/WFC3 measurements, expressed as $1.3 - 1.6\mu$m absorption in scale heights. Top panel: all 34 planets, showing a weak correlation between water amplitude and equilibrium temperature, with a Spearman correlation coefficient $r_S = 0.43$. The sample of Sing et al. (2016) is shown by the shaded blue circles. Bottom panel: the sample divided into bins 150K wide. In this case the correlation is significant with $r_S = 0.85$. The colours show the masses of the planets indicating the degeneracy between temperature and mass in the sample. Figure from Fu et al. (2017).

a primary motivation for this thesis and is a goal of the LRG-BEASTS survey, which is described in Chapter 2.

Stellar contamination

While the slopes seen in the optical transmission spectra of exoplanets are often interpreted as being caused by atmospheric opacity sources, it is also possible for activity on the host star to give rise to similar features. Star spots that are on the stellar surface but lie away from the transit chord (*unocculted* spots) make the star darker and redder which can lead to an overestimate of the planet's size, making the transit depth larger towards bluer wavelengths. McCullough et al. (2014) re-analysed the Rayleigh scattering slope of HD 189733b presented in Pont et al. (2013) and found that unocculted star spots provided a comparable fit to the data. Similarly, Oshagh et al. (2014) found that plages (bright regions) on the stellar surface could also reproduce the Rayleigh scattering slope of HD 189733b if they were occulted during the planet's transit. Occulted plages would cause dips in the transit light curve, getting larger towards bluer wavelengths, and cause a deeper transit to be observed than due to the planet's atmosphere alone. These studies have prompted subsequent transmission spectroscopy studies to perform a careful analysis of the possible influences of stellar activity on their results. The results presented in this thesis also take this into consideration and this analysis is presented in Chapters 4, 5 and 6.

1.5.4 Hot Jupiters

As noted in section 1.5.1, hot Jupiters have very large atmospheric scale heights. This makes them excellent targets for transmission spectroscopy and the majority of atmospheric detections have come from this class of planets. This is a very interesting population of planets, with many questions remaining regarding their inflated radii (section 1.3.3), formation (section 1.4) and atmospheres (section 1.5). These questions demand further study of hot Jupiters and this has been a primary motivation for the work presented in this thesis and is the driver for the LRG-BEASTS survey. For this reason, I devote the next chapter, Chapter 2, to my survey aimed at characterising a large number of such planets.

Chapter 2

The LRG-BEASTS survey of hot Jupiter atmospheres

The Low Resolution Ground-Based Exoplanet Atmosphere Survey using Transmission Spectroscopy (LRG-BEASTS), has become the focal point of my PhD research. The primary motivation behind this survey was to obtain ground-based transmission spectra of a sample of hot Jupiters to allow for statistical studies. Additionally, however, we also sought to test the capabilities of 4-metre class telescopes, which have typically been underused for such observations and could, in principle, allow for more telescope time to be brought to bare to expand the sample more quickly. LRG-BEASTS began in 2016 and has since been awarded a combined 26 nights on the WHT and NTT, covering a sample of 15 hot Jupiters (Fig. 2.1). A table listing all these planets along with their associated parameters is given in Chapter 7.

As can be seen in Fig. 2.1, this sample covers a large region of parameter space and will allow us to test the correlations between the presence of clouds and hazes and fundamental planetary parameters. As discussed in Chapter 1, there is evidence that hotter planets are more likely to be cloud-free than cooler planets. However, these studies have been based on small sample sizes and it is necessary that we increase the sample of characterised exoplanets to test such correlations. LRG-BEASTS, together with HST and other ground-based surveys such as the Gran Telescopio Canarias (GTC) exoplanet transit spectroscopy survey (e.g. Parviainen et al., 2016; Chen et al., 2017), ACCESS (Rackham et al., 2017), the Gemini/GMOS Transmission Spectral Survey (Huitson et al., 2017) and the VLT/FORS2 survey (e.g. Nikolov et al., 2016; Gibson et al., 2017) will expand the sample of characterised exoplanets.

To date, LRG-BEASTS has revealed a Rayleigh scattering haze in the at-
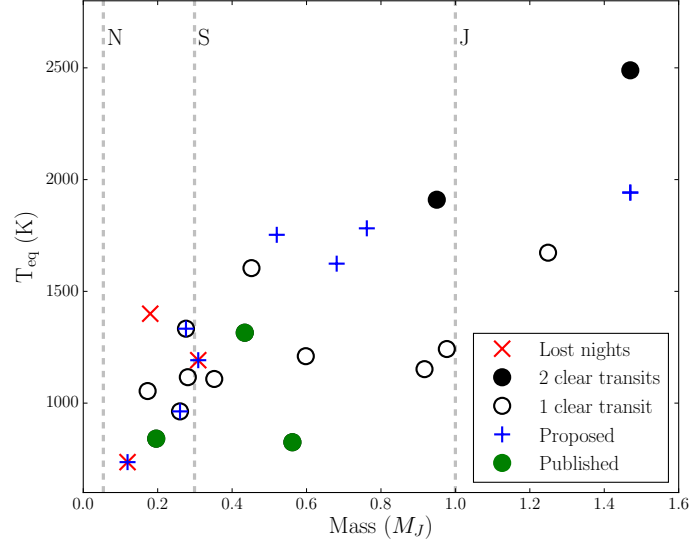
Figure 2.1: The full LRG-BEASTS sample including targets lost to poor weather and proposed targets under review. The masses of Neptune (N), Saturn (S) and Jupiter (J) are shown for reference and indicate the broad parameter space we're probing. As of 1st April 2018, the sample consists of 15 planets with at least one transit observed.

mosphere of HAT-P-18b (Chapter 5 and Kirk et al., 2017); detected clouds in the atmosphere of WASP-52b (Louden et al., 2017), which was consistent with my earlier study using ULTRACAM (Chapter 4 and Kirk et al., 2016); and ruled out a previously claimed detection of pressure-broadened potassium in the atmosphere of WASP-80b, instead finding an atmosphere most consistent with haze (Chapter 6 and Kirk et al., 2018).

2.1 Target selection

To select my targets, I consider a number of different factors. Firstly I base my target selection on a figure of merit that ranks all known exoplanets by the contribution of one atmospheric scale height to the transit depth. The signal is dependent on the scale height of the exoplanet and the relative size of the planet to its host. I calculate the signal as (Bento et al., 2014)

$$\text{Transit signal} = \frac{2R_P k T_{eq}}{\mu g R_*^2} \tag{2.1}$$

where R_P is the radius of the planet, k is the Boltzmann constant, T_{eq} is the planet's equilibrium temperature, μ is the mean molecular weight, which is
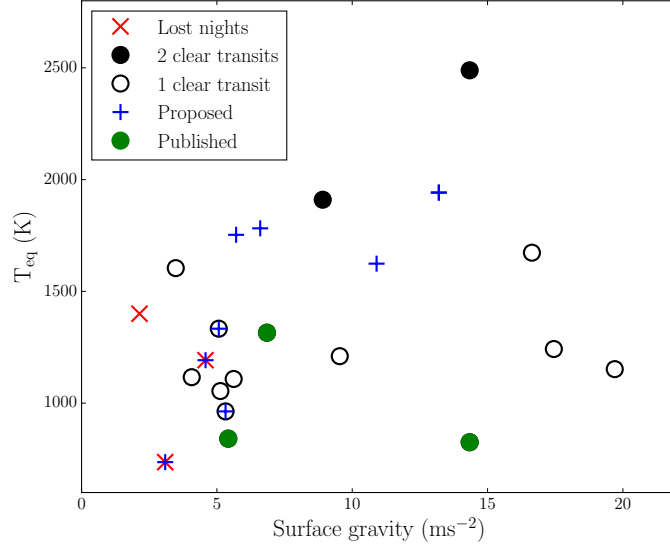
Figure 2.2: The same sample as Fig. 2.1 but this time with the surface gravity on the x axis.

assumed to be the same as for Jupiter (2.3 times the mass of a proton), g is the planet's surface gravity and R_* is the radius of the host star. While my assumption that all planets have the same mean molecular weight is not valid (my catalogue contains Earth-sized planets which will likely have higher mean molecular weights), this is not the dominant factor in equation 2.1. This equation indicates that hot planets with large radii, low surface gravities and small host stars give the highest predicted transmission signals. Fig. 2.2 demonstrates how our sample has primarily focussed on exoplanets with low surface gravities, as they are expected to have large absorption signals.

I maintain a spreadsheet of all known exoplanets ranked by this method, which I periodically update. For all planetary parameters, I use the very well-maintained TEPCAT catalogue[1] (Southworth, 2011). Table 2.1 shows the current top 10 best planets ranked by this signal[2]. In Fig. 2.3, I plot all exoplanets with absorption signals ≥ 100 ppm against their host stars' V magnitudes.

Of course, this is not the only consideration when selecting targets for observation. I also require targets to be bright. In general this means the host star should have a V magnitude $\lesssim 12$, although if the predicted signal is exceptionally large I will consider fainter targets. The largest limitation is the availability of suitable comparison objects. For ground-based studies, I correct for telluric effects

---

[1]http://www.astro.keele.ac.uk/jkt/tepcat/
[2]valid as of 08/02/2018

Table 2.1: The 10 exoplanets with the largest predicted transmission signals as of 08/02/2018. This ranking is based on the contribution of 1 atmospheric scale height to the transit depth and does not take into account the suitability of the targets for observation.

Planet name	Predicted signal (ppm)
WASP-107b	530
WASP-127b	482
WASP-39b	452
WASP-17b	445
GJ 1214b	407
WASP-52b	405
HATS-43b	397
HAT-P-32b	376
TrES-4b	363
WASP-31b	355
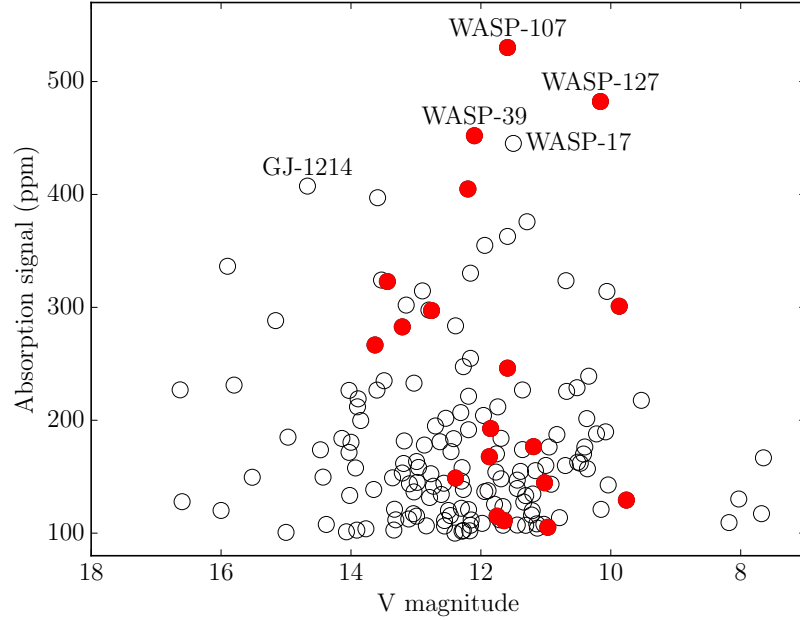
Figure 2.3: A plot of all exoplanets with predicted absorption signals above 100 ppm against the host stars' V magnitudes. All of the LRG-BEASTS planets from Fig. 2.1 are shown by the red points. The names of the 5 planets with the largest absorption signals are included.

by dividing the target's light curve by a comparison star's light curve. However, the comparison star should be similar in colour (ideally $\Delta(B - V) < 0.5$) and in magnitude ($\Delta m < 1$) although I do sometimes have to settle for less than ideal comparisons if a target is particularly attractive for transmission spectroscopy. It is not unusual for no comparison to be available at all within the instrument's field of view and in this case I have to remove such targets from consideration.

In order to search for comparison stars, I wrote a PYTHON script that takes as command-line inputs, the target's name and the chosen instrument (ACAM or EFOSC). The script then queries the UCAC4 catalogue (Zacharias et al., 2012) with the target name and performs a cone search based on the slit length of the chosen instrument. This provides the user with a list of possible options from which the user is prompted for a choice. With the choice, the script calculates the position angle needed to rotate the two stars such that they lie along the same CCD row and opens up a DS9 window with the target and comparison plotted with a accurate representation of the slit overlaid. An example is shown below for HAT-P-18b.

```
% python ~/python/obs_tools/ds9_finder_tool.py hat-p-18 -inst ACAM -s 40


  _r      _RAJ2000    _DEJ2000    Vmag    Bmag     pmRA       pmDE
 arcm       deg         deg       mag     mag    mas / yr  mas / yr
------ ---------- ---------- ------ ------ -------- --------
0.0004* 256.346461  33.012481 12.655 13.662    -16.3     -33.4
3.4487 256.285489  33.038756 11.247 12.542     -5.4      -3.6
4.4775 256.272199  32.971390 13.863 14.712      4.8     -16.7
4.5737 256.437274  33.009001 13.145 13.740      6.4      5.9
5.3203 256.262897  32.958181 13.019 13.870    -23.5      28.5
6.0168 256.226918  33.015166 13.955 14.620    -16.4      -3.0
6.0183 256.227347  33.021693 13.105 13.920    -18.3     -25.6

*this is the target
===
Which comparison to use?     1


===
Target coords =  256.346461 33.012481
Target coords =  17h05m23.1506s +33d00m44.9316s
Target Vmag = 12.65
Target B-V = 1.01
```
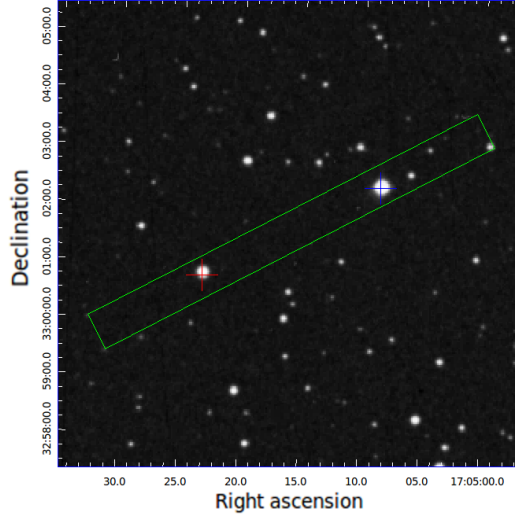
Figure 2.4: Finding chart generated using my custom PYTHON script. In this example, HAT-P-18 is shown by the red cross-hair and the comparison is in blue. A to-scale depiction of the 40 arcsec slit is overlaid in green.

```
=
Comparison coords =  256.285489 33.038756
Comparison coords =  17h05m08.5174s +33d02m19.5216s
Comparison Vmag = 11.25
Comparison B-V = 1.30
=
Position angle = 297.219
Separation = 3.449 arcmin
Midpoint coords =  17h05m15.834s +33d01m32.2266s
===
```

In this case, `-s 40` means we're using the 40 arcsec slit rather than the 27 arcsec slit. The finding chart created by running the above script is shown in Fig. 2.4.

Finally, I also consider previously published measurements of each prospective target. Targets that have not been studied before or that have been found to show unusual features will be included while targets that have been studied extensively (e.g. HD 189733b) will be excluded. However, repeat observations are important as transmission spectroscopy is susceptible to systematics and stellar activity can be misinterpreted as an atmospheric signal (see Chapter 1). Indeed, my

51

study of WASP-80b (Chapter 6 and Kirk et al., 2018) highlighted the importance of repeat observations.

With the above considerations I then use the NASA Exoplanet Archive's Transit Ephemeris Service[3] to search for all transits of interesting targets in a particular observing season that are observable from a chosen observatory. As a cross-check I compare the output with a similar search performed with the Exoplanet Transit Database[4]. I then subsequently cut the resulting transits, by enforcing that transit ingress, mid-point and egress all occur at an airmass $\leq$ 1.7, and that a full transit with 1-2 hours of out-of-transit baseline is observable between evening and morning twilight. I further make sure that any resulting transits are far away from the moon (> 30 degrees) if the night is bright.

If this results in few potential transits, I will relax the constraints on airmass (but never observe at airmass > 2) and also allow the transit to occur nearer to twilight. These nights then become backup nights.

The first results of LRG-BEASTS are presented in Chapters 5 and 6.

---

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

[4]http://var2.astro.cz/ETD/

Chapter 3

Methods and Instrumentation

In this chapter, I outline the methods, techniques and instrumentation used to obtain and analyse the data presented in Chapters 4, 5 and 6. I will begin by briefly introducing CCDs and discuss their calibration in section 3.1. I introduce the technique of spectroscopy in section 3.2 and then move onto a description of the instruments used in sections 3.3 and 3.4. Finally, I discuss the tools I have developed to reduce and analyse the data obtained in sections 3.5, 3.6 and 3.7, respectively.

3.1   CCDs

Charge-coupled devices (CCDs) have been the main light sensors in astronomy since the 1980s. CCDs are silicon semi-conductors that operate using the photoelectric effect, with incident photons liberating electrons which are subsequently recorded. The electrons are stored in pixels which are shifted row-by-row to be read out and an image is reconstructed. CCDs can be *back-illuminated*, meaning that the photons are incident on the rear silicon face of the CCD rather than onto the face where the pixel gate structures are located. This is done to maximise the sensitivity of the CCD as blue photons are absorbed by the gate structures. However, back-illuminated CCDs can suffer from *fringing*, which is caused by interference between incoming light and light which has been internally reflected off the back surface of the CCD. A thicker and/or higher optical resistance silicon, such as in deep-depleted CCDs, can be used to trap incoming photons before they reflect and interfere, hence reducing the effects of fringing.

Once the electrons have been released into the pixel's potential well, and the integration has finished, the electrons are shifted row-by-row and read out. During the readout process, the voltage measured by each pixel is amplified and the signal

digitised, typically in ADU (Analogue-to-Digital Unit). The conversion factor from electrons to ADU is called the gain, which is given in units of e^- ADU^{-1}. Therefore the values in FITS images are not the number of detected photons but instead are given in ADU. This is important when I come to calculate the error in my extracted photometry (section 3.6).

CCD readout time is approximately proportional to the number of pixels read out. In order to reduce the readout time, and increase the efficiency of the observing duty cycle, we can window and bin CCDs to reduce the number of pixels read out. The readout speed is also configurable but a faster readout leads to higher noise.

A more in-depth review of CCDs and their use in astronomy is presented by Howell (2006).

3.1.1 Bias frames, dark frames and flat fields

When reducing CCD data, we typically need to correct for the baseline offset of the CCD (a bias frame), correct for thermally generated electrons (a dark frame), and correct for pixel-to-pixel sensitivity variations in the CCD (a flat field).

CCDs have an offset applied to them in the form of a bias voltage. If there was no offset applied to the CCD, then pixels with few photo-electrons could return negative counts due to the readout noise associated with the pixel, and negative counts cannot be processed by an analogue to digital converter. This problem is overcome with the bias voltage so that even if there are few photo-electrons, the pixel will contain a positive count. A bias frame is a zero second dark frame, taken without illuminating the CCD, and is used to correct for this offset. In this frame, each pixel will have a value which follows a Gaussian distribution centred on the mean bias level ADU. Structure in the bias can be dependent on the windowing. It is therefore important to take bias frames with the same windowing setup, and readout speed, as the science images that they will be used to correct. However, this variation is constant between images. The biases are affected by read noise and so many biases are median combined to create a master bias. This master bias is then subtracted from all other data images to remove the base count level of the CCD. Other images can also be used to infer the mean bias level that are themselves not bias frames. For standard windows, there is often an overscan region at the edge of the CCD which is not illuminated and records the mean offset, although it does not capture the pixel-to-pixel offset variations. However, in the case of my science observations, I often use a smaller than standard window to reduce the readout time and therefore do not always record the overscan region.

A dark frame is similar to a bias but with a non-zero integration time. It is an image with the shutter closed which is used to record thermal signal (dark current) within the CCD. This signal is a result of thermally generated electrons that cause charge to be recorded in the CCD. Dark current is not usually a significant source of noise for cryogenically-cooled modern detectors, and typically contributes only a few electrons per pixel per hour. Dark frames include the bias, and so it is not necessary to take separate bias frames if darks have been taken so long as the darks have the same exposure time as the science frames and flat fields they are being used to correct. If not, the bias must be subtracted from the dark frames to leave the thermal image. This thermal image can then be scaled to the appropriate exposure time to correct the science images and flat fields.

Flat fields are used to remove the pixel-to-pixel sensitivity variations of the detector and to remove the results of imperfections in the optical path, such as dust. Due to the second point, is important to take flat fields with the same filter (in the case of photometry) and slit (in the case of spectroscopy) as the science images they will be used to correct. When taking a flat field, we would like to illuminate the detector evenly, although in practice this can be difficult as we often struggle for blue photons. This is especially true in the case of spectroscopy, which is discussed in more detail in section 3.3.2. To do this, images are taken of a light source, typically of light reflected from the inside of the telescope dome (dome flats), light from an internal lamp (lamp flats) or of the sky at twilight (sky flats). Flat fields are also used to correct for vignetting at the edge of the telescope field of view.

In the case of spectroscopic flats, the spectrum of the source used to create the flat field needs to be removed. Sources used to create lamp flats usually have featureless blackbody spectra which can be removed with a spline or polynomial but which lack blue light. Sky flats, however, contain emission from the sky which cannot be as easily removed. I discuss both these processes in more detail in section 3.6.

3.2   Spectroscopy

Spectroscopy is a fundamental tool for astronomy. By dispersing incident light we can measure the intensity of the light as a function of wavelength. The resolving power of a spectrograph R is given by

$$R = \frac{\lambda}{\Delta\lambda} \tag{3.1}$$

where $\Delta\lambda$ is the limiting resolution, that is the difference in wavelength
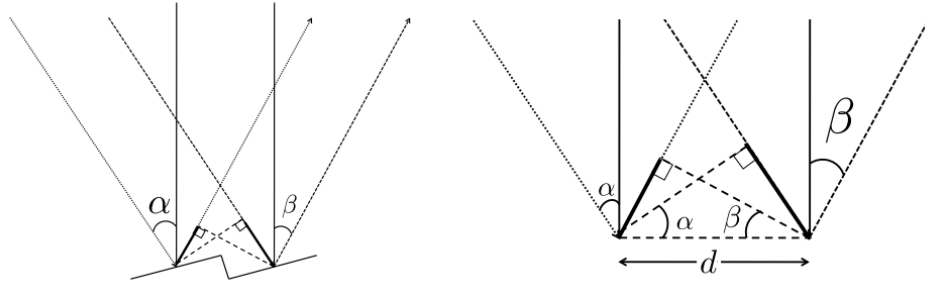
Figure 3.1: Schematic of a diffraction grating, with the right plot showing a zoom in on the left plot. Two light beams are incident at an angle α to the grating normal. The light is then diffracted at an angle β from the grating normal. The grooves are separated by a distance d. The bold black lines in the right-hand zoom indicate the extra path distance travelled by the second ray of light, whose lengths are given by $d \sin \alpha$ and $d \sin \beta$, respectively. Figure from Huitson (2013).

between two spectral lines of the same intensity that can be distinguished.

Maxima in the dispersed light occur when the path difference between two diffracted beams of light is equal to an integer number of wavelengths such that the interference is constructive. These maxima are called the *orders* of the diffraction pattern, and run from zero to n.

The setup of a diffraction grating is shown in Fig. 3.1. This figure can be used to yield the difference in path lengths travelled by two beams of light incident on a diffraction grating to give the grating equation

$$m\lambda = d(\sin \alpha + \sin \beta) \tag{3.2}$$

where m is the diffraction order, λ is the wavelength of incident light, d is the groove spacing, α is the angle of incident light to the grating normal and β is the angle of diffracted light to the grating normal, as described by Palmer & Loewen (2005).

The resolving power of a diffraction grating is related to the order m of the diffracted light and the total number of grooves illuminated on the surface of the grating, N by

$$R = mN \tag{3.3}$$

Figure 3.1 actually depicts a *blazed* diffraction grating whereby incident light is diffracted by grooves in the grating instead of passing through slits. The angle of these grooves is called the *blaze angle*, θ_B in Fig. 3.2. This is related to the
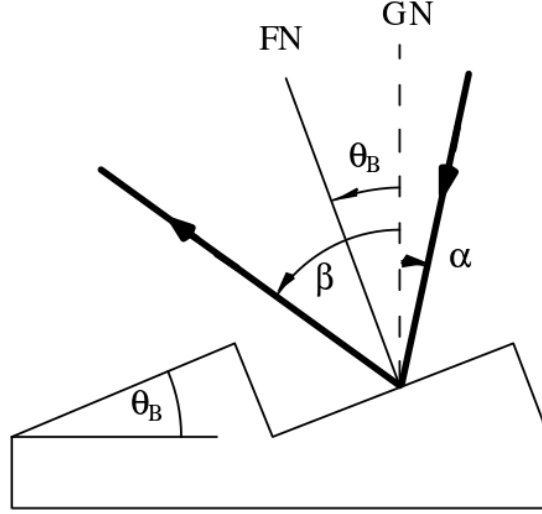
Figure 3.2: Figure indicating the blaze angle, θ_B of the grating. The angles of incidence and diffraction are given by α and β respectively. GN is the grating normal and FN is the facet normal. Figure from Palmer & Loewen (2005).

blaze wavelength λ_B (the wavelength of the maximum intensity for a given order m) through

$$\lambda_B = \frac{2d}{m} \sin \theta_B \tag{3.4}$$

The spectrographs used in this thesis employ a *grism*, which is a diffraction grating on the hypotenuse side of a right-angled triangular prism. The prism acts to refract the light back into line for a central wavelength. This is useful as the same camera can be used for spectroscopy and imaging by removing the grism without having to move the detector.

The science results presented in Chapters 5 and 6 were based on long slit spectroscopic data. Long slit spectroscopy is performed with a single, long slit that is placed in the optical path, before the light is dispersed. The 27 arcsec slit used in Chapter 5 has a length of 6.8 arcmin (with a physical length of 90.5 mm) and the 40 arcsec slit used in Chapter 6 has a length of 7.6 arcmin (with a physical length of 101.1 mm). Long slit spectroscopy lends itself well to ground-based transmission spectroscopy since it allows for a comparison star to be observed simultaneously with the target. This is necessary to correct for changes due to observing conditions and the effects of observing through Earth's atmosphere (telluric effects). It is also possible to request custom made slits which are much wider (tens of arcsec) than

standard (a few arcsec), thereby avoiding differential slit losses between the target and comparison. The disadvantages of wide slits, however, is that the spectral resolution varies with the seeing and the wavelength calibration varies with guiding errors.

The grism spectrographs used in this thesis provide low resolution spectra ($R \approx 400$) across the optical regime. We chose to use these spectrographs as they are relatively simple in their optical design and they can observe a broad range of wavelengths simultaneously, which is particularly important when looking for Rayleigh scattering signatures in exoplanetary atmospheres (as described in Chapter 1).

While I employ long-slit spectroscopy with grism spectrographs in my research, other spectroscopy techniques and instruments exist.

Echelle spectrographs have very high blaze angles and often very high diffraction orders are observed to increase the resolution of the spectrum. The incredibly successful planet hunting HARPS (Mayor et al., 2003) spectrograph on the European Southern Observatory's 3.6 metre telescope is an echelle spectrograph, with a resolution of 115,000.

Multi-object spectroscopy (MOS) allows the observer to place slits at the locations of stars on the CCD without being restricted by a single slit. In this way MOS can extract spectra for many objects within a field of view, allowing for a greater number of comparison stars. MOS has been used on several occasions to study exoplanet atmospheres, such as the recent observation of clouds in the atmosphere of WASP-4b using the Gemini Multi-Object Spectrometers (Huitson et al., 2017).

Integral field units (IFUs) feed each pixel into a spectrograph so that each pixel has an associated spectrum. This results in a datacube so that the 2D image has a third dimension corresponding to the individual pixels' spectra. Recently, Hoeijmakers et al. (2018) detected CO and H_2O in the atmosphere of β Pictoris b using the SINFONI integral field spectrograph on the VLT.

3.3   The Willam Herschel Telescope (WHT)

The William Herschel Telescope is a $4.2\,\mathrm{m}$ Alt-Az telescope at the Roque de los Muchachos Observatory, La Palma, Spain. Its first light was in 1987 and it is currently the largest optical telescope in the Northern Hemisphere for which UK astronomers have direct access.

The WHT currently has three instruments, a near-infrared imager and spec-

trograph (LIRIS, Long-slit Intermediate Resolution Infrared Spectrograph), a medium resolution single-slit spectrograph (ISIS), and ACAM, an optical imager and low resolution spectrograph. Previously, ULTRACAM was a visitor instrument on the WHT but it is now at the New Technology Telescope, which is discussed in section 3.4.

The observations and results in Chapter 4 were made using ULTRACAM on the WHT, which is discussed in more detail in section 3.3.1.

The observations and results presented in Chapters 5 and 6 were obtained with ACAM, which is discussed in more detail in section 3.3.2. ACAM will be largely displaced by the installation of WEAVE on the WHT, which is expected to begin commissioning in 2018. WEAVE is a multi-object spectrograph designed for large surveys. Its fibre acquisition unit will be mounted on the prime focus of the WHT with the spectrograph on one of the Nasmyth platforms. Its wide, 2 degree, field of view will allow it to take spectra of 1000 stars at a time and one of its main science goals is to provide radial velocity follow-up to the GAIA satellite.

On average, 25% of nights a year on the WHT are lost to bad weather, although the number of nights lost per month varies considerably with less than 5% of time lost in the summer and over 35% of time lost in winter (Fig. 3.3). Since exoplanet transits are time-critical observations, and as transmission spectroscopy requires precision on the order of 100 ppm, the loss of time and effects of weather can compromise my science objectives.

3.3.1 ULTRACAM

ULTRACAM (Dhillon et al., 2007) is a high-speed CCD photometer. It uses two dichroics allowing one to simultaneously observe in three colours (Fig. 3.4). ULTRA-CAM is able to reach frame rates of up to 500 Hz due to its use of frame-transfer CCDs which provide a storage area for charge to be digitised while the next exposure is integrating. It was primarily designed for rapidly varying variable stars, including eclipsing white dwarfs. Its fast readout enables efficient photometry of bright stars on a big telescope and its ability to observe at three wavelengths simultaneously make it a useful tool for transmission spectroscopy. Stellar activity can masquerade as absorption from an exoplanet's atmosphere, as discussed in Chapter 1, which can be alleviated by observing at multiple wavelengths simultaneously. ULTRACAM has a field of view of 5 arcmin, which allows for suitable comparison stars to be simultaneously observed, and has a plate scale of 0.3 arcsec pixel^{-1}.

In Chapter 4, I present observations of WASP-52 taken with ULTRACAM. For these observations, ULTRACAM was mounted at the Cassegrain focus of the
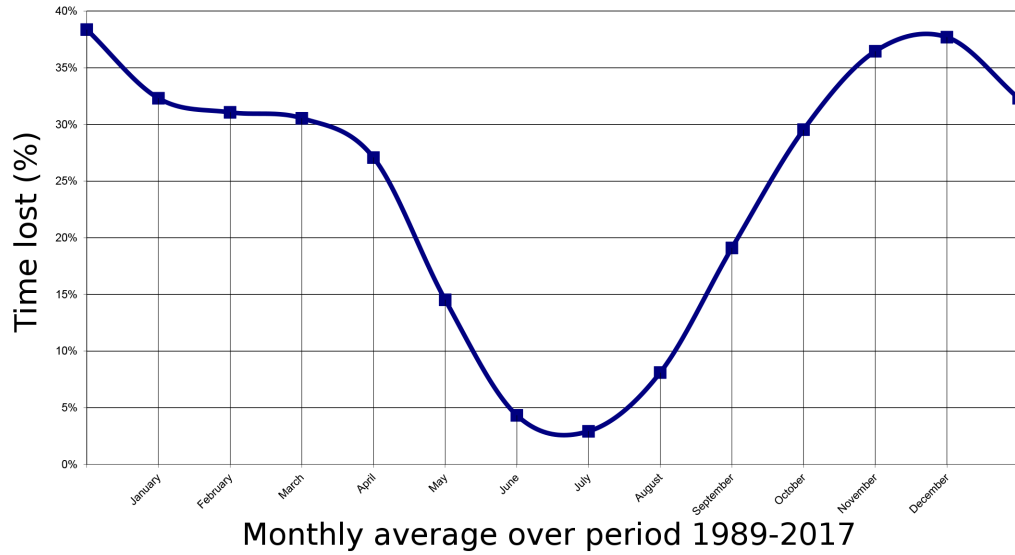
Figure 3.3: Average fraction of observing time lost per month at the WHT since 1989. Figure adapted from http://www.ing.iac.es, accessed 15/11/2017.
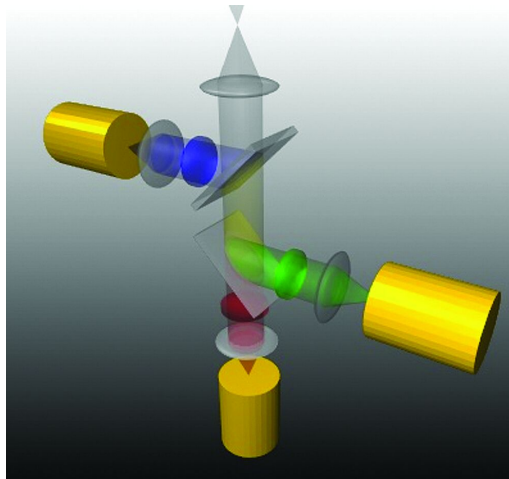
Figure 3.4: A schematic of ULTRACAM showing the two dichroics used to split incoming light into 3 wavelengths that are re-imaged onto 3 CCDs. Figure from http://slittlefair.staff.shef.ac.uk/teaching/phy241/lectures/L07/index.html, accessed 15/11/2017.
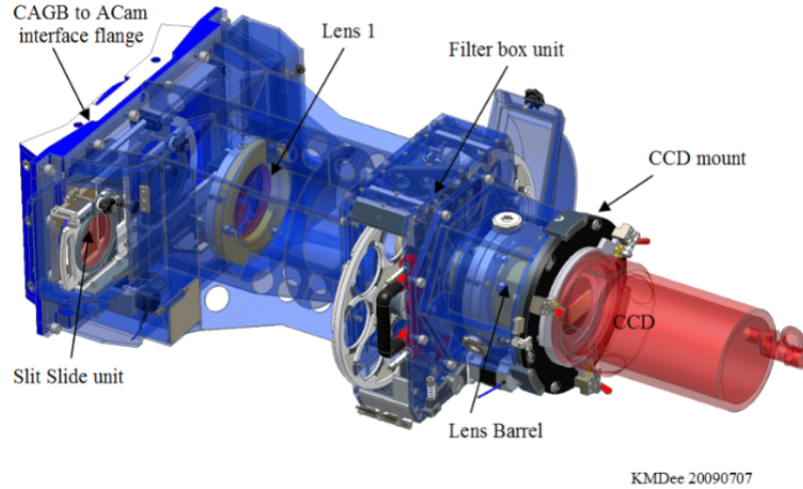
Figure 3.5: A schematic of the ACAM instrument. Light is indicident from the left in this schematic before passing through the slit wheel, filter wheels and being imaged onto the CCD. Figure from http://www.ing.iac.es, accessed 08/12/2017.

WHT as a visitor instrument. As ULTRACAM is a photometer, not a spectrograph, I will refer to these observations as transmission *photometry*. Transmission photometry has been used a number of times in the literature, including with ULTRACAM (e.g. Copperwheat et al., 2013; Bento et al., 2014) and GROND, a 7-band photometer (e.g. Southworth et al., 2012; Mancini et al., 2013).

ULTRACAM data reduction is performed using a code written by Tom Marsh[1], which is introduced by Dhillon et al. (2007). This performs aperture photometry and relies on an input file which is discussed in more detail in Chapter 4.

3.3.2 ACAM

ACAM is mounted at the folded-Cassegrain focus of the WHT (until the installation of WEAVE) and can be used both for broad- and narrow-band imaging and for low-resolution spectroscopy using a Volume Phase Holographic disperser. In imaging mode, ACAM has a field of view of 8 arcmin but for spectroscopy this is limited by the length of the slit used, which is 6.8 arcmin for the 27 arcsec slit and 7.6 arcmin for the 40 arcsec slit. It has a resolution $R \approx 400$, which I measured empirically, and has a pixel scale of 0.25 arcsec pixel^{-1}, with a dispersion of ~ 3.3 Å pixel^{-1}. A schematic of the instrument is shown in Fig. 3.5.

ACAM can be windowed in both x and y, with the ability to use multiple

---

[1]http://deneb.astro.warwick.ac.uk/phsaap/software/ultracam/html/index.html

61

Table 3.1: Measured parameters of the ACAM instrument on the WHT and the EFOSC2 instrument on the NTT. All values are from the relevant instrument-specific web pages and correspond to fast readout mode.

	WHT/ACAM	NTT/EFOSC2
Gain (e^- ADU^{-1})	1.86	1.38
Readnoise (e^-)	6.5	12.6
Dark current (e^- per pixel per hour)	3	7
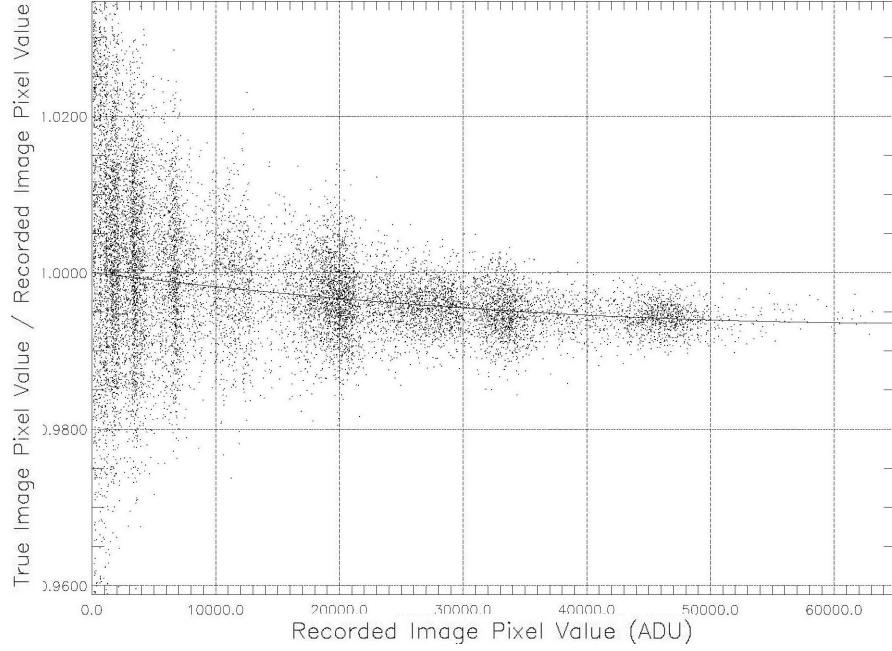
Figure 3.6: A plot of ACAM's linearity. Figure from http://www.ing.iac.es, accessed 11/12/2017.

windows, which reduces the readout time. My observations always use fast readout, and the combination of this with windowing leads to readout times of $< 10\,$s. Table 3.1 gives the gain, readout noise and dark current for ACAM as listed on the instrument's web pages[2]. In fast readout mode, the camera's response to incoming photons is linear to within 0.7% over the range $0 - 65\,$k ADU (Fig. 3.6).

A plot of ACAM's throughput is shown in Fig. 3.7 and is given relative to ISIS's throughput in Table 3.2. ISIS covers a similar wavelength range to ACAM and is more sensitive to bluer wavelengths (Table 3.2). However, we chose not to use ISIS as it uses a dichroic to split incoming light into a red and blue arm, which could in principle introduce different systematics between the two arms. It is also

---

[2]http://www.ing.iac.es/Engineering/detectors/g3_ultra_auxcam.html

Table 3.2: Throughput of ACAM relative to ISIS at different wavelengths. Data from http://www.ing.iac.es, accessed 15/11/2017.

	3600Å	4300Å	5500Å	6500Å	8200Å
ACAM/ISIS	0.6	0.9	2.1	1.7	1.3
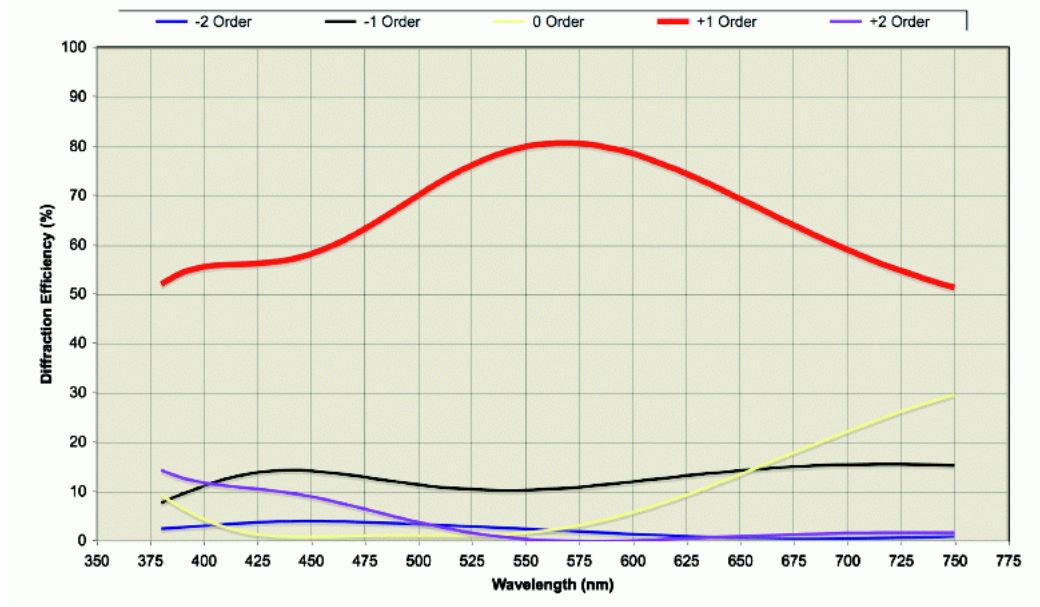
Figure 3.7: Diffraction efficiency of ACAM's Volume Phase Holographic (VPH) grating. Figure from http://www.ing.iac.es, accessed 08/12/2017.

less sensitive than ACAM at the wavelengths of the alkali absorption features of sodium and potassium. ACAM was preferred due to its simple design, reducing the opportunities for instrument related systematics, and also as it is the WHT's only permanently mounted instrument, which helps scheduling.

Between 3300 and 6600 Å, ACAM is uncontaminated by second-order light, with some contamination from blue second-order light between 6600 and 9500 Å. Such contamination can be removed by using an order blocking filter. However, the ACAM transmission spectra presented in Chapters 5 and 6 were taken without an order blocking filter. The motivation for observing without an order blocking filter was to reduce the number of elements within the optical path, thereby reducing the opportunities for systematic noise. Since the stars studied in this thesis are not blue objects and ACAM is not very sensitive to blue light, I do not expect this to have a significant effect. Nevertheless, it is the goal of future ACAM observations to observe two transits of the same exoplanet, with and without an order blocking filter, to quantify the effects this has on the transmission spectrum.
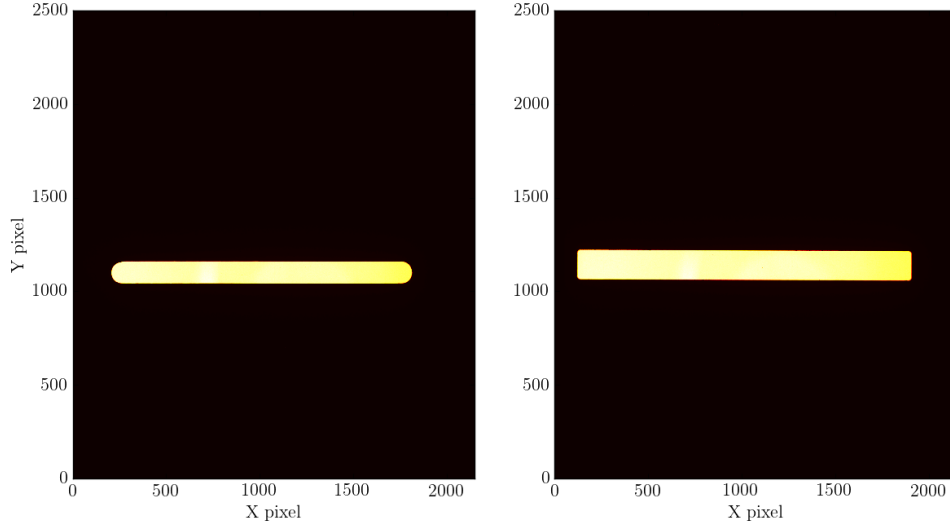
Figure 3.8: ACAM flat fields taken through the 27 arcsec slit (left panel) and 40 arcsec slit (right panel).

ACAM is well suited to transmission spectroscopy due to its wide wavelength coverage (3300-9500 Å) and the availability of very wide slits which avoid the effects of differential slit losses between the target and comparison stars. There are two very wide slits available for ACAM, a 27 arcsec slit which was kindly manufactured by the Isaac Newton Group (ING) in December 2013 for observations made by our team, and a 40 arcsec slit. The 40 arcsec slit was manufactured for observations of 55 Cancri e (by a different team) which, due to its very bright 5th magnitude host star, required a large defocus to avoid saturation. For my observations of HAT-P-18, presented in Chapter 5, I used the 27 arcsec slit. However, this became damaged soon after my observations and more recent observations have made use of the 40 arcsec slit. Fig. 3.8 shows flat fields taken with both the 27 and 40 arcsec slit in place.

3.4 The New Technology Telescope (NTT)

The New Technology Telescope (NTT) is a 3.58 m Atl-Az telescope located at ESO's La Silla Observatory, Chile. It was the prototype for the Very Large Telescope (VLT, Paranal), and the Telescopio Nazionale Galileo (TNG, La Palma) is a replica of this telescope. It currently has two permanently mounted instruments, a large field infrared spectrograph and photometer (Sofi) and the ESO Faint Object Spectrograph and Camera (EFOSC2).

64

I do not present results from this instrument here, so I will not go into detail about this instrument, other than to mention it is a low resolution grism spectrograph, similar to ACAM. For this reason, it has contributed, and will continue to contribute to, the LRG-BEASTS survey, which is described in Chapter 2. I also note that my custom-built long slit reduction pipeline, which is presented in section 3.6, has been written to reduce both ACAM and EFOSC2 data.

3.5   Manual guiding corrections

While observing, I monitor the positions of the target and comparison in real-time and in both the spatial and dispersion directions. This is done using a script that was created by John Rostron that fits Moffat profiles to the spectral traces in x and user-selected stellar absorption lines in y. From these fits, the locations of the target and comparison stars in x and y can be monitored at the sub-pixel level. This allows me to make manual guiding corrections to the telescope throughout the night to ensure the stars stay well-centred in the slit, thereby avoiding potential systematics. I have added some functionality to the code to allow for easier selection of absorption lines and also added plots of FWHM and maximum counts to the output.

Examples of such fits to a target's trace and stellar absorption line are shown in Fig. 3.9 and Fig. 3.10 respectively. Fig. 3.11 shows how the x and y positions change across the course of the observation. This figure also includes the manual guiding corrections which are made when the star is observed to drift by $\gtrsim 1$ pixel ($= 0.25$ arcsec).

3.6   Custom built long slit spectroscopy pipeline

While software exists to reduce long slit spectroscopy data, such as IRAF and Molly, I chose to write my own reduction pipeline in Python. The motivation behind writing my own pipeline was to get a better understanding of the data and the process of spectroscopy reduction. I also chose to keep the reduction as simple as possible to help identify and trace sources of systematics more easily.

I built my pipeline to rely on few user defined variables, each of which is easy to comprehend. This code was used to reduce the data presented in Chapters 5 and 6. As time progressed, the code became more sophisticated in its approach and therefore there were subtle differences between each implementation which will be discussed in the relevant science chapters.

My pipeline relies on normal aperture extraction, instead of optimal extrac-
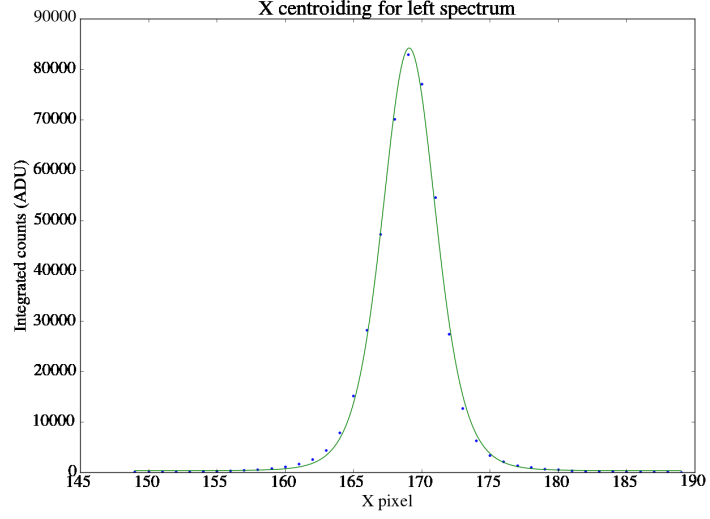
Figure 3.9: A Moffat fit to the spectral trace of the target in the spatial direction. The extracted spectrum is in blue and the fit is in green. The same procedure is performed for the comparison star.
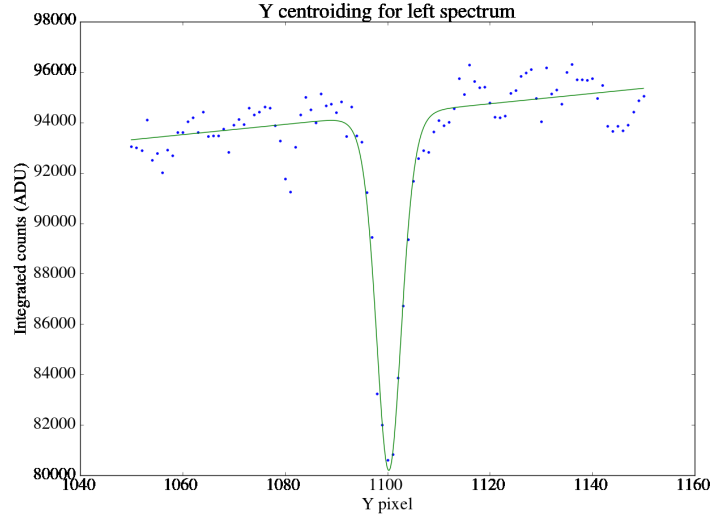
Figure 3.10: A Moffat fit to an absorption line in the target's spectrum in the dispersion direction. The extracted spectrum is in blue and the fit is in green. The same procedure is performed for the comparison star.
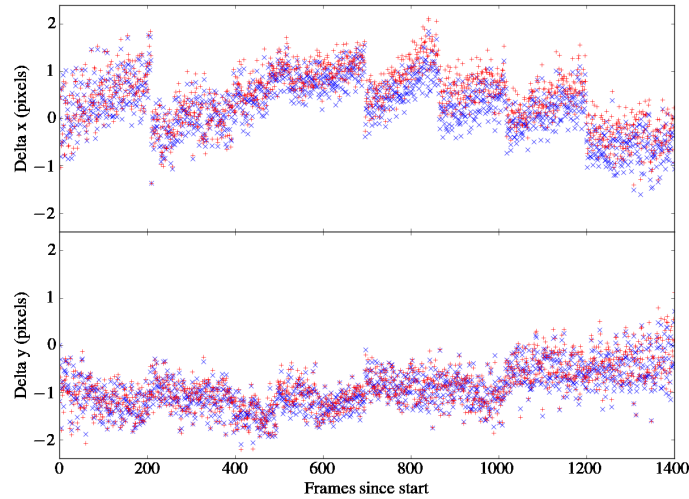
Figure 3.11: Top panel: the measured x locations of the target (blue pluses) and comparison (red crosses) from fitting Moffat profiles to the traces (Fig. 3.9). Bottom panel: the measured y locations of the target (blue pluses) and comparison (red crosses) from fitting Moffat profiles to stellar absorption lines (Fig. 3.10). In both panels the offsets are measured in pixels, where 1 pixel $= 0.25$ arcsec. This figure also shows the positions at which manual guiding corrections were made, seen as discrete steps.

tion which assigns weights to pixels before summing spectra in the spatial direction (Horne, 1986). Optimal extraction is particularly useful for low S/N observations but is not necessary for the bright targets I observe as part of this thesis. It is, however, an aim of mine to include optimal extraction in future versions of this pipeline to quantify the difference in quality of the extracted spectra.

In this section, I will begin by discussing how I de-bias and flat field the raw images, then move onto the extraction of the spectra and finish by describing how I calibrate the wavelength scale. The example reduction and figures in the subsequent sections correspond to ACAM data only. The reduction process for EFOSC2 data is identical but the flats and biases are different to those presented here.

3.6.1 Correcting the raw images

Master biases and master flats are relatively simple to create, as they just comprise of the median stacking of a number of images. However, obtaining a high S/N flat and removing the spectrum of the source used to generate it is more complicated.

I experimented with three different flat fields to correct the ACAM data presented in this thesis. These included lamp flats using a tungsten lamp (lamp flats), twilight spectroscopic sky flats without a slit (slitless sky flats), and twilight
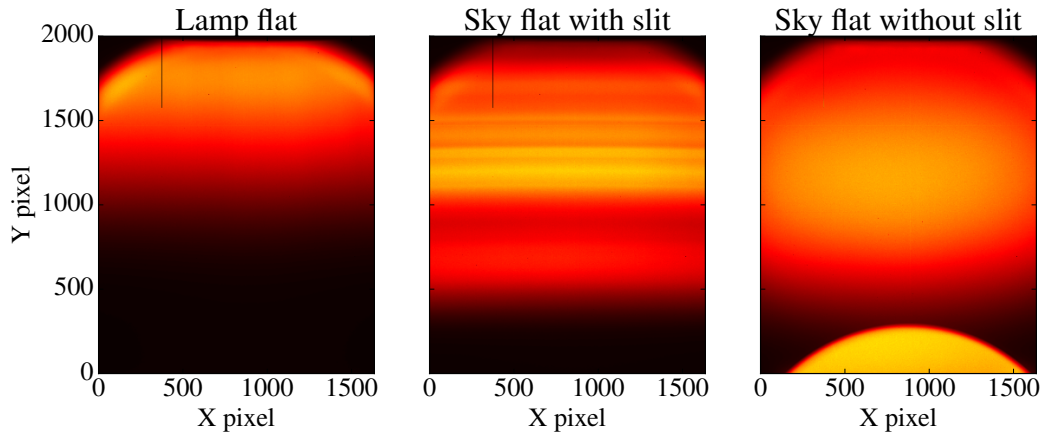
Figure 3.12: A comparison of example master flats created from a tungsten lamp flat (left panel), a twilight sky flat taken with a 40 arcsec slit (middle panel), and a twilight sky flat taken with no slit (right panel). These images show the flats before the removal of the sources' spectra. Wavelengths run from blue at the bottom to red at the top of the CCD. In each flat, the x-axis corresponds to the spatial direction and the y-axis corresponds to the spectral/dispersion direction.

spectroscopic sky flats with a slit (sky flats). Median combined examples of all 3 of these types of flat are shown in Fig. 3.12.

Fig. 3.12 highlights the issues faced here. The lamp flat contains very few blue photons, the sky flat contains strong sky lines and the slitless sky flat contains a strong semi-circular feature at the blue end of the CCD. This feature is undispersed, zeroth-order light which would otherwise be masked with the inclusion of the slit mask.

Fig. 3.13 is another representation of these 3 flats. This figure shows a vertical slice along the flat which highlights the issues raised above.

Fig. 3.14 shows an attempt at removing the spectrum of the tungsten lamp from the flat field. To do this, I followed the same method as used for reducing spectra with Molly[3]. This involved collapsing the flat field along the spatial direction, taking the logarithm of this collapsed function and then fitting a polynomial to this. This fit is shown in Fig. 3.14 (left column). I found that neither a polynomial nor a cubic spline offered sufficient fits to these spectral responses (Fig. 3.14). Fig. 3.14 highlights the lack of blue photons from the tungsten lamp, which is why these flats were not used to correct the data presented in Chapters 5 and 6.

I did not experiment further with the slitless sky flats due to the issue of the zeroth order light (Fig. 3.12). However, it is my aim in future reductions to

---

[3]http://www.astro.keele.ac.uk/jkt/GrSpInstructions/GrSpInstructions.html
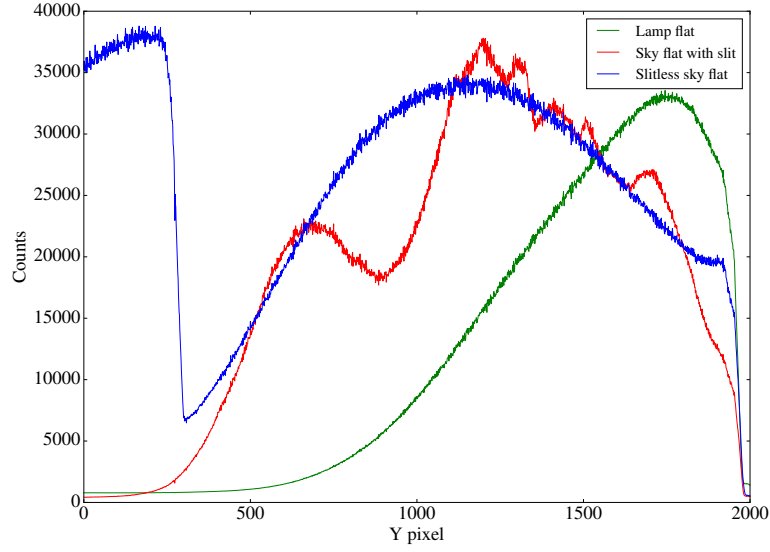
Figure 3.13: A vertical slice of the three master flat frames shown in Fig. 3.12. These are one pixel wide slices taken in the centre of the CCD. The x-axis corresponds to wavelength, running from blue (left) to red (right).
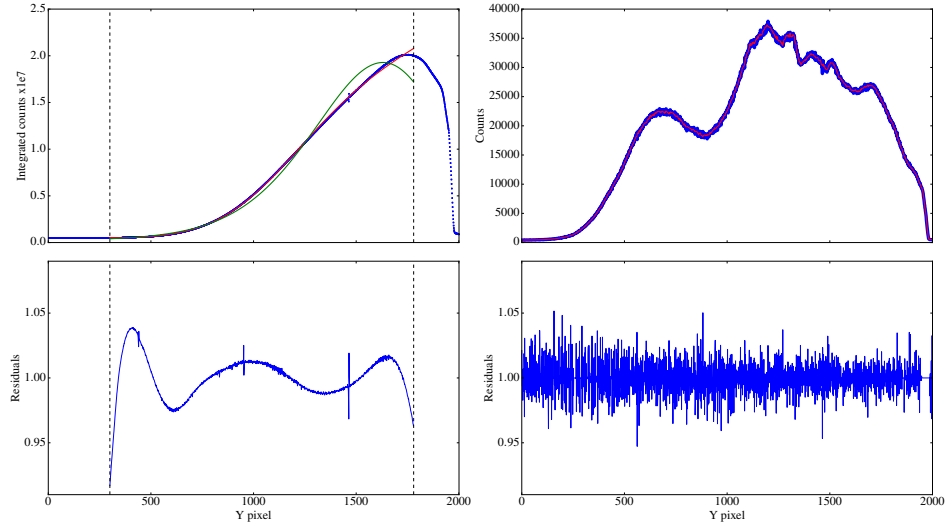
Figure 3.14: Left column: the lamp flat collapsed in the spatial direction is shown by the blue points in the top panel. The red line shows a 5th-order polynomial fitted to the response and the green line shows a cubic spline. The vertical dashed lines indicate the region that the fit was performed over in an attempt to improve the quality of the fit. The residuals to the 5th-order polynomial are shown in the lower panel. Right column: the twilight sky flat taken through the slit. A slice one column wide along the spectral direction is shown by the blue points in the upper panel. A running median of these points is shown by the red line. The spectral response divided by the running median is shown in the bottom panel.

experiment with these flats by using just the portion of the flat uncontaminated by zeroth order light.

The sky flats taken with the slit gave higher counts in the blue than the lamp flats (Fig. 3.13) while not being affected by the zeroth order light (Fig. 3.12), however they did contain sky emission features (Fig. 3.13). Given the poor fits using a polynomial and spline to the lamp flat response (Fig. 3.14) and that the sky flats contained much more structure (Fig. 3.13) I experimented with other methods for the sky flats.

To remove the sky lines, I first tried using a running median. I experimented with the size of the window used to calculate the median and found that 7 points per window were able to fit out the sky features while keeping the pixel to pixel sensitivity variations in the flat (Fig. 3.14, right column). This running median was calculated for each column individually, rather than the collapsed response, due to the curvature of the sky lines on the CCD (Fig. 3.12). An example master sky flat having divided by this function is shown in Fig. 3.15, which shows that sky lines have been removed but the pixel-to-pixel sensitivity variations remain.

In addition to using a running median, I also experimented with a Gaussian filter implemented through SCIPY's `ndimage` image processing library. This is equivalent to applying a low pass filter and therefore removes the high frequency signals, i.e. the pixel to pixel sensitivity variations, and instead leaves the sky lines which are more smoothly varying. I was then able to divide the master sky flat by the low frequency image, leaving behind the pixel-to-pixel sensitivity variations (Fig. 3.15, right panel). This technique preserves the bad pixel columns within the flat field, unlike the column-by-column running median (Fig. 3.15). However, it does blur out the features across columns which is a problem with this technique. This flat was not used other than for experimentation.

It is not essential to flat field the spectra presented in this thesis as they are spread over several thousand pixels, meaning that the pixel-to-pixel sensitivity variations will tend to average out. Therefore in all reductions I perform, I also experiment without flat fielding and compare the noise levels between these runs and runs where I have used a flat field.

3.6.2 Extracting the photometry

My photometric extraction pipeline relies on two files, a PYTHON script that performs the extraction and an extraction input file. This code has been designed to be simple to execute, such that it can be run during an observation, producing a near real-time transit light curve. To begin, I introduce an example extraction file below and then
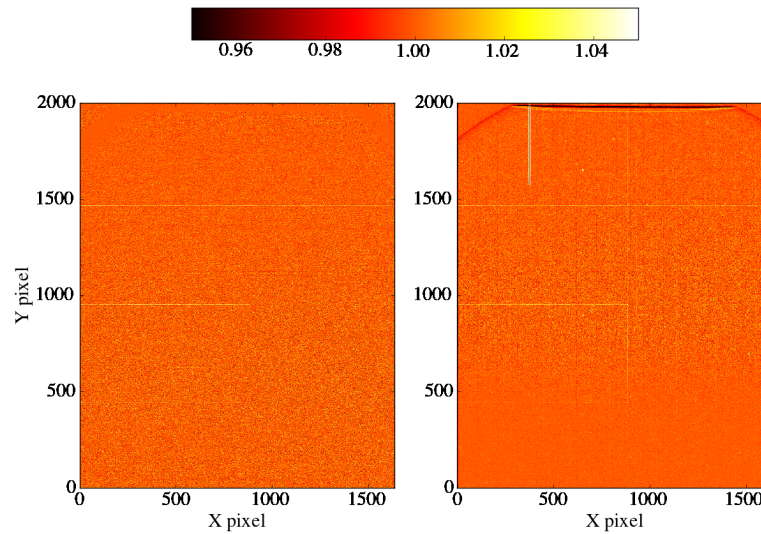
Figure 3.15: Left panel: a master sky flat after division by the running median shown in Fig. 3.14. Right panel: a master sky flat after division by a Gaussian filtered image.

I discuss how the code uses each parameter to trace the spectrum on the CCD and perform the photometry. The example extraction used below is the extraction as performed for the analysis of the WASP-80b data presented in Chapter 6.

```
# Extraction input file

science_list = science.lis
master_bias =
master_flat =

instrument = ACAM
nwindows = 1

verbose = 0
overwrite = 0

row_min = 300
row_max = 1900

trace_guess_locations = 250,1070
trace_search_width = 30
```

71

```
trace_gaussian_width = 5
trace_poly_order = 4
co_add_rows = 0

masks =
mask_width =

aperture_width = 26
background_offset = 20
background_width = 100
poly_bg_order = 2
```

One of the principal motivations behind writing this extraction pipeline was to have few input parameters that are all intuitive to understand. The `science_list` parameter is a list of science images over which the extraction runs. This can also be defined as the name of a single science file. If this option is selected, the code will then load in all subsequent files within the current working directory and assume these are also science images. This is useful when performing reductions during an observation as the files within the directory will be continually updating. The `master_bias` and `master_flat` parameters are the paths to both these files but either or both can be left blank if these corrections are not wanted. Again, this is useful when observing, when we are interested in a more qualitative view of the data.

The `instrument` parameter is currently set up to account only for either 'ACAM' or 'EFOSC', and depending which is selected, the correct instrument-specific parameters, such as the gain, are used. The code is set up to work for observations taken with 1 or 2 windows, set with the `nwindows` parameter.

When performing the reduction, diagnostic plots can be displayed by setting the `verbose` parameter to > 0, with the number indicating the number of seconds that each plot is displayed for before automatically closing. The example diagnostic plots shown in this section were created by changing this parameter from 0. `overwrite` is used to overwrite the output in the current working directory. This is mostly used when performing on-the-night reductions if wanting to restart a reduction with a different setup.

The `row_min` and `row_max` parameters define the minimum and maximum rows along the spectral direction at which to extract the spectrum between, as for ACAM the spectrum does not extend across the full vertical range.

The `trace_guess_locations` set the rough position of each star's trace in x-pixels (the spatial direction), and the number of locations given corresponds to the number of stars that will be extracted. In this way, I can extract spectra for all the stars on the CCD, not just those of the target and comparison. `trace_search_width` defines the width of the box that is searched around the guess locations to find the exact position of the trace. With these starting positions and a search box, the code then fits a Gaussian to the spatial profile within this box, with an estimate of the standard deviation of the Gaussian given by `trace_gaussian_width`. An example of this is shown in Fig. 3.16. The script does this iteratively row-by-row to extract the centre of the trace in x-pixel coordinates. In the event that no satisfactory fit with a Gaussian is found, the pixel position of the maximum count is assumed to be the centre of the trace. This is rare and for a typical night, such cases will only happen in total for ~ 20 rows for both the target and comparison traces (which extend over 1500 rows) across a few hundred frames. With the x-pixel coordinates as a function of y-pixel, the script fits a polynomial to these values, with an order given by `trace_poly_order`. The choice of the polynomial's order is made through a qualitative look at the residuals (Fig. 3.17), and I find that a 4th order is a good default to use. `co_add_rows` is a parameter that can be used to combine a number of rows along the spectral direction before trying to fit the trace. This could be useful for low signal cases but is typically not needed for the bright targets presented here.

Fig. 3.17 shows an example fit, with residuals, to the trace position along the spectral direction. The fitted trace is shown overlaid on a science image in Fig. 3.18. The residuals show some outliers which are not well fit by the 4th-order polynomial in this case, and some larger residuals at the location of strong absorption lines. However, these residuals are typically less than 0.2 pixels, and well within the typical aperture width (~ 30 pixels).

The extraction code is also set up to deal with contaminating stars which can be masked from the extraction using the `masks` and `mask_width` parameters. The first parameter defines the location of the contaminants with respect to the star of interest. For example, if there were contaminating stars located 50 pixels to the left and 100 pixels to the right of one target of interest, and contaminants located at 75 pixels to the left and 125 pixels to the right of a second target of interest, this would be defined as `masks = -50,100;-75,125`, with a semi-colon separating masks applied to different stars. The `mask_width` parameter then defines how many pixels should be masked at these locations. This masking assumes the curvature of the contaminant's spectral trace matches that of the target, which is a fair assumption
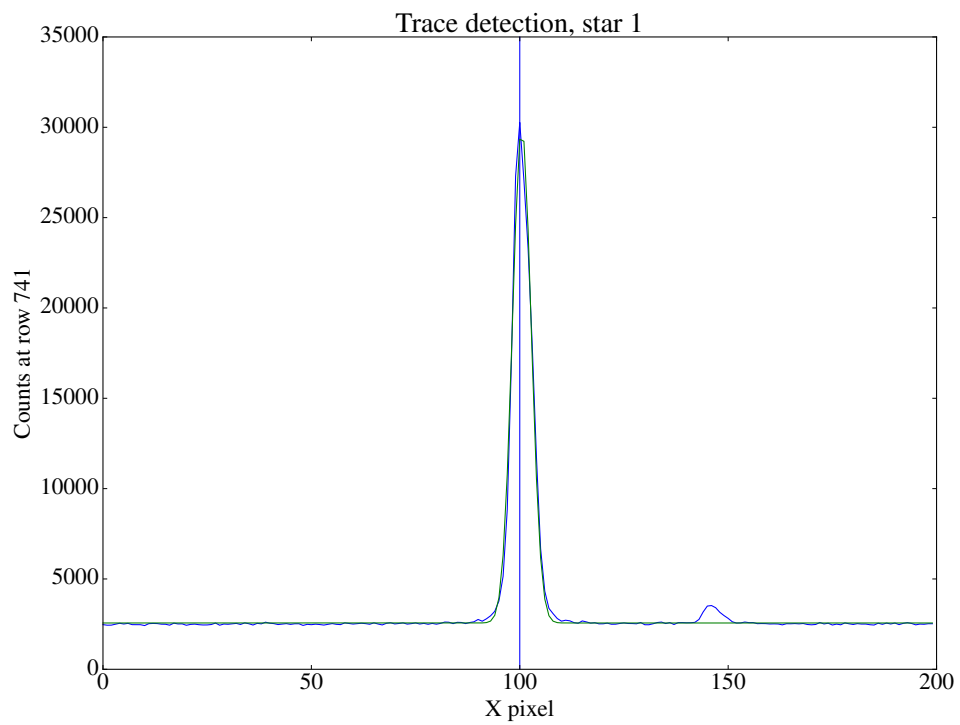
Figure 3.16: An example of a Gaussian fit (green line) to the spatial profile of the stellar trace (blue line). The star of interest is at $x = 100$ and the bump at $x \approx 150$ is a second, fainter star which is ignored by the fit.
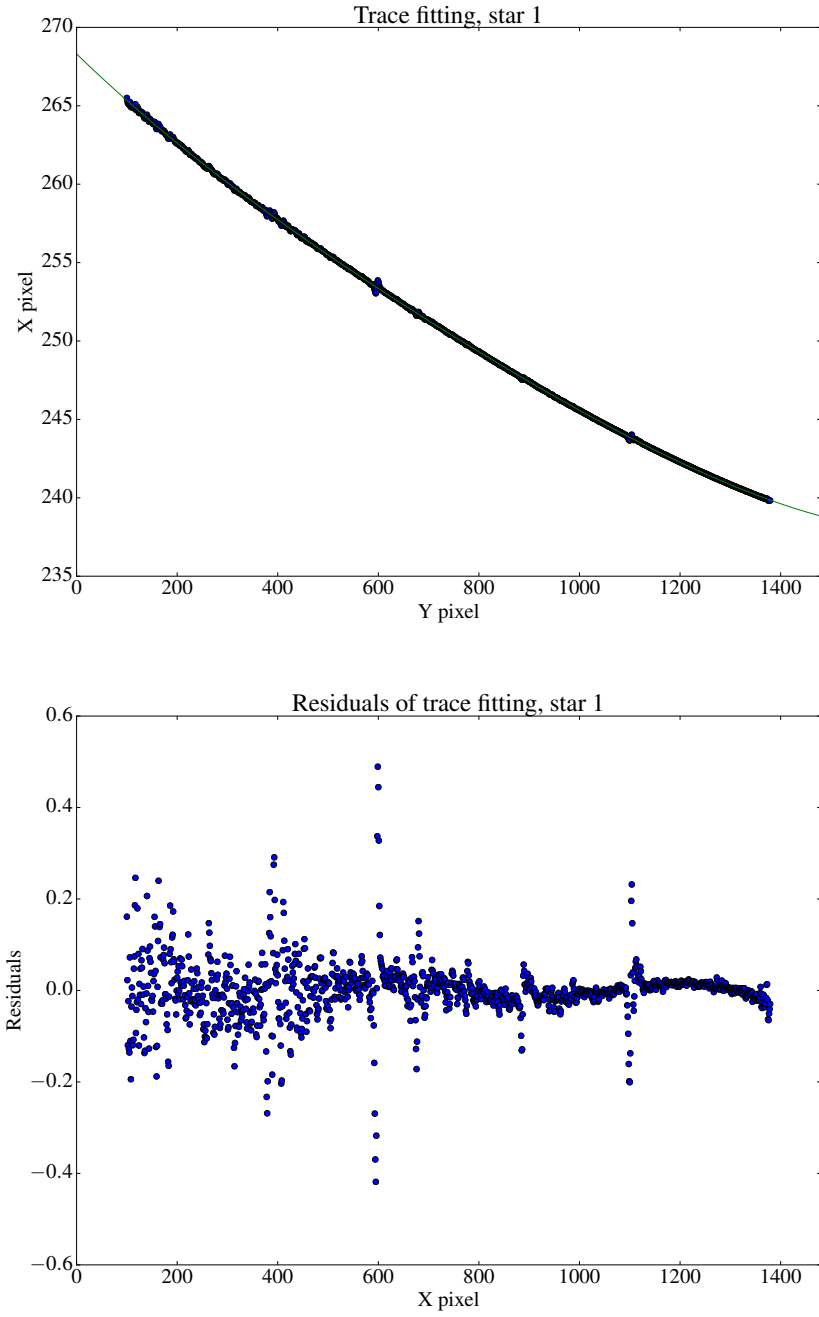
Figure 3.17: Top panel: plot showing a 4th-order polynomial (green) fitted to the central positions of the stellar trace resulting from the method shown in Fig. 3.16. Bottom panel: the residuals of the fit to the positions in the top panel.

given the usually small separations considered and the generous mask widths used.

After the traces have been located and the masks defined, the code performs aperture photometry row-by-row at the location of the trace. Initially the background is calculated by fitting a polynomial to two background regions either side of the trace, which is typically linear or quadratic. The width of these regions is defined by the `background_width` parameter and they are offset from the trace by a number of pixels set by the `background_offset` parameter. If masks have been defined, the masked pixels are ignored by the polynomial. Furthermore, any pixels that have values more than 3 standard deviations from the median of the combined background region are ignored from the fit, as these could be cosmic rays. Following this, the polynomial is then interpolated across the aperture to estimate the background within the aperture. This polynomial fit is then subtracted from the data, and the counts are summed within the aperture, which has a width set by `aperture_width`. The script is not currently set up to handle partial pixels and instead uses integer pixels for the aperture. An improved treatment of partial pixels will be included in future versions of the code.

Fig. 3.18 shows the apertures used to extract the target flux and estimate the sky background. In this figure it is clear that there are several stars which need to be masked from the background estimation. Fig. 3.19 indicates the pixels that were masked from the background fit.

In the above example, I only include the plots for the extraction of the target trace, since the code is set up to deal with one star at a time. However to extract the comparison star's flux, which is the bright star to the right in Fig. 3.18, the code automatically follows the same procedure.

The main purpose of this script is for simplicity of input, rather than speed of extraction. For the example above, the script is able to extract the target and one comparison spectrum for 10 science frames in 43 seconds, and can therefore be done at the telescope in real-time. This means a full night's data typically takes around 20 minutes to reduce. However, given the automation of the trace locating, the overall time taken to go from raw FITS files to extracted spectra is comparable to other extraction software.

3.6.3 Calculation of errors

To calculate the error in the extracted photometry, the script takes into account the shot noise (Poisson noise) associated with photon counting and the read noise of the instrument. The error per pixel is given as
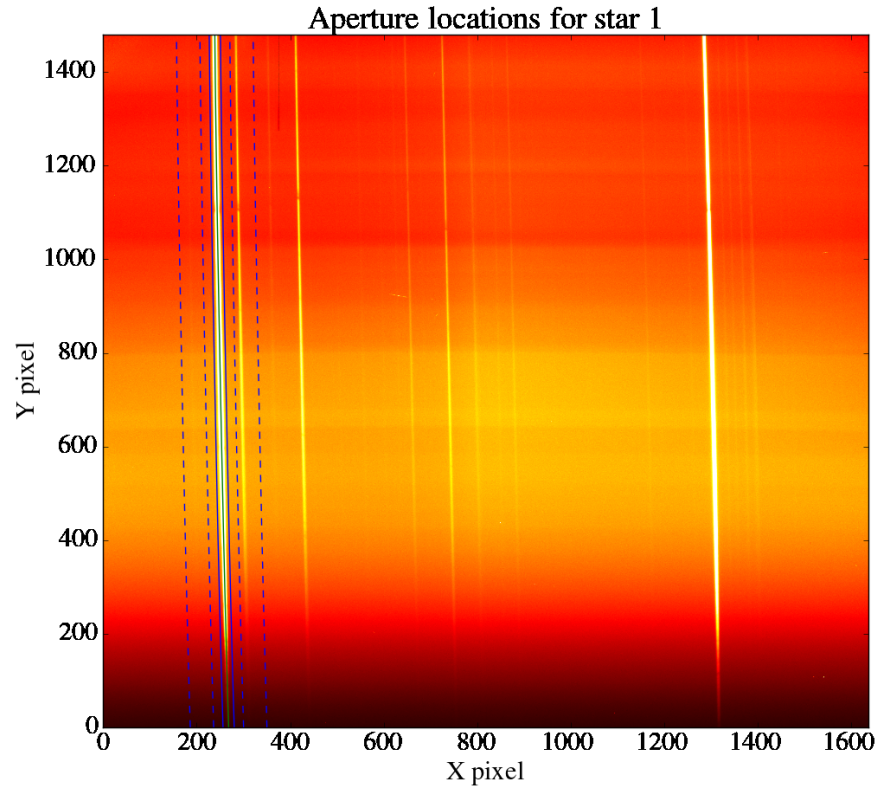
Figure 3.18: Locations of the apertures used to extract the flux of the target trace. The green line indicates the fit to the centre of the trace, the solid blue lines indicate the extraction aperture used and the dashed blue lines indicate the region in which the background is estimated. The contaminating stars falling within the background region are masked from the background estimation, as shown in Fig. 3.19.
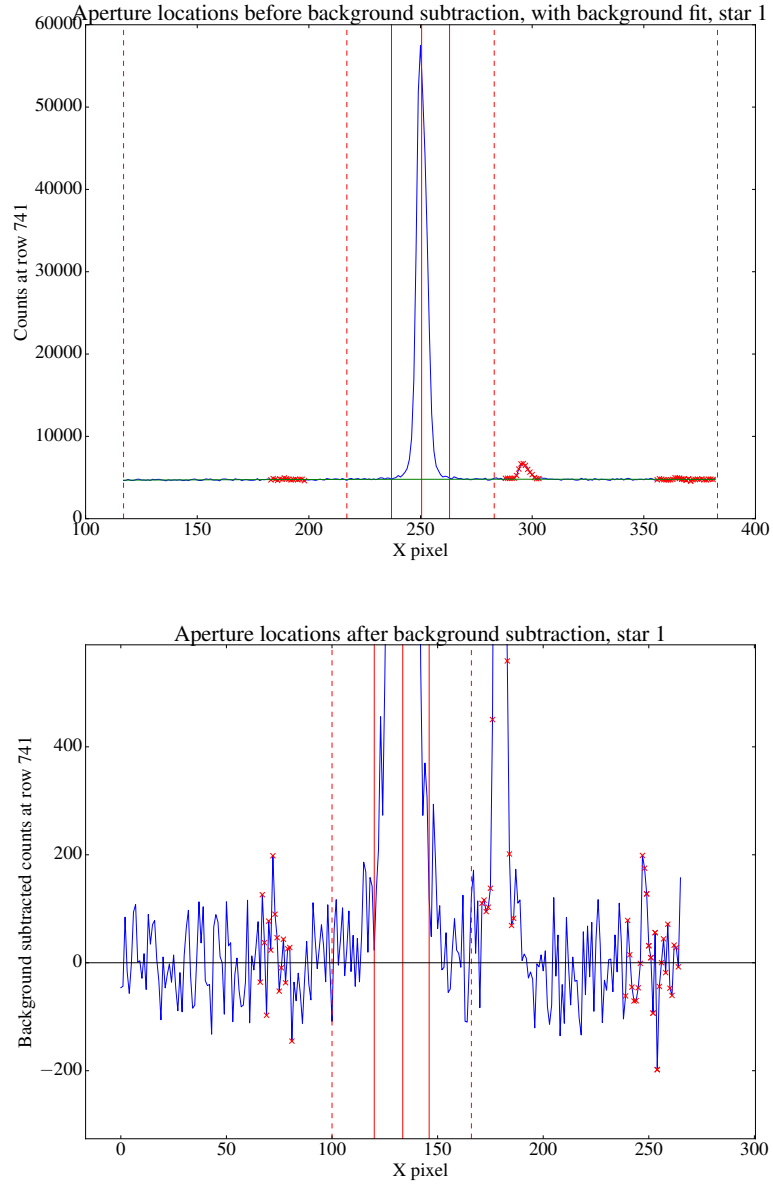
Figure 3.19: Top panel: the spatial profile of the target trace is shown in blue. The central, vertical solid red line indicates the fitted centre of the trace, with the surrounding solid red lines indicating the aperture over which the flux is integrated. The dashed red lines show the regions in which the background is estimated, with the red crosses showing the pixels that are ignored as they are the locations of contaminating stars. The green line indicates the quadratic polynomial fitted to the background regions, which is interpolated across the target aperture. Bottom panel: the residuals once the polynomial fitted to the background has been removed.

$$\sigma = \sqrt{F + R^2} \qquad (3.5)$$

where F is the combined number of recorded electrons from the star and sky background and R is the readout noise in units of e^- pixel^{-1}. Dark current is included in F since no dark correction is made but this is also negligible (Table 3.1). To find F, we need to convert from units of ADU to e^- which is done by multiplying by the gain, and is calculated before the flat field correction. The gain, readout noise and dark current for ACAM and EFOSC2 is listed in Table 3.1.

To calculate the error in the differential transit light curve, the errors in the target and comparison's spectra are added in quadrature.

Equation 3.5 does not take into account scintillation. However, the contribution from scintillation can be calculated through (Dravins et al., 1998)

$$S = 0.09 D^{-2/3} (\sec Z)^W exp(-h/h_0)/(2T)^{1/2} \qquad (3.6)$$

where D is the diameter of the aperture in centimetres, $\sec Z$ is the airmass, W is an exponent dependent on the wind direction relative to the pointing direction, h is the altitude of the telescope above sea level in metres, h_0 is the Earth's atmospheric scale height in metres (8000 m), and T is the integration time in seconds. Since the direction of the wind can change during an integration, I assume the maximum value for W of 2, the case where the telescope is pointing parallel to the wind direction. The WHT has an aperture diameter of 420 cm and the Roque de los Muchachos observatory is located at 2420 m above sea level.

Using the reduction of HAT-P-18b as an example, with an aperture width of 30 pixels and an integration time of 60 s, the mean fractional error from Poisson noise, read noise and scintillation is 3.8×10^{-3}, 2.8×10^{-3} and 1.6×10^{-4} respectively. Therefore scintillation does not form a major source of error, however, it is not negligible and will be included in future reductions.

3.6.4 Post-extraction processing

Ancillary outputs of the extraction script include the airmass, the FWHM of each trace, the recorded sky values in each of the regions to the left and right of each trace, the maximum counts along the length of the trace, and the measured x-pixel locations of the traces at each y-pixel. These values can be used as diagnostics of the extraction process and also as input parameters for subsequent decorrelation analyses.

The main output of the extraction is the extracted flux for each of the stars.

79

This flux is given as a function of y-pixel position and so post-extraction processing is required to convert these positions into calibrated wavelengths. To demonstrate this process, I again use the data of WASP-80b which the example extraction was performed on in section 3.6.2.

Initially, I begin with removing cosmic rays from the spectra. While the extraction code does make an attempt to remove cosmic rays that fall within the background regions by performing a sigma clip, it does not do the same for cosmic rays falling within the aperture. I experimented with using the Laplacian edge detection method of van Dokkum (2001) in order to mask cosmic rays within the entire frame before extraction. However, I found that the parameters governing this required tweaking from frame to frame which is undesirable when reducing hundreds of frames. For this reason, I decided to deal with the cosmic rays falling within the aperture during post-extraction.

To locate cosmic rays, I normalise all of the night's spectra by a spectrum free of cosmic rays, and locate all points that deviate from the normalised spectra. This is effective at locating cosmic rays as they produce large residuals in the normalised spectrum. In more recent reductions I have added a function which uses a median filter to locate cosmic rays, using the `medfilt` function within the SCIPY PYTHON library (Jones et al., 2001). The median filter calculates the median in a sliding box of a defined width, typically ~ 7 pixels, which slides across the extracted 1D spectrum. This running median can then be used to locate pixels that deviate significantly from the median, which are potentially cosmic rays.

Before removing each cosmic ray, I check that the flagged cosmics are indeed real as sometimes large absorption features do not divide perfectly, particularly as this is done before the spectra are aligned in pixel space. Once the flagged events have been confirmed, the remaining real cosmics are replaced by linearly interpolating between the nearest neighbouring pixels that are unaffected. An example of this process is shown in Fig. 3.20.

Once all the spectra have been cleaned of cosmic rays they can be aligned in pixel space. This is done by cross-correlating a number of absorption features within each spectrum with a reference spectrum. Sub pixel resolution is achieved by fitting a quadratic to the peaks in the cross-correlation functions and solving for the minimum, using a PYTHON script written by James McCormac. The reference spectrum is typically chosen to be near the mid point of the observations. The cross-correlation of each absorption feature with the reference returns the shift in pixels, allowing them to be aligned into the frame of the reference. An example of these shifts for a number of different features is shown in Fig. 3.21. In this case, the
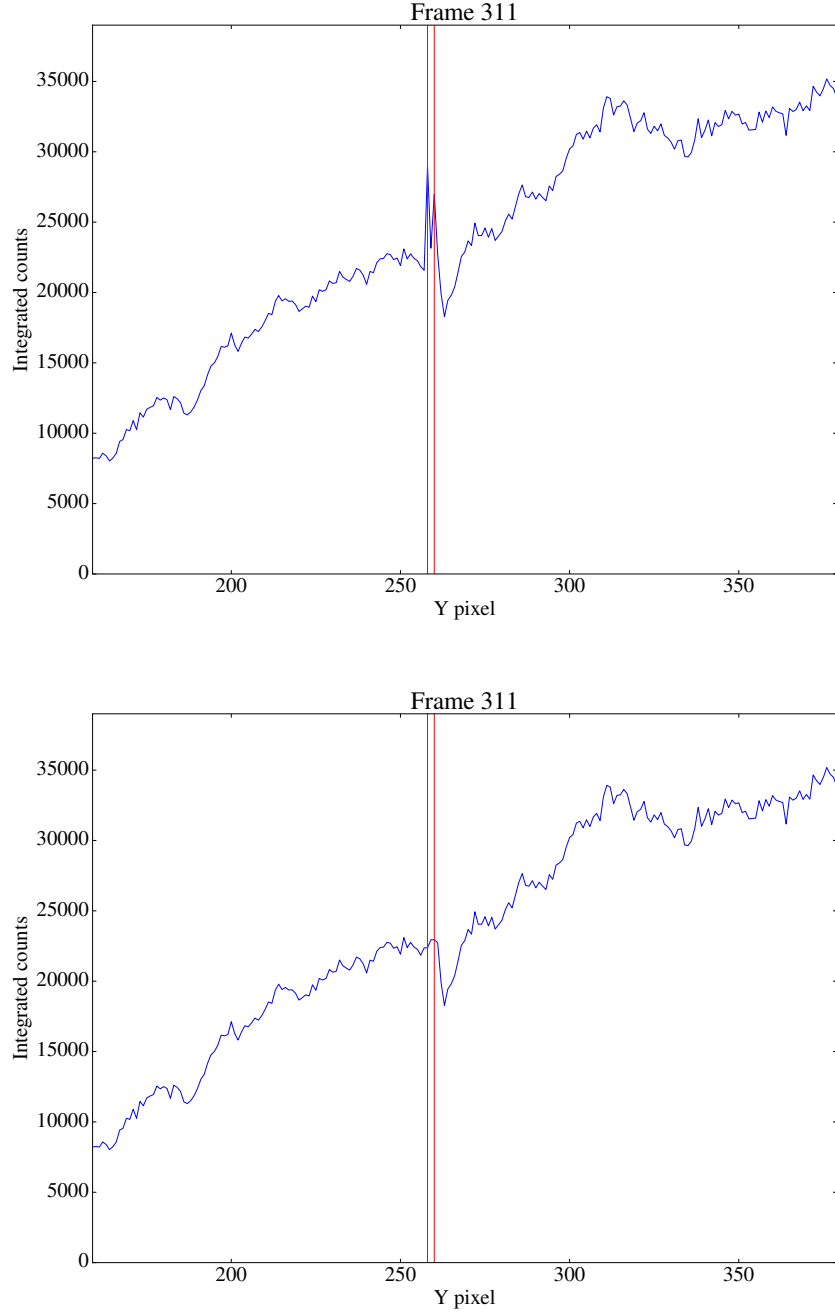
Figure 3.20: Top panel: examples of cosmic rays (at the positions of the vertical red lines) in the extracted stellar spectrum (blue line). Bottom panel: the same portion of the spectrum following the removal of the cosmic rays using the method described in the text.
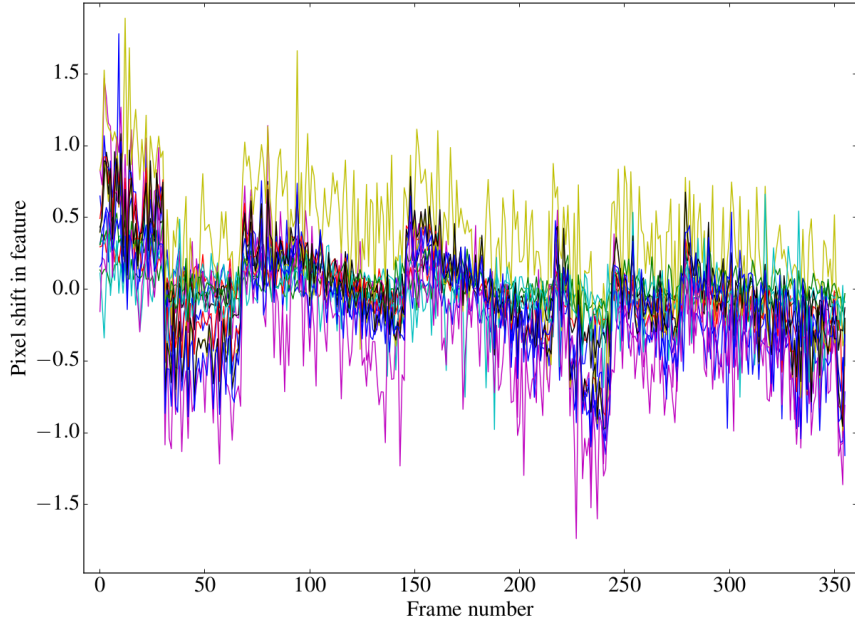
Figure 3.21: A plot showing the measured shifts of 15 absorption features over the course of a night's observations with respect to a reference frame. These features are located across the spectrum and so produce different amplitude shifts, as features nearer the edges of the spectrum are typically shifted by a larger amount. This plot signifies that there are a number of steps in these features, which are associated with manual guiding corrections made to the telescope during the observations.

shifts are characterised by high frequency shifts in the features' positions, associated with the error in the shift measurement, and lower frequency steps, associated with guiding corrections made to the telescope during the observations.

With the shift in pixel position resulting from the cross-correlation and the reference pixel of the absorption features, I fit a polynomial to describe the shifts as a function of y-pixel, resulting in a 'pixel solution' for each frame's spectrum. An example of this is shown in Fig. 3.22.

Next, each frame's spectrum is resampled onto the same grid as the reference spectrum by resampling the 'pixel solutions' onto the pixel grid of the reference. The cross-correlation step was necessary in order to relate each individual frame's pixel solution to that of the reference frame. The resampling is necessary as the spectra of the target and comparison need to be well aligned in wavelength space for accurate differential spectroscopy. The factors contributing to the shifts in the features include both astrophysical and instrumental effects but is mainly due to refraction in the Earth's atmosphere, which causes the curvature of the spectra on the CCD and is slightly different for the two stars. The instrumental effects include
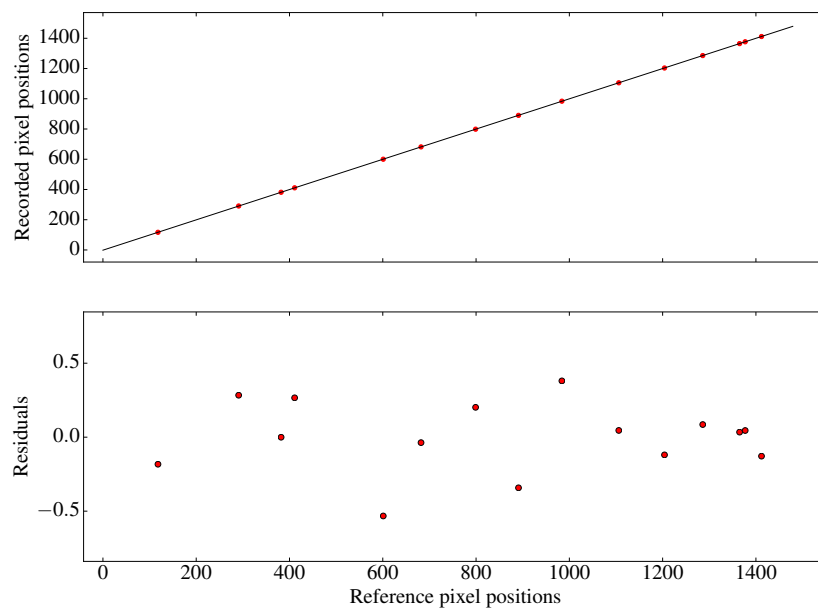
Figure 3.22: Top panel: a plot of the measured pixel positions of certain features for an example frame against the reference frame's pixel positions. These positions are indicated by the red points and the black line shows the third-order polynomial used to describe this relation. Bottom panel: the residuals, in pixels, once the polynomial in the top panel has been subtracted.
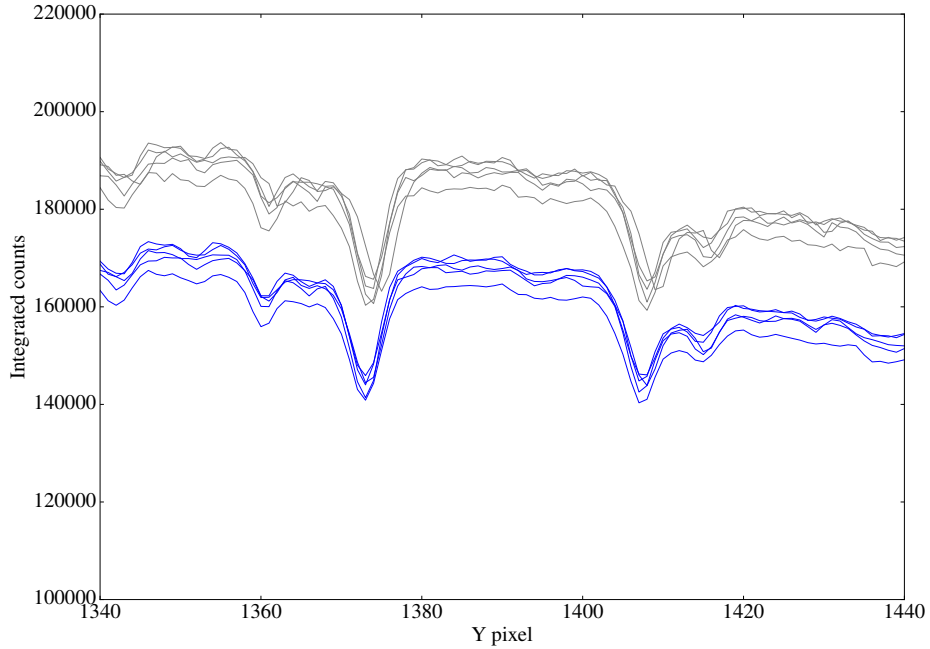
Figure 3.23: An example portion of a number of spectra extracted from different frames for the same object. The grey lines indicate the spectra before resampling and the blue lines are the same spectra following the resampling method discussed in the text. These spectra are offset in y for clarity.

flexure of the instrument as the telescope observes through differing elevations. The astrophysical effects are negligible at the resolution used here ($R \approx 400$) but include the Doppler shift in the absorption lines due to the orbital motion of the exoplanet and the Earth, and for other applications this effect would be measured. However, my science goals are not concerned with radial velocity measurements and so I do not make scientific use of the shifts in the absorption lines.

To perform the resampling, I make use of PYSYNPHOT[4] which conserves flux during the resampling process. An example portion of a spectrum before and after resampling is shown in Fig. 3.23.

With the spectra aligned in pixel space, I next perform a wavelength calibration to assign each pixel a wavelength in Å. I begin with using arc spectra to get an initial estimate of the wavelength solution. A narrower slit is used for the arcs as the wide slit produces wide arc lines which cannot be used for calibration. The narrow slit also avoids saturation which occurs when we try to take arc spectra with a wide slit.

---

[4]Distributed by STScI, //pysynphot.readthedocs.io/en/latest/
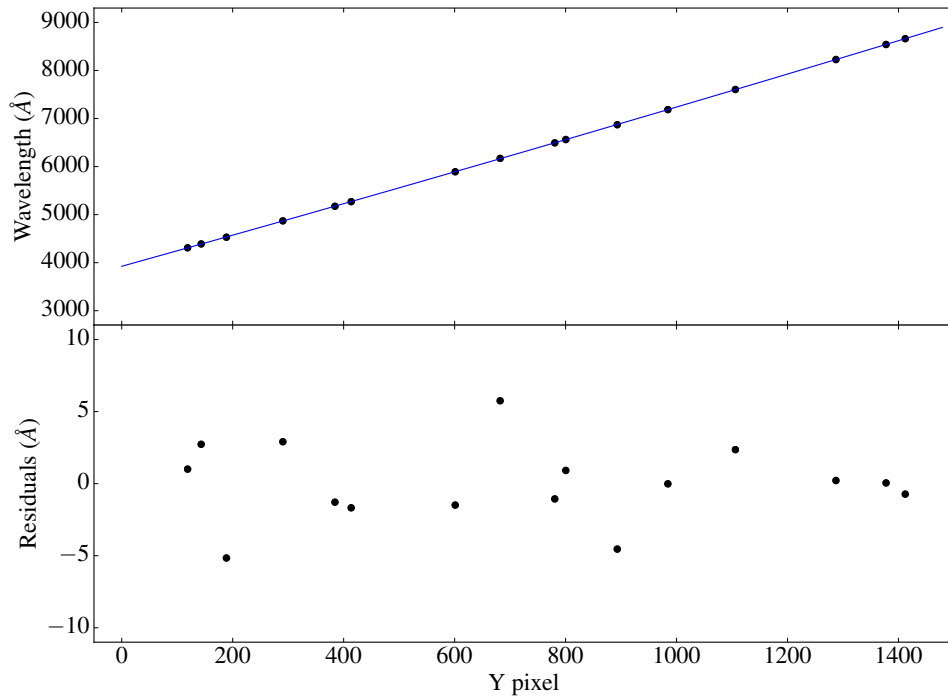
Figure 3.24: Example wavelength calibration. Top panel: the black points show the fitted centres, in y-pixels, of 16 absorption lines in an extracted stellar spectrum against the wavelength of these lines. The blue line is a quadratic polynomial fitted to the black points. Bottom panel: the residuals showing the difference between the quadratic polynomial and the measured line centres at the locations of the 16 lines.

To perform the arc calibration, the arc spectra are first extracted at the locations of the stellar traces. I then compare the locations of the peaks of the lines in the arc spectra to the instrument-specific tabulated arc values. Using the tabulated wavelength in Å and the measured peak position in the arc spectra, I use a cubic polynomial to describe the pixel position as a function of wavelength.

This polynomial serves as a useful start point to the wavelength calibration that I perform on the stellar spectra themselves, although the arc solution is offset due to the different slit used. To perform this absolute wavelength calibration, I use synthetic telluric and stellar spectra to identify absorption lines within the stellar spectra. I then fit a Gaussian to each of these absorption lines to retrieve the centre of the line in terms of pixels. With this I can then relate the true wavelength of the absorption feature to the pixel position of the absorption feature using a polynomial. An example of this is shown in Fig. 3.24. This results in wavelength calibrated and aligned spectra, examples of which are shown in Fig. 3.25. This figure shows the end product of this extraction and calibration process.

85
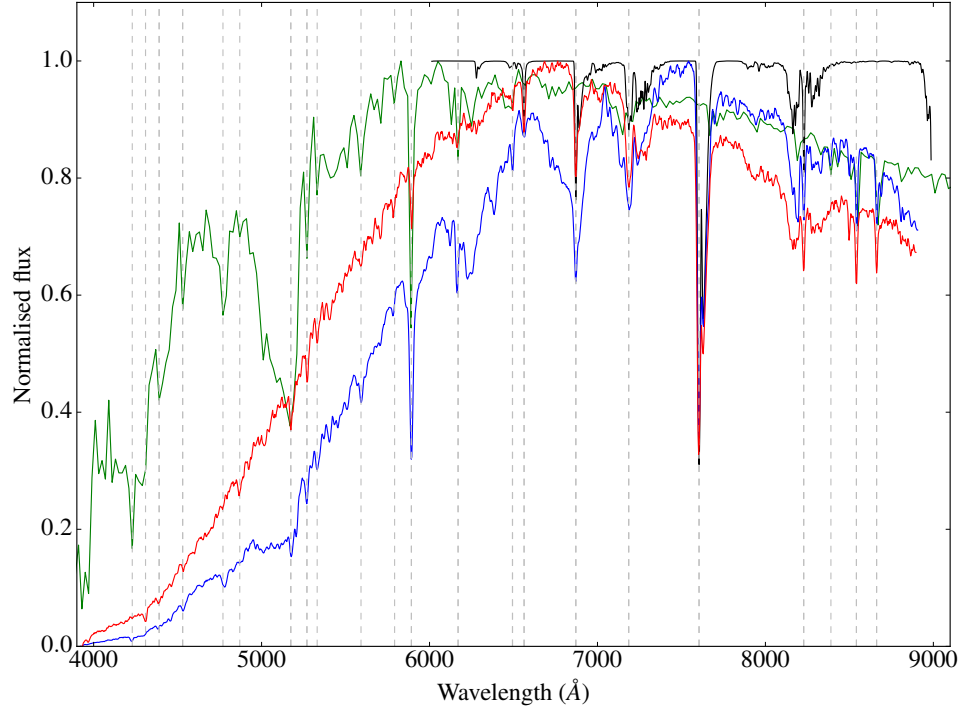
Figure 3.25: Example wavelength calibrated spectra for the target and comparison (blue and red solid lines, respectively). The green line shows an ATLAS9 template stellar spectrum (Kurucz, 1993) and the black line shows a template telluric spectrum (generated by Amanda Doyle using UCLSYN; Smith & Dworetsky, 1988; Smith, 1992; Smalley et al., 2001), both of which were used for the identification of lines in the extracted spectra. The vertical dashed lines indicate the central positions of lines used for calibration. All spectra in this plot have been normalised to unity for clarity.

Following the alignment of the spectra and the wavelength calibration, I am able to construct wavelength-dependent light curves for different portions of the spectrum. The method I chose to do this has varied between my analyses and so I describe this process in more detail in Chapters 5 and 6.

3.6.5   Comparison with independent pipelines

As an end to end test to check for major problems, I also compared the output from my pipeline with pipelines written independently by Matteo Brogi and Tom Louden in analyses of their own data.

Similar to my pipeline, the reduction performed by Brogi used a simple aperture extraction but with the sky background calculated from the median of two regions either side of the extraction aperture. In this instance, I performed a straight line background fit across the extraction aperture.

Both my pipeline and that of Brogi were run on the same WHT/ACAM data set, with the resulting white light curves plotted in Fig. 3.26 (top panel). This figure shows that the night was badly affected by clouds passing overhead, seen as drops in transmission in Fig. 3.26 (bottom panel), but that the outputs agree very well. The reason for the differing standard deviations in Fig. 3.26 could be due to the different aperture widths used in the two reductions.

The pipeline of Louden was used to analyse WHT/ACAM data of WASP-52b and is described in Louden et al. (2017). It also performs simple aperture photometry. The comparison between the white light curves resulting from my pipeline and Louden's pipeline is shown in Fig. 3.27. This again shows good agreement with my pipeline and also shows the presence of a large star spot resolved in both white light curves.

These comparisons encouraged me that my pipeline was performing correctly.

3.7   Light curve fitting

Using the outputs from the reduction tools presented above, the remaining analysis primarily concerns fitting the resulting transit light curves. To do this, I make use of the analytic transit light curves of Mandel & Agol (2002). In this paper, the authors derived analytic functions for transit light curves using both quadratic and non-linear limb darkening laws as described in Chapter 1. These functions have been implemented in PYTHON by several groups, but for my analyses I downloaded these functions from Ian Crossfield's PYTHON library[5]. In addition to the analytic

---

[5]https://people.ucsc.edu/~ianc/python/index.html
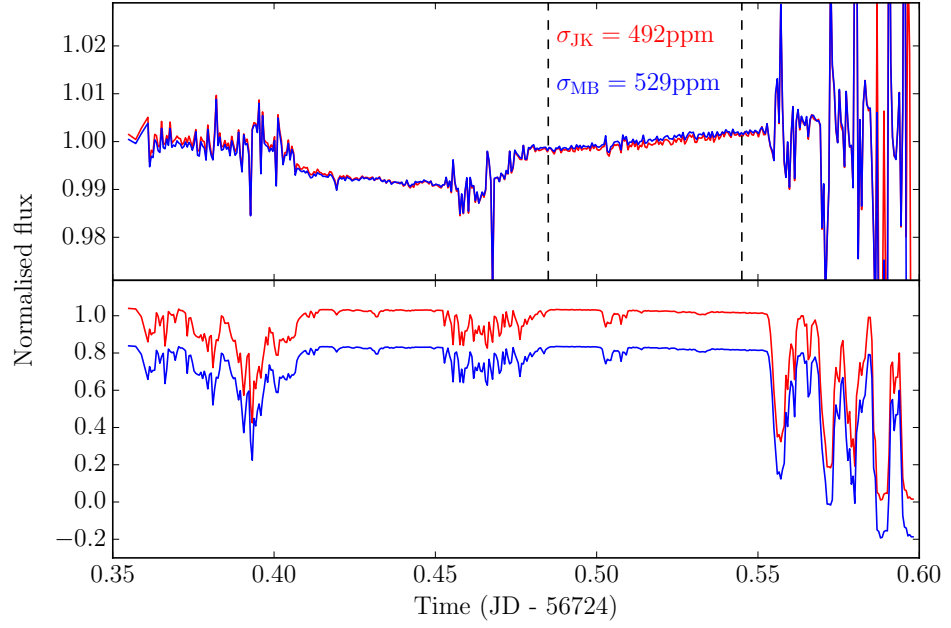
Figure 3.26: A comparison between the outputs of my pipeline with a pipeline written independently by Matteo Brogi run on the same night's data. Top panel: the white light curves from my pipeline (red) and Brogi's (blue). The standard deviations of a low-noise out-of-transit region (demarked by the vertical dashed lines) are shown on the plot. Bottom panel: the normalised flux for the target star resulting from my pipeline (red) and Brogi's pipeline (blue), offset by -0.2 for clarity. This panel shows the effects of clouds passing overhead causing a drop in transmission. This causes the noise in the white light curves shown in the top panel.
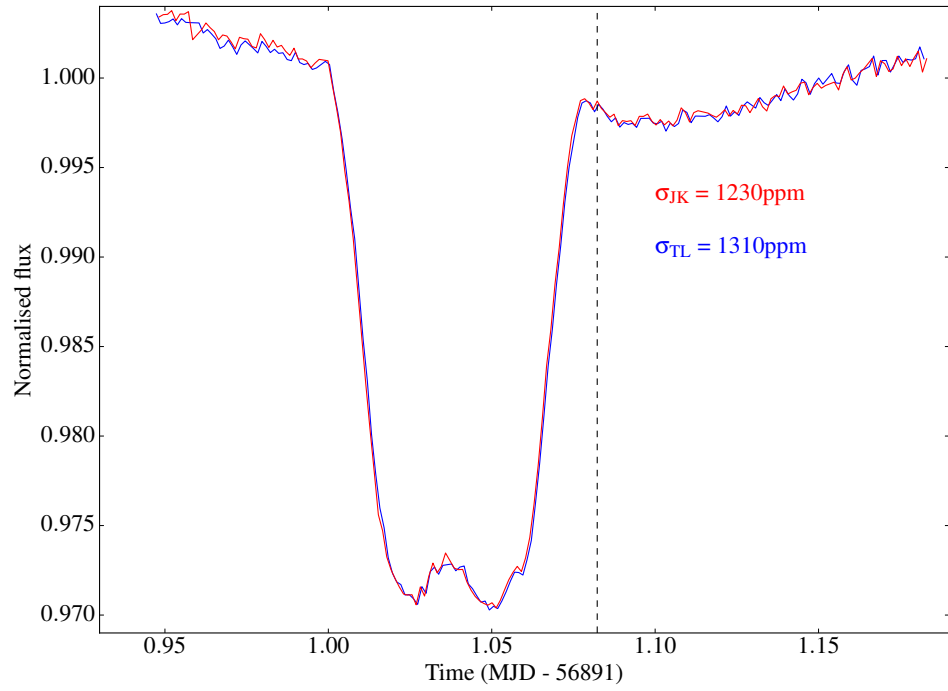
Figure 3.27: A comparison between WASP-52b's white light curves resulting from my reduction pipeline (red) and that of Tom Louden (blue), and which is published in Louden et al. (2017). The standard deviations of the light curves to the right of the vertical dashed line are given. These standard deviations are dominated by an airmass-related colour effect as these light curves do not have any detrending applied. Both light curves show the presence of a large star spot mid-transit.

transit light curve, I also simultaneously fit a systematic noise model to each light curve, which have varied between my analyses and are discussed in more detail in Chapters 4, 5 and 6.

3.7.1 Bayes' theorem

Bayesian statistics can be generalised as the degree of belief that one has in an outcome. A good, recent overview of Bayesian statistics and their use in exoplanet astronomy can be found in Parviainen (2017). Bayesian statistics describes the probability of an event given prior knowledge, and differs to frequentist statistics, which is the frequency of an outcome of a set of identical experiments. Bayes' theorem is applied to give the *posterior probability distribution*. It is written as

$$P(M|D) = \frac{P(D|M)P(M)}{P(D)} \tag{3.7}$$

where $P(M|D)$ is the posterior probability distribution, and is the probability of the model, M, given a set of data, D. $P(D|M)$ is the *likelihood* and is the probability of the data given the model. $P(M)$ is the *prior* probability and $P(D)$ is the *evidence*.

Priors can be used to encapsulate information that we already know about the system, such as information from previous experiments, or used to exclude unphysical values. In general, priors can either be *uninformative*, designed to contain as little information as possible, or *informative*, such as a normal distribution with a mean and standard deviation using previously measured values. It is important to understand what effects the prior distribution has on the posterior distribution and allow the data to drive the posterior, not the prior.

Let's consider the following example of Bayes' Theorem[6]. Three factories, A, B and C, manufacture lamps. The probability of a lamp being manufactured in factory A is 0.35, in B is 0.4 and in C is 0.25. The probability that a lamp from factory A is defective is 0.015, from factory B is 0.02 and from factory C is 0.025. Given this, let's work out the probability that a defective lamp (Def) is from factory C.

In this case, we want to calculate P(C|Def) and we know P(A), P(B), P(C), P(Def|A), P(Def|B) and P(Def|C). In this case, Bayes' Theorem gives

$$P(C|Def) = \frac{P(Def|C)P(C)}{P(Def)} \tag{3.8}$$

---

[6]Example from https://onlinecourses.science.psu.edu/stat414/node/43, accessed 12/12/17

and P(Def) can be written

$$P(Def) = P[(Def \cap A) \cup (Def \cap B) \cup (Def \cap C)] \tag{3.9}$$

which can be expanded to give

$$P(Def) = P(Def|A)P(A) + P(Def|B)P(B) + P(Def|C)P(C) \tag{3.10}$$

and substituted into equation 3.8 to give

$$P(C|Def) = \frac{P(Def|C)P(C)}{P(Def|A)P(A) + P(Def|B)P(B) + P(Def|C)P(C)} \tag{3.11}$$

We can now insert the numbers above to give

$$P(C|Def) = \frac{0.025 \times 0.25}{0.015 \times 0.35 + 0.02 \times 0.4 + 0.025 \times 0.25} \tag{3.12}$$

$$P(C|Def) = 0.32 \tag{3.13}$$

i.e. there is a 32% chance that the defective lamp was manufactured in factory C.

Now let's return to equation 3.7. In the case that the data can be described by a parametric model $m(\theta, x)$, plus some normally distributed, uncorrelated (white) noise σ, the likelihood can be written as

$$P(y_i|\theta, x_i, \sigma_i) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp-\frac{(y_i - m(\theta, x_i))^2}{2\sigma_i^2} \tag{3.14}$$

where θ are the model parameters, x_i are the covariates of the observation (such as the mid-exposure time and airmass in the case of transit observations), and σ_i is the uncertainty in the observation. For computational reasons it is often more practical to work with the *log-likelihood*, which is given by

$$\ln P(y|\theta, X, \sigma) = -\frac{1}{2}\left(n\ln 2\pi + \sum_i^n \ln \sigma_i^2 + \sum_i^n \frac{(y_i - m(\theta, x_i))^2}{2\sigma_i^2}\right) \tag{3.15}$$

for n observations.

For the purposes of this thesis, I employ Bayes' theorem through Bayesian

inference, that is to find the posterior probability distribution. To perform Bayesian inference, I use Markov chain Monte Carlo (MCMC) algorithms to sample from posterior probability distributions.

3.7.2 MCMC

MCMCs are designed to efficiently sample from posterior probability distribution functions in multi-dimensional parameter space. MCMC algorithms draw random samples for each of the model parameters (this is a Monte Carlo method), in a stochastic way that depends only on the current parameter values with no memory of earlier steps (this is the Markov chain).

MCMCs commonly employ a Metropolis-Hastings sampling algorithm (Metropolis et al., 1953; Hastings, 1970), which operates as follows. At each step in the chain, the algorithm calculates the probability ratio of the new value $P(M(\theta')|D)$ to the old value $P(M(\theta)|D)$. This probability ratio is then compared with a random number between 0 and 1. If the probability ratio is greater than this randomly generated number, the proposed location is accepted and $\theta' \to \theta$, otherwise the proposed location is rejected and θ remains unchanged. In this way, better fitting values are more likely to be accepted. Typically the acceptance fraction of proposed locations should be between ~ 0.2 and 0.5 for efficient sampling (e.g. Gelman et al., 1996). Since the model evidence, $P(D)$, cancels out in the posterior probability ratio, MCMC methods neatly avoid the computationally complex calculation of $P(D)$.

To implement MCMC within my work, I use the EMCEE package within PYTHON (Foreman-Mackey et al., 2013). Instead of evolving a single chain, EMCEE evolves an ensemble of *n walkers* allowing n locations within the parameter space to be sampled at each time step. This significantly decreases the autocorrelation time (increasing convergence speed). The chains will typically take some amount of time before they are sampling the high probability parameter space, which will increase if the chains are initiated far away from this region. This *burn-in* phase can be seen visually as the walkers explore the parameter space before settling into a region of higher probability. The burn-in phase is discarded before construction of the posterior probability distribution function.

3.7.3 Bayesian Information Criterion

The Bayesian Information Criterion (BIC) is a useful method to compare the goodness of fit between models of differing numbers of free parameters. In the case that two models produce an equally good fit, but one model uses an additional free pa-

Table 3.3: The strength of the evidence favouring one model over another given the ΔBIC between the two models. Table from Kass & Raftery (1995).

ΔBIC	Evidence against higher BIC
0 to 2	Not worth more than a bare mention
2 to 6	Positive
6 to 10	Strong
> 10	Very strong

rameter, this model will be disfavoured in a comparison of the BIC between the two models.

The BIC can be written as

$$BIC = \chi^2 + k\ln(n) \tag{3.16}$$

where k are the number of model parameters, n is the size of the sample, and χ^2 is given as

$$\chi^2 = \sum_i^n \frac{(data_i - model_i)^2}{\sigma_i^2} \tag{3.17}$$

where σ is the error in the data.

As noted in section 3.7.2, MCMC methods do not calculate the model evidence $P(D)$. As a result, the ΔBIC between models is often used to provide a useful approximation to the ratio of two evidences. Table 3.3 gives guidelines to the strength of evidence against a model with a higher BIC.

3.8   Gaussian Processes

Gaussian processes (GPs) are a powerful method to model correlated (red) noise and to determine its contribution to the uncertainties in model parameters. GPs are now becoming increasingly used in exoplanet astronomy, from transmission spectroscopy (e.g. Gibson et al., 2012a,b; Evans et al., 2015), to stellar activity modelling in radial velocity and photometry data sets (e.g. Haywood et al., 2014; Rajpaul et al., 2015), to detrending K2 data (Aigrain et al., 2016). More detailed reviews of GPs can be found in Rasmussen & Williams (2006), Gibson et al. (2012a), and Roberts et al. (2013).

A GP models the covariance between data points. It is defined by a covariance function (*kernel*) and a set of hyperparameters. The scalar log-likelihood function in equation 3.15 can be written more generally as

$$\ln P(y|\theta) = -\frac{1}{2}(n \ln 2\pi + \ln |\Sigma| + r^{\mathrm{T}} \Sigma^{-1} r) \tag{3.18}$$

where r are the residuals, given by $(y_i - m(\theta, x_i))$, and Σ is the covariance matrix given by

$$\Sigma_{i,j} = k(x_i, x_j) + \sigma^2 \delta_{ij} \tag{3.19}$$

where δ_{ij} is the Kronecker-delta function, $k(x_i, x_j)$ is the kernel, x is the data and σ is the white noise term.

For white noise, the covariance matrix only includes the diagonal term, with off-diagonal elements equal to zero. Incorporating a red noise kernel implies correlation between points, and therefore the off-diagonal elements are non-zero.

Two commonly used kernels for GPs are the *squared exponential* kernel and the *Matérn* class of kernel, and these are the kernels employed in this thesis. The choice of kernel will affect the noise properties of the GP (Fig. 3.28) and so it is important to choose a kernel that best describes the noise in the data. However, even given this choice, the non-parametric nature of the GP still makes it more powerful than parametric methods for modelling correlated noise (Gibson et al., 2012a).

The squared exponential kernel is defined as

$$k_{\mathrm{SE}}(x_i, x_j) = a^2 \exp\left(-\left(\frac{x_i - x_j}{\tau}\right)^2\right) \tag{3.20}$$

where a is the output-scale amplitude and τ is the length-scale, which defines the distance over which points are correlated. The parameters a and τ are referred to as the *hyperparameters* of the GP and are marginalised over.

The Matérn class of kernel is similar to a squared exponential kernel but allows for higher frequency variance (Fig. 3.28). It is described by

$$k_{\mathrm{M}}(x_i, x_j) = a^2 \left(1 + \sqrt{\frac{3(x_i, x_j)^2}{\tau^2}}\right) \exp\left(-\sqrt{\frac{3(x_i, x_j)^2}{\tau^2}}\right) \tag{3.21}$$

where a and τ are again the amplitude and length-scales respectively.

Various other kernels exist in addition to the squared exponential and Matérn kernels, including periodic kernels. These can be particularly useful when modelling the effects of stellar activity on light curves (e.g. Haywood et al., 2014).

Fig. 3.28 shows a GP with a white noise kernel, squared exponential kernel, a Matérn kernel, and an example of a periodic kernel.

The processing time of a GP goes as $\mathcal{O}(n^3)$ and therefore scales poorly with increasing n (the number of input measurements). Due to this limiting factor, there have been attempts to speed up the execution of GPs. Indeed, the GEORGE PYTHON library used here employs a Hierarchical Off-Diagonal Low-Rank (HODLR) solver which allows the GP to run in $\mathcal{O}(n \log^2 n)$ time (Ambikasaran et al., 2014). To generate the starting values for the hyperparameters, it can be beneficial to optimize the log-likelihood with respect to the hyper-parameters (e.g. Gibson et al., 2012a). To do this, I use a Nelder-Mead simplex algorithm (Nelder & Mead, 1965), implemented through the SCIPY PYTHON library. This algorithm uses a simplex to evaluate the function space in question. In 2 dimensions, a simplex is defined by a triangle, in 3 dimensions by a tetrahedron and so on. The simplex then transforms at each iteration, by reflecting the point of the simplex where the function is at its largest through its opposite face to a lower point in the function space (Press, 2007). It is helpful to imagine the simplex oozing downhill in this function space to find the minimum (Press, 2007).

The optimized values for the hyperparameters resulting from the Nelder-Mead algorithm can then be used as starting values for the MCMC. This ensures that the GP hyperparameters are starting at reasonable values, shortening the convergence time.

I have found GPs to be very useful at modelling the noise in the light curves presented in Chapters 5 and 6. I describe how they have been used in each case in the relevant chapters.
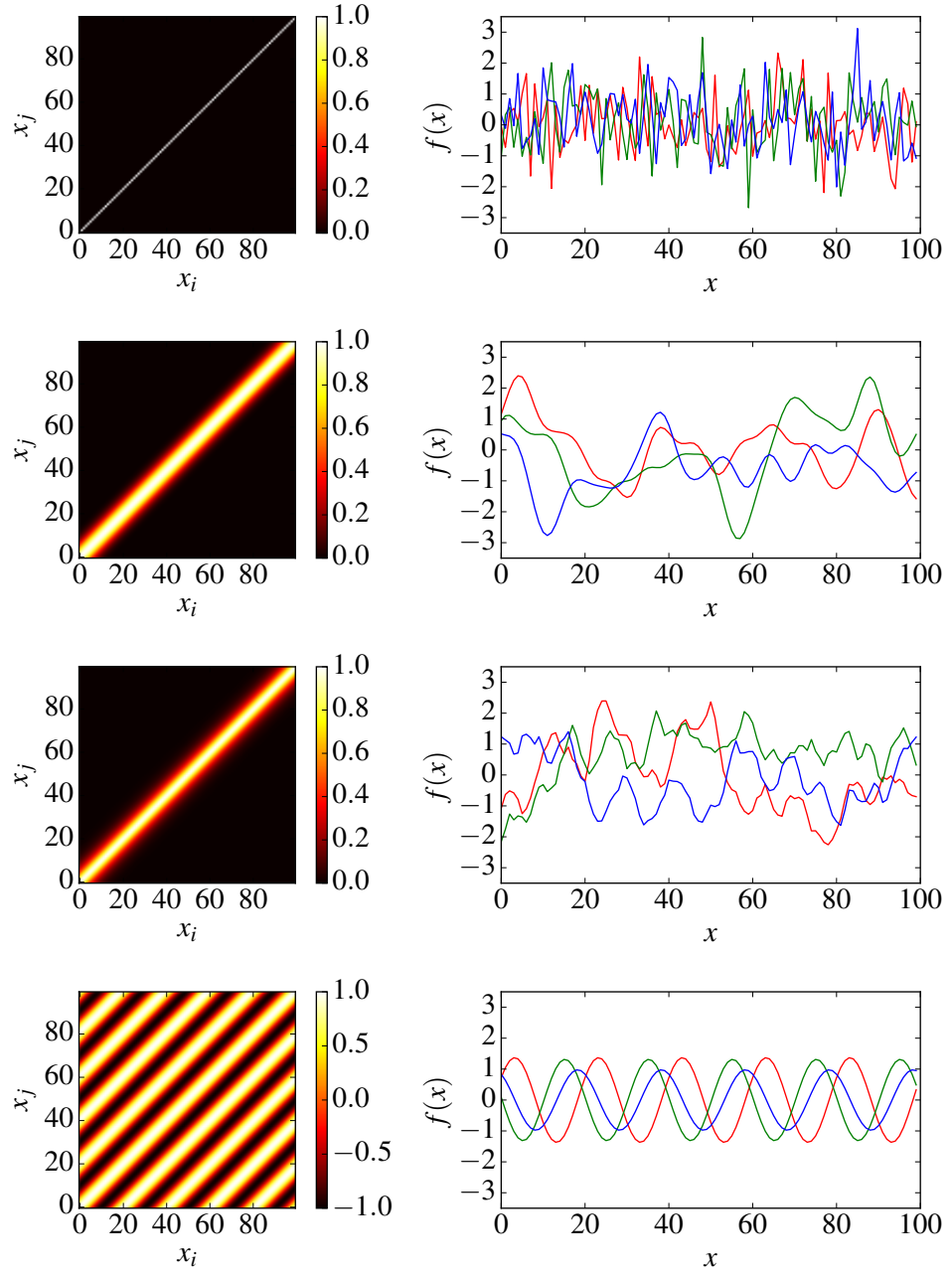
Figure 3.28: The left column shows the covariance matrices for different kernel classes, with the right column showing 3 randomly drawn samples from a GP defined by these kernels and evaluated at 100 time steps. Top row: a white noise kernel, with a variance of 1. Second row: a squared exponential kernel with an amplitude (a) of 1, and length scale (τ) of 20. Third row: a Matérn kernel with $a = 1$ and $\tau = 20$. Bottom row: an example of a periodic kernel with a period of 20.

Chapter 4

Transmission photometry of WASP-52b

In this chapter, I measure the transmission spectrum of the extremely inflated hot Jupiter WASP-52b using the ULTRACAM instrument mounted on the WHT, as introduced in Chapter 1. Simultaneous photometric observations were made in Sloan Digital Sky Survey u', g' and a filter centred on the sodium doublet (Na I). I find that Rayleigh scattering is not the dominant source of opacity within the planetary atmosphere, instead finding a transmission spectrum more consistent with wavelength-independent opacity, which I ascribe to the presence of clouds. I detect an in-transit anomaly that I attribute to the presence of stellar activity and find that this feature can be more simply modelled as a bright region on the stellar surface akin to solar faculae rather than spots. A spot model requires a significantly larger planet/star radius ratio than that found in previous studies. My results highlight the precision that can be achieved by ground-based photometry with errors in the scaled planetary radii of less than one atmospheric scale height, which is comparable to the precision achieved with the *Hubble Space Telescope*.

The results presented in this chapter were published by Kirk et al. 2016, MNRAS, 463, 2922.

4.1   Introduction

As detailed in Chapter 1, transmission spectroscopy has been successfully conducted at low and high resolution both from space and the ground. High resolution measurements from the ground do not require a comparison star to be observed simultaneously with the target as the planet's spectrum can be resolved from the telluric

spectrum (e.g. Brogi et al., 2018). However, low resolution ground-based measurements require a comparison star to be observed simultaneously with the target so that the effects of observing through Earth's atmosphere can be removed. This can lead to systematics arising from differential slit losses between the target and comparison if one or both of these stars begins to drift on the CCD. For this reason, such studies often use very wide slits, such as the 27 and 40-arcsec wide slits used in this thesis. An example of the effect of differential slit losses was shown in Sing et al. (2012), where the authors used a 5 arcsec-wide slit for GTC/OSIRIS observations of XO-2b.

Another method to avoid this problem, however, is to use photometers to perform broad-band observations at multiple wavelengths simultaneously. As mentioned in Chapter 1, the multi-band photometers GROND and ULTRACAM have both been used multiple times to perform such observations.

Among the results from multi-band photometry are: the detection of a steep bluewards slope in the transmission spectrum of a very hot Jupiter from GROND observations (WASP-103b; Southworth et al., 2015; Southworth & Evans, 2016), the lowest mass exoplanet to have a claimed detection of an atmosphere (GJ 1132b, $1.6\,M_\oplus$; Southworth et al., 2017), a detection of a Rayleigh scattering slope in a Neptune sized planet (GJ 3470b; Dragomir et al., 2015), enhanced absorption around the sodium feature of WASP-17b (Bento et al., 2014), and transmission spectra most consistent with clouds (e.g. Mallonn et al., 2015b).

By observing the planetary radius at carefully selected wavelengths we can probe for the existence of Rayleigh scattering and the broad absorption wings of the Na I doublet that are expected to be present in the cloud-free atmospheres of hot Jupiters (Seager & Sasselov, 2000). In this chapter, I present a transit of the hot Jupiter WASP-52b in three wavelengths using the triple beam CCD photometer ULTRACAM.

WASP-52b (Hébrard et al., 2013) is an extremely inflated hot Jupiter, with a mass of $0.46\,M_J$ and radius of $1.27\,R_J$ giving it a mean density of $0.299\,\mathrm{g\,cm^{-3}}$ ($= 0.22\,\rho_J$). It orbits its $0.87\,M_\odot$ K2V host star with a period of 1.75 d. Because of the combination of the inflated planetary radius with the small radius of the host ($0.79\,R_\odot$), it shows a deep transit in the WASP photometry (2.7 %). From the table of system parameters (Table 4.2; Hébrard et al., 2013), and assuming a Jupiter mean molecular mass of 2.3 times the mass of a proton, the scale height of WASP-52b is calculated to be 731 km. This makes WASP-52b an exceptional target for transmission spectroscopy as the difference in the transit depth corresponding to one atmospheric scale height is 440 ppm, at least three times greater than that

of HD 189733b.

At the time of these observations, in 2012, WASP-52b was in the top 5 exoplanets for transmission spectroscopy, as ranked by my order of merit (see Chapter 2). However, my transmission spectrum, published in 2016, was the first concerning WASP-52b. The reason for the absence of studies before this point is unclear but could perhaps be due to the host star's activity which, as shown in this chapter, has significant effects on the transit light curves.

4.2   Observations

WASP-52b was observed on the night of 2012 September 7 using the ULTRACAM instrument (Dhillon et al., 2007) on the WHT. These observations were taken by Stuart Littlefair as part of WHT programme WHT/2012B/34 (PI: Wheatley). ULTRACAM is particularly useful for the ground-based application of transmission spectroscopy as it simultaneously takes measurements at three different wavelengths, enabling the transit depth to be measured as a function of wavelength (e.g. Copperwheat et al., 2013; Bento et al., 2014). The use of frame transfer CCDs allow for high frame rates (up to $500\,\mathrm{Hz}$) with little dead time ($24\,\mathrm{ms}$), which is useful to avoid saturation when observing bright stars and enables many more sky flats to be taken.

The observations were made using Sloan Digital Sky Survey (SDSS) u' ($\lambda_{\mathrm{central}}$ = $3557\,\text{Å}$, FWHM = $599\,\text{Å}$) and g' ($\lambda_{\mathrm{central}}$ = $4825\,\text{Å}$, FWHM = $1379\,\text{Å}$) filters and a filter centred on the sodium doublet ($\lambda_{\mathrm{central}}$ = $5912\,\text{Å}$, FWHM = $312\,\text{Å}$). These filters were selected to probe for Rayleigh scattering and TiO by observing the difference in transit depth between the u' and g' bands and to search for the broad wings of the sodium doublet with the Na I filter. The filter response curves are shown in Fig. 4.1.

The observations were performed with moderate telescope defocusing ($\sim$ 3 arcsec) in windowed mode with exposure times of $0.76\,\mathrm{s}$ in the red and green channels and a cadence of 0.79 seconds. Telescope defocussing can be useful to avoid saturation of bright stars but also to reduce flat-fielding errors as the stellar PSF is spread over a larger number of pixels (e.g. Southworth et al., 2009). Because of the reduced photon count in the blue channel, multiples of 10 frames were averaged on-chip before readout, leading to a $7.9\,\mathrm{s}$ cadence in this channel.

The moon was at 54% illumination on the night of the observations and I analysed the data with an airmass varying from $1.41 \rightarrow 1.06 \rightarrow 2.00$.

ULTRACAM's field of view allows for simultaneous observations of a compar-
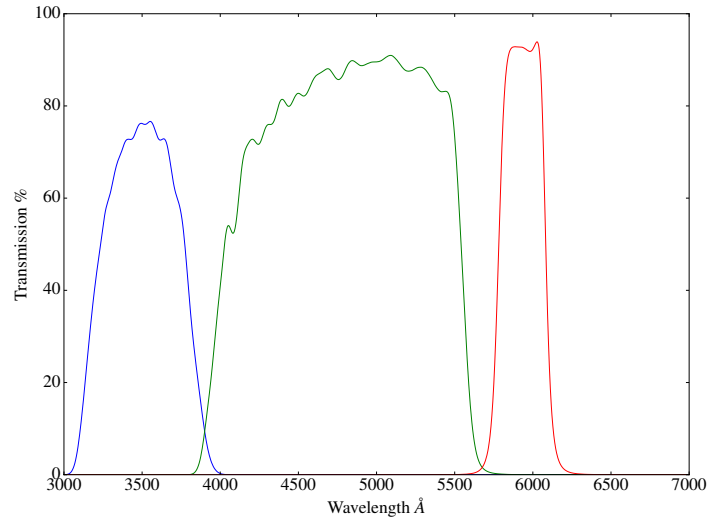
Figure 4.1: Filter response curves of the 3 filters used to perform the observations in this chapter: SDSS u' (blue); SDSS g' (green) and Na I (red).

ison star which is necessary for differential photometry. The comparison star has a similar colour to WASP-52 but is brighter (BD+08 5023, 23:14:12.026 +08:50:56.28). This star has a V magnitude of 10.59 and $B - V$ colour of 0.86, whilst WASP-52 has a V magnitude of 12.22 and $B - V$ colour of 0.82. The comparison star was checked and found to be photometrically stable.

4.3 Data Analysis

4.3.1 Data reduction & ULTRACAM pipeline

All the data were reduced using the ULTRACAM data reduction pipeline[1] (Dhillon et al., 2007). ULTRACAM uses 'runs' which are series of frames taken consecutively with the same setup. For example, a number of biases taken with the same setup would constitute a different run to a number of flats. The basic reduction procedure involves median combining biases within a run, and median combining flat fields within a run. Then, using these calibration frames, aperture photometry is performed at user-defined aperture locations, with parameters set in a reduction file called `reduce.red`. The aperture selection is described in more detail below.

Flat fields were taken on the night of the transit without windowing. However, the science data were taken with 4 different windows. Because of this, separate master biases had to be created for both the flats and science frames, as window-

---

[1]http://deneb.astro.warwick.ac.uk/phsaap/software/ultracam/html/index.html
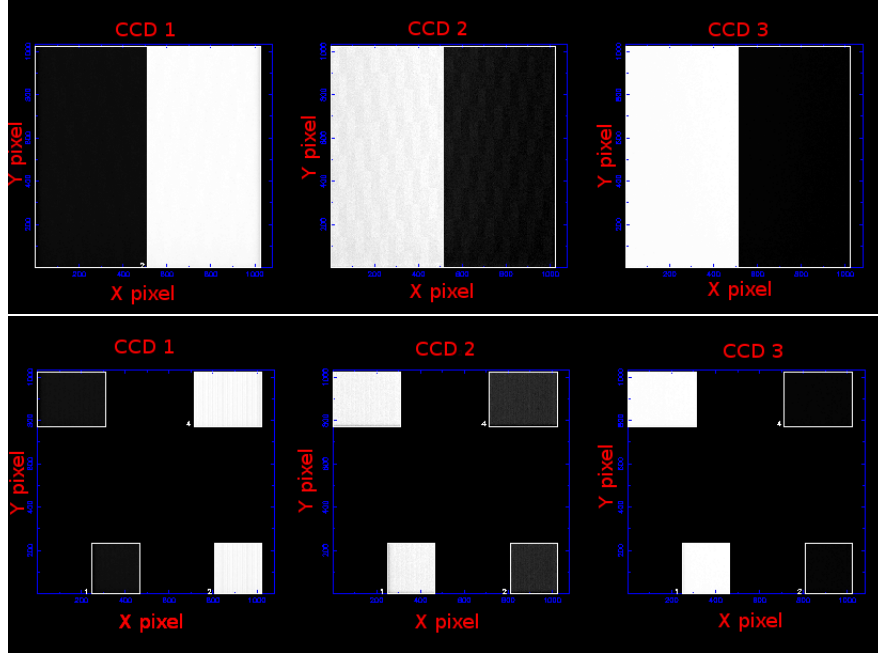
Figure 4.2: Top row: the median combined master biases used to correct the flats. Bottom row: the median combined master biases used to correct the science images. In each row, the three panels correspond to the 3 filters used, running from red (left panel) to blue (right panel).

ing will affect the bias close to the edges of the window. The fact that the flats were taken with a different window setup did not mean they could not be used to calibrate the science frames however. This is because the flat field measures the illumination levels and pixel-to-pixel sensitivity variations, which do not change when introducing windows. The master bias used to correct the flat fields was created by median combining 101 bias frames, and for the science master bias, 100 bias frames were median combined to create a master bias. These two master biases are shown in Fig. 4.2. The bias is different between the two halves of each chip as two readout amplifiers are used, one for each side of the chip, which speeds up the readout. These amplifiers have slightly different bias levels causing the difference in Fig. 4.2.

The master flat was created by median combining 176 twilight sky flats, after subtraction of the relevant master bias. The master flat is shown in Fig. 4.3. The blue diamond effect can be seen in the u' band (CCD 3). This is due to surface imperfections on the silicon chip. Bad pixels can be seen in all three bands, particularly in the g' band (CCD 2), and we try to avoid these when acquiring stars. Vignetting can also be seen in all three bands. The vertical lines in the middle of the chips are a result of the manufacturing process, where there is a slightly larger
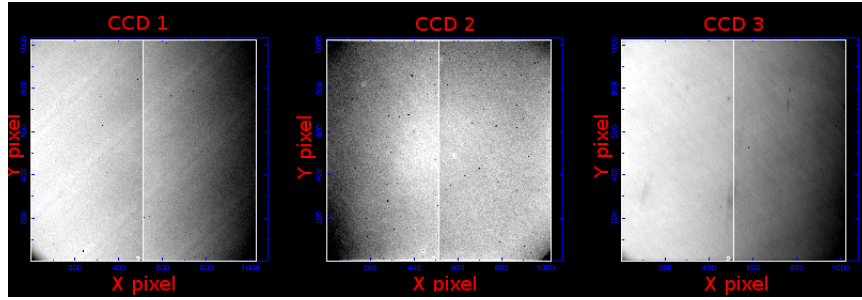
Figure 4.3: The median combined master flats used for each filter from red (left panel) to blue (right panel).

pixel at the division between the two sides of the chip.

Following the creation of the master flats and biases, the next step in the ULTRACAM reduction pipeline is to create an aperture file. This works by loading in an example science frame and using the cursor to select the locations of the stars on the chip. The ULTRACAM software then refines the location of the star from the cursor selection by fitting a Gaussian to the profile within the aperture. The software is built to deal with as many apertures as the user selects, which can be linked between stars. The linking of apertures was primarily designed for white dwarf binaries which can essentially disappear during eclipse, but is also useful if a faint star is obscured by cloud passing overhead.

In the case of these observations, I placed apertures around each of the stars visible in the field of view (Fig. 4.4). The two faintest stars (apertures 5 and 6, Fig. 4.4) were linked to nearby bright objects in case the faint stars were obscured during a drop in transmission. Aperture 2 was selected as a reference star (green aperture in Fig. 4.4), when experimenting with variable apertures. Variable apertures adjust to the seeing, and were experimented with. However, as these were defocussed observations, seeing variations were not expected to be important and I was concerned that variable apertures across the 3 CCDs could introduce systematics that differed between the channels. For these reasons I opted to used fixed apertures for the final reductions.

To test the effects of the aperture radius on the photometry, a number of reductions were performed with different aperture radii. An aperture with a radius of 18 pixels was found to optimise the signal to noise ratio in the resulting photometry. The background annulus was selected to encompass enough area to get a good estimate of the sky level while not being too close to the target aperture. The selected sky annulus had an inner radius of 23 pixels and outer radius of 27 pixels.

Following these reductions, I constructed differential light curves by dividing

102
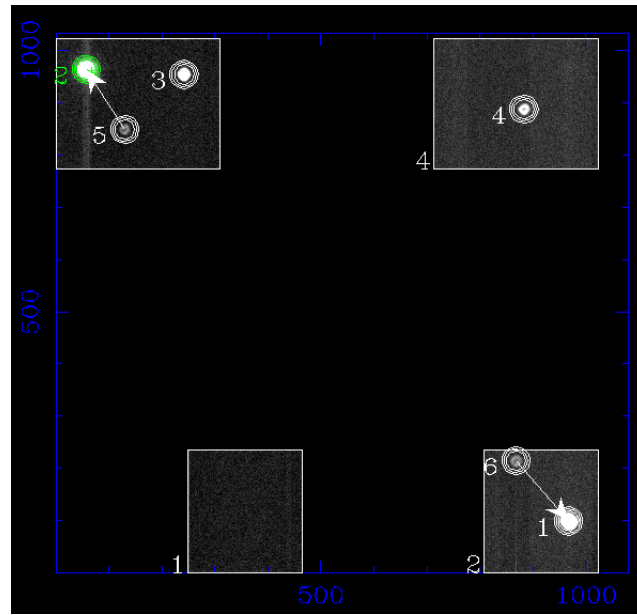
Figure 4.4: Aperture positions used in the reduction of the data. This is an example frame taken with the g' filter. Aperture 1 is the target and aperture 2 is the bright comparison object. The faint stars (apertures 5 and 6) are linked to nearby bright stars. This is done to avoid losing the position of the star in case the transmission decreases, as described in the text. Aperture 2 is green as it has been flagged as a reference star, as described in the text. This figure also displays the 4 windows used during these observations.

the target's light curve by a comparison star's light curve. Combinations of fainter stars in the field were tested as comparison stars but led to more scatter in the differential light curve than the use of the single, bright comparison.

These light curves showed dips in transmission associated with cloud passing overhead and the affected data were removed before analysis using an iterative process. The running median was calculated over a sliding box of 600 data points in the red and green arms and 60 points in the blue. The running median was then subtracted from the raw data and σ clipping was performed on the residuals using the median absolute deviation (MAD). Points lying at $\geq 6\sigma$ from the MAD were clipped. This process was then repeated with a smaller sliding box of 400 data points in the red and green and 40 in the blue with a final σ cut at 5σ from the MAD. The σ clipping was performed on both the target and comparison star independently and only those frames that passed the sigma cut for both were kept, resulting in the removal of 20 % of the data. The outcomes of this process are shown in Fig. 4.5.

After the σ clipping and binning of the red and green channels to the cadence of the blue channel, there were 2494 data points in each of the three light curves at a cadence of 7.9 s giving us excellent sampling even after the cleaning of the data.

4.3.2 Light curve fitting with analytic model

I initially fitted the differential light curves with analytic limb-darkened transit light curves (Mandel & Agol, 2002) using a Markov chain Monte Carlo (MCMC) algorithm, implemented through the EMCEE (Foreman-Mackey et al., 2013) PYTHON package. A quadratic limb darkening law was used and fit simultaneously with a long time-scale trend. In the red and green channels, this trend was fit with a second-order polynomial whilst in the blue it was fit as a function of airmass since it appeared to be related to extinction. In order to fit the airmass term in the blue channel, an extinction coefficient that varied quadratically in time was used, which replicated the trend well. This trend was of the form

$$\log_{10}(F) = a + (b + c\Delta t + d\Delta t^2)X \qquad (4.1)$$

where F is the flux, Δt is time (defined as Time(MJD) - 56177.9), a, b, c, d are the coefficients of the model, and X is the airmass. This is the same function as used by Rostron (2015) to fit ULTRACAM observations of WASP-33b. An extinction coefficient that varied linearly in time was also tested but could not fit the sharp downturn in the blue light curve at the end of the night (Fig. 4.7). Fig. 4.6

104
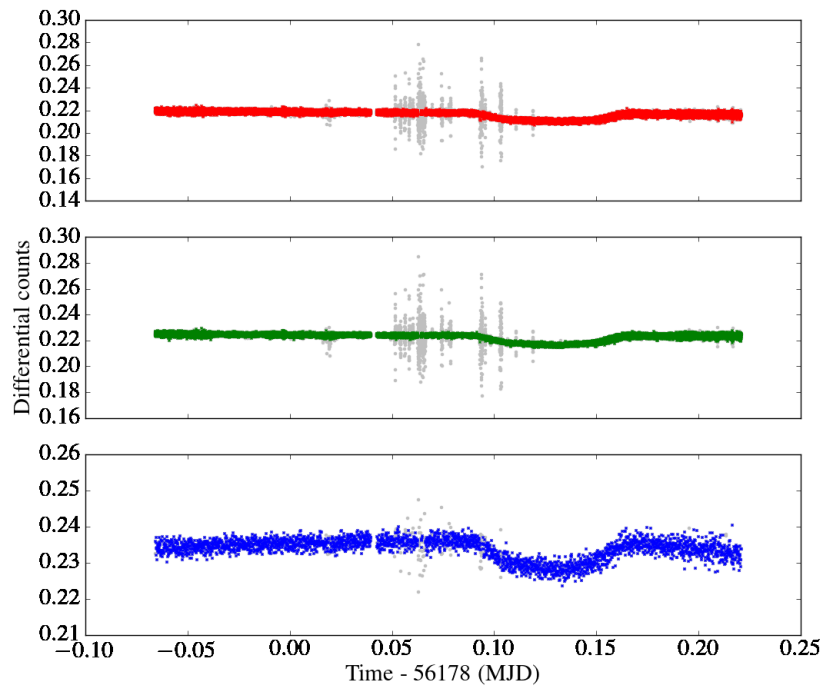
Figure 4.5: The result of the iterative cloud removal process as described in the text. In each panel the coloured points indicate the data remaining after the removal of cloudy frames and the grey points show the data before any cloud removal. The three panels correspond to the red (top), green (middle) and blue (bottom) channels.

105

shows the different airmass functions that were tested. This figures demonstrates that a quadratically varying in time extinction coefficient was necessary to fit the flux-airmass relation in the blue channel.

When fitting the differential light curves, the scaled semi major axis, a/R_*, inclination of the orbit and the time of mid transit, T_c, were tied across the three light curves. The parameters that were fit individually in each of the channels were the ratio of planet to star radius R_P/R_*, the second limb darkening coefficient $u2$, and the parameters defining the long time-scale trend. The first limb darkening coefficient, $u1$, was held fixed in the fitting due to the degeneracy between the limb darkening parameters, as described in Chapter 1, which can affect the light curve solution (e.g. Southworth, 2008). The limb darkening coefficients were taken from the tables of Claret & Bloemen (2011). Uniform priors were adopted for all the model parameters, with the MCMC walkers started at the values from Hébrard et al. (2013).

The resulting fits of the analytic limb-darkened transit light curves are shown in Fig. 4.7. The strongest residual across the whole light curve in all three wavelengths is seen during transit and is consistent across the three bands. This residual is akin to the planet occulting areas of stellar activity. I considered the possibility that this anomaly could have been associated with the use of incorrect limb darkening coefficients so I ran a fit without holding $u1$ fixed but found that this could not replicate the feature.

4.3.3 Fitting of star spot model

The presence of the in-transit anomaly motivated the use of star spot models. Star spot occultations have been seen in the transit light curves of several planets, such as HD 189733b (Pont et al., 2007; Sing et al., 2011a), TrES-1b (Dittmann et al., 2009; Rabus et al., 2009) and CoRoT-2b (Huber et al., 2009; Wolter et al., 2009; Silva-Valio et al., 2010).

If a planet occults a spot (a region cooler than the surrounding photosphere) it blocks less of the stellar flux than compared with its transit across the hotter pristine stellar disc. This results in a bump during transit and therefore a smaller derived planetary radius. Star spot activity is not unexpected for WASP-52 since Hébrard et al. (2013) found modulations in its light curve and chromospheric emission peaks in the Ca II H+K lines. Using these modulations they calculated the rotation period of WASP-52 to be 16.4 d.

Spot crossing events have been modelled with a variety of software, such as PRISM (Tregloan-Reed et al., 2013), SOAP-T (Oshagh et al., 2013) and SPOTROD
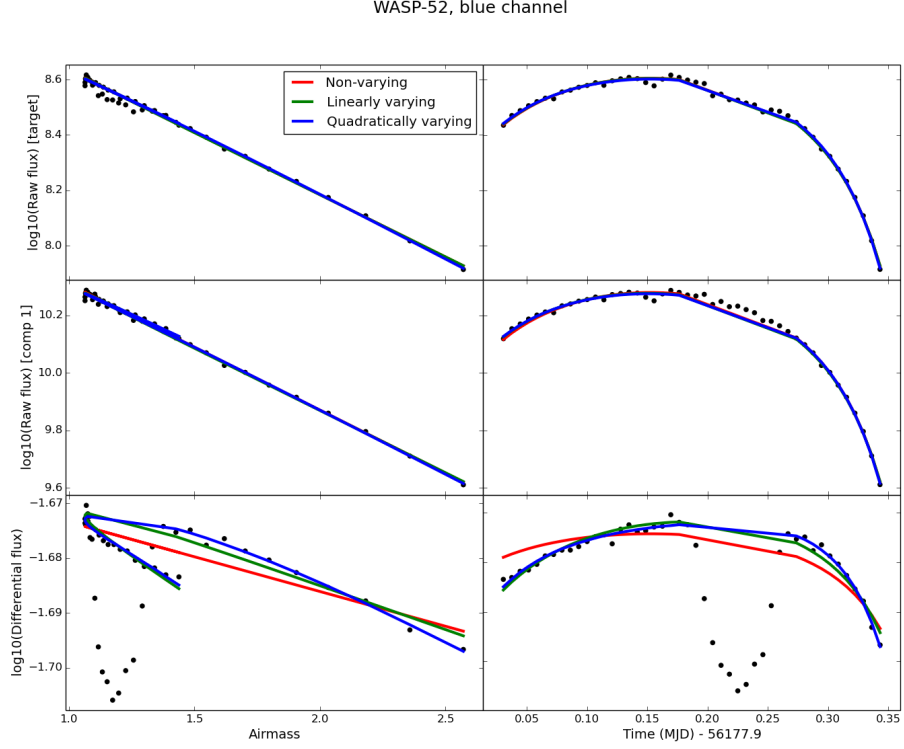
Figure 4.6: These plots show different time-varying extinction coefficients fitted to the blue channel's data. The left column shows the logarithm of the target's flux (top panel), comparison's flux (middle panel) and differential flux (bottom panel) against airmass. The differential flux, in particular, shows the hook-like feature indicating the time-varying nature of the extinction coefficient. If the extinction coefficient was constant in time, the flux would follow airmass along a straight line. The solid, colour lines in all of these plots indicate fits to the out of transit data (after masking the in-transit data) with a non-varying extinction coefficient (red line), linearly varying extinction coefficient (green line) and quadratically varying extinction coefficient (blue line). The right hand column shows the same plots but against time, not airmass.
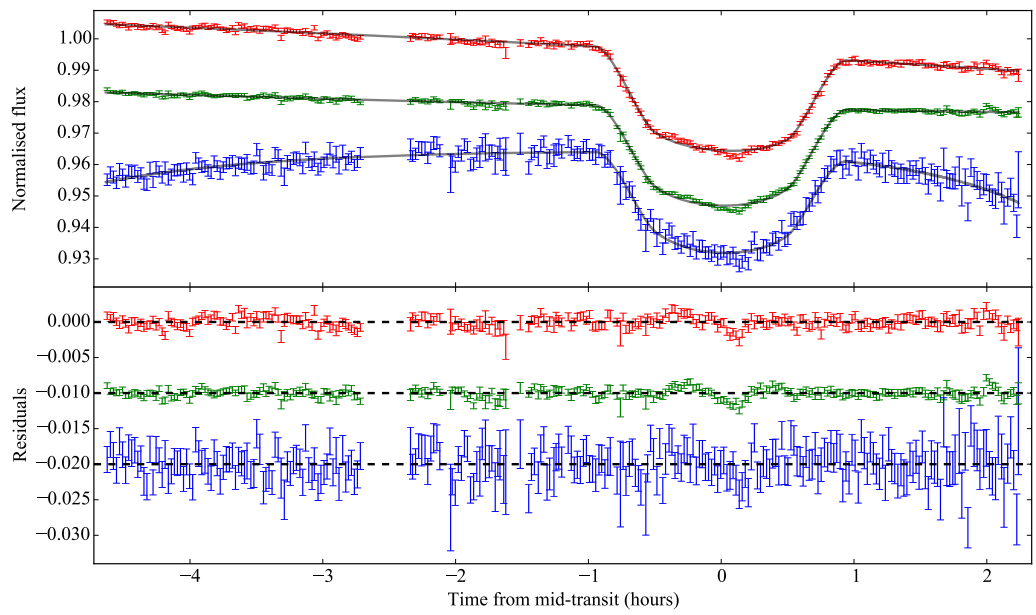
Figure 4.7: MCMC fits of analytic quadratic limb-darkened transit light curves (Mandel & Agol, 2002) revealing the presence of an in-transit anomaly. The upper panel shows the fits to each of the three wavelengths; Na I (red), g' (green) and u' (blue), which are offset for clarity. The lower panel shows the residuals from these fits, which are again offset for clarity.

(Béky et al., 2014). For this analysis I used SPOTROD to construct quadratically limb-darkened transit light curves that included the effects of spot crossings. I incorporated SPOTROD into an MCMC framework and fitted the light curves simultaneously and with the same long time-scale trends as before. I chose to use SPOTROD due to the speed of its integration, which uses polar coordinates in the projection plane. The integration with respect to the polar coordinate is done analytically so that only the integration with respect to the radial coordinate needs to be performed numerically. To calculate the projection of the planet on the stellar surface, SPOTROD calculates the arrays of planar orbital elements ξ and η, using the formalism of Pál (2009), and assumes the same limb darkening law for the spot as for the star.

Since my study of WASP-52b a recent paper introduced PYTRANSPOT, a code to model multi-wavelength light curves that are affected by spots (Juvan et al., 2018). This improves on the pre-existing software for modelling spot-affected light curves, which are normally only set-up to deal with a single light curve.

When using my own fitting routines, the system parameters were again fit across the three light curves simultaneously (although this time fitting for impact parameter, b, rather than the inclination, as required by SPOTROD) and with the addition of the parameters defining the spots. The fitting of one spot was tested but was unable to fit both bumps on either side of the transit mid-point, therefore two spots were used in further analysis. The parameters defining each spot were the longitude, latitude, radius ratio of spot to star, and ratio of the spot flux to stellar flux (with 1 being a spot with the same flux as the pristine photosphere and 0 being a spot with zero flux). I held $u1$ fixed as before but now also put Gaussian priors on $u2$ with means equal to the values from Claret & Bloemen (2011) and standard deviations from the propagated errors in the effective temperature and surface gravity of the host star. This prior was necessary as the limb darkening and spot models can play off each other in trying to fit the transit shape.

The MCMC was initiated with the system parameters equal to those in Hébrard et al. (2013) and was run for 10 000 steps in burn in and another 10 000 steps in the production run. There were 31 fitted parameters with 124 walkers. For the spot starting parameters, I started one spot in the western hemisphere and one in the eastern hemisphere (as this configuration was necessary to explain the features either side of the transit, Fig. 4.7), and both at latitudes near the transit chord. I initiated them with radii 10% of the stellar radii and with contrasts of 1 (i.e. the same flux as the surrounding photosphere). This flux ratio was selected so as not to impose a spot on the star if the model did not need it. After the burn in

phase the error bars in the data points were rescaled to give a reduced χ^2 of unity.

After the first MCMC chain, a second MCMC was run but this time with the parameters that were tied across channels fixed to the results from the first run (a/R_*, b, T_c and the spot sizes and positions). Correlations with these parameters cause R_P/R_* to move up and down together across the three wavelengths, contributing to the uncertainty in the absolute planetary radius in each of the bands. Since we are concerned with the shape of the transmission spectrum, we are interested in the relative radii between the bands and not the absolute planetary radius, thus motivating the second run of the MCMC with fixed system parameters.

I present the best fitting spot model in Fig. 4.8, after the second MCMC run, with the results in Table 4.1 and transmission spectrum in Fig. 4.16 (blue squares)[2].

With the sizes and contrasts calculated from SPOTROD, I was able to create a schematic of the stellar surface (Fig. 4.9) and consider what filling factor would reproduce the derived contrasts (section 4.4.1 and Fig. 4.10). Fig. 4.9 displays the large regions of stellar activity along the transit chord. The second spot crossing event composed of a smaller region of higher contrast (0.2 in the g' band, Table 4.1).

4.3.4 Fitting of bright region model

As an alternative interpretation, the in-transit anomaly was also modelled as a bright feature analogous to solar faculae. Solar faculae are bright regions on the solar photosphere which display limb brightening behaviour (e.g. Unruh et al., 1999). They are often co-spatial with sunspots but not perfectly so (e.g. Haywood et al., 2016).

The effects of occultations of bright regions in exoplanet transits have been discussed by Oshagh et al. (2014) and could lead to an observable anomaly in transit data. There has not yet been any conclusive evidence of a facula occultation in a transit light curve although Mohler-Fischer et al. (2013) found evidence for a hotspot in GROND photometry of HATS-2b. They detected a bright feature in the Sloan g band, which covered the Ca II lines, that was consistent with a chromospheric plage occultation.

SPOTROD was also used to model the facula scenario but instead of two individual spots with flux ratios < 1, a single feature was modelled with a flux ratio of > 1. This fitting method produced the fits shown in Fig. 4.11, with the results in Table 4.1. This model was able to reproduce the in-transit anomaly with an equally

---

[2]Note: x positions of the spots are defined to take values between ± 1 which are analogous to longitudes of $\pm 90°$ whilst y positions take values between 1 and 0 corresponding to latitudes of $\pm 90°$. Spot contrast is defined as the ratio of spot flux to stellar flux so a value of unity gives a spot of equal brightness to the surrounding photosphere.
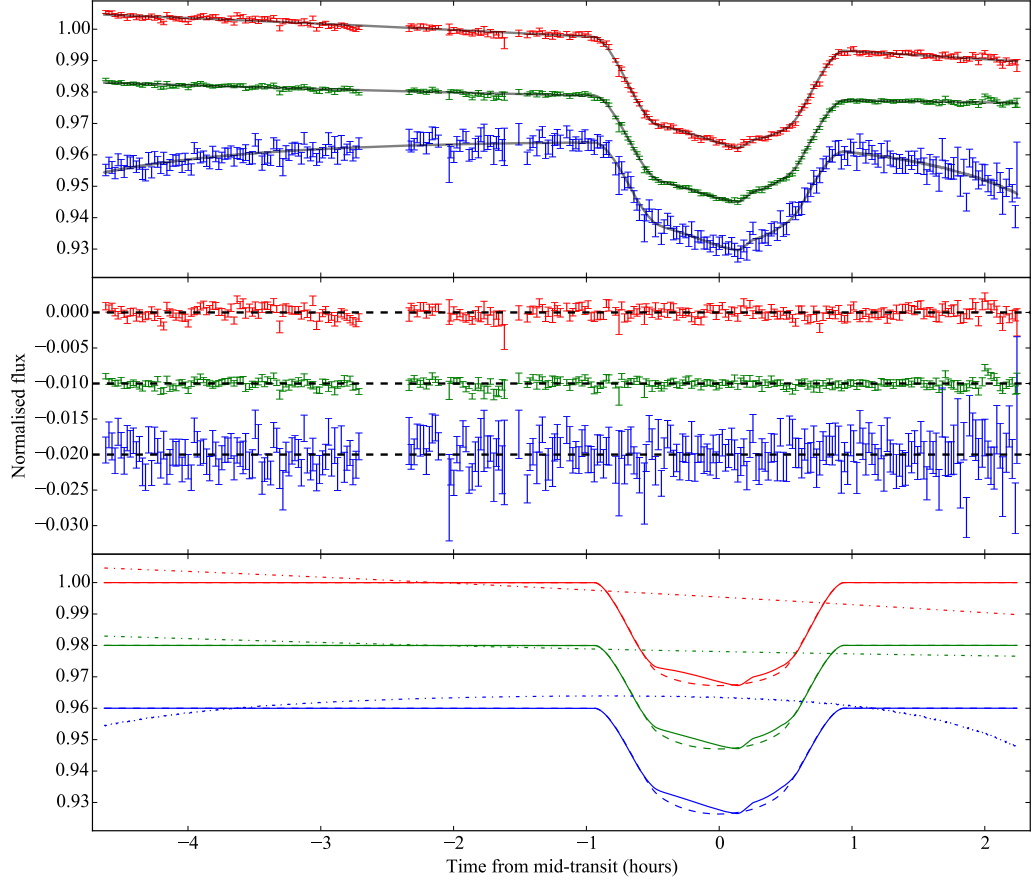
Figure 4.8: MCMC fit of spot model with the occultation of two areas of stellar activity at temperatures lower than the pristine photosphere. The top panel shows the best-fitting model overlaid on the Na I (red), g' (green) and u' (blue) light curves which are offset for clarity. The middle panel shows the residuals to these fits and the bottom panel shows the constituents of each model, which are again offset for clarity. In the bottom panel, the solid line shows the best fitting spot model with the dashed line representing how the model would appear in the absence of spots. The dot-dashed line shows the long time-scale term fitted by a second-order polynomial in each of the Na I and g' bands and as a function of airmass in the u' band.
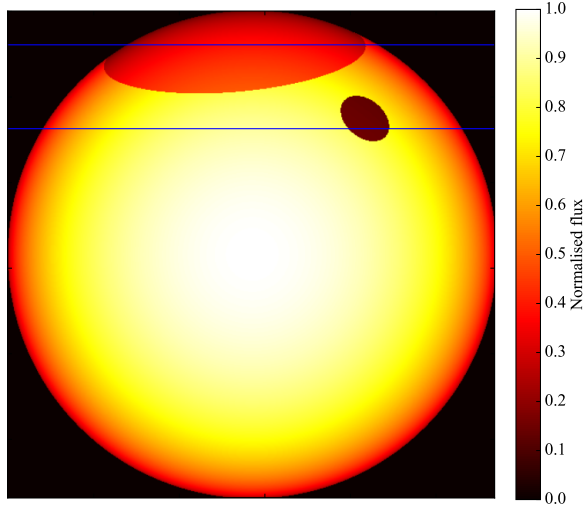
Figure 4.9: Schematic illustration of how the stellar surface may appear with the active regions derived from the fitting of the spot model and contrasts consistent with those in the g' band. The blue lines indicate the planet's transit chord. Spot #1 is the larger spot at higher latitudes and spot #2 is the smaller, darker spot.
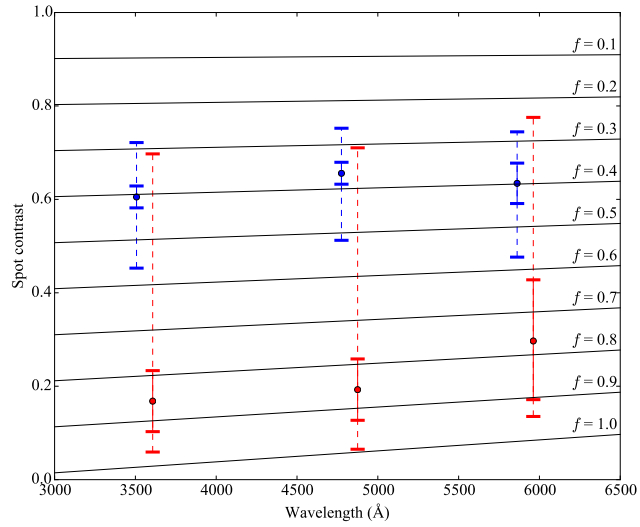
Figure 4.10: Plot of the derived spot contrasts for both spot #1 (blue) and spot #2 (red) in each of the three wavelengths, which are offset in the x-axis by -50 and $+50$ Å for clarity. The dashed error bars show the error in the absolute spot contrasts whilst the solid error bars give the relative uncertainties between the bands for my best-fitting system parameters. Over-plotted are the calculated contrasts for a spot with $\Delta T = 1500$ K (consistent with predictions for K stars) with varying filling factors f (solid black lines).
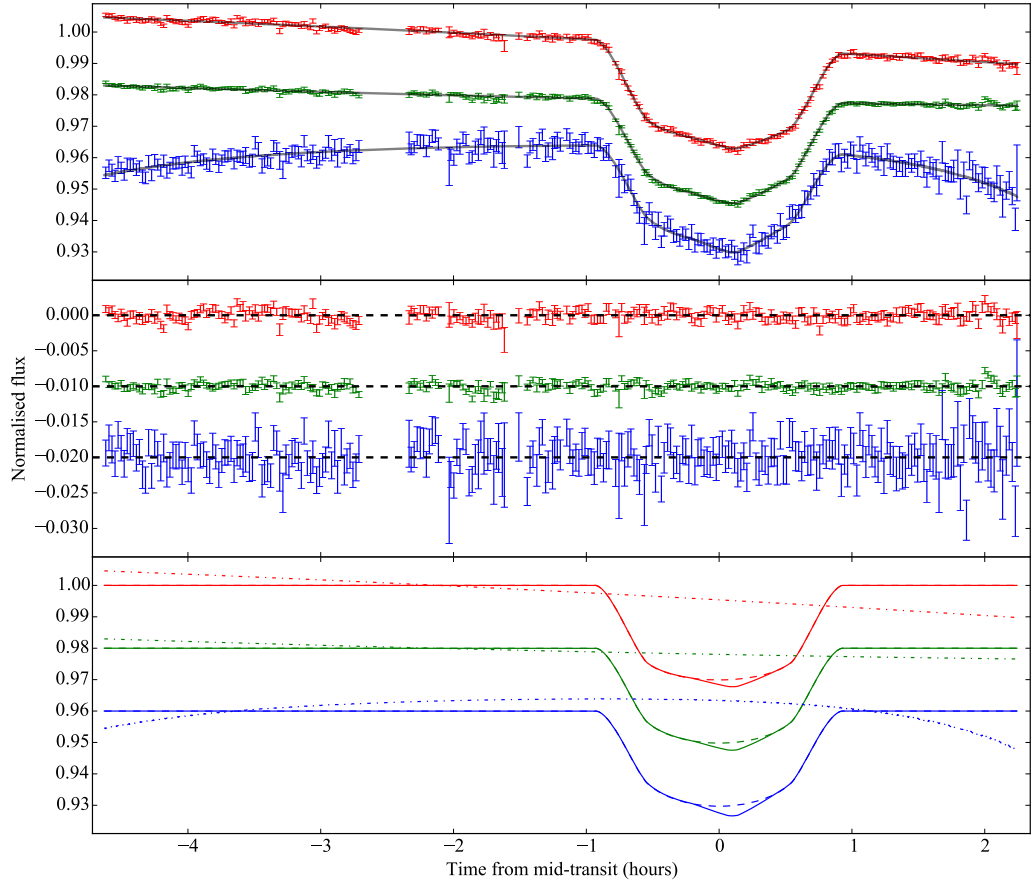
Figure 4.11: The same representation as in Fig. 4.8 but this time showing the result of the fitting of a bright spot, analogous to a facula, on the stellar surface. This bright spot leads to a dip in the in-transit data (bottom panel, solid lines) as compared to a transit free of such activity (bottom panel, dashed line).

good fit as the two-spot model but with six fewer parameters. The transmission spectrum resulting from the facula model is also shown in Fig. 4.16 (red triangles). In contrast to the flat transmission spectrum resulting from the fitting of spots, the fitting of a facula led to a slope in the planetary radius increasing towards the red (Fig. 4.16).

Similar to Fig. 4.9, I created a schematic of how the stellar surface would appear with the results from the facula model. This is shown in Fig. 4.12 and indicates the extent of the active region occulted by the planet.

Table 4.1: Results from fitting the transit light curves with both spots and faculae to model the stellar activity.

Parameter	Symbol	u'	g'	Na I
Two-spot model				
Time of mid-transit (MJD)[a]	T_c	-	56178.12742 $\pm$0.00005	-
Scaled semi-major axis[a]	a/R_*	-	6.99 $\pm$0.04	-
Impact parameter[a]	b	-	0.656 $^{+0.006}_{-0.007}$	-
Radius ratio	R_P/R_*	0.1759 $^{+0.0005}_{-0.0004}$	0.1751 $\pm$0.0004	0.1757 $\pm$0.0008
Limb darkening coeff.[b]	u_1	0.643	0.529	0.381
Limb darkening coeff.	u_2	0.169 $\pm$0.008	0.233 $\pm$0.008	0.290 $^{+0.009}_{-0.010}$
Spot 1 x position[a]	x_1	-	-0.19 $\pm$0.03	-
Spot 1 y position[a]	y_1	-	0.91 $^{+0.03}_{-0.04}$	-
Spot 1 radius ratio[a]	r_1/R_*	-	0.36 $^{+0.04}_{-0.03}$	-
Spot 1 contrast	ρ_1	0.61 $\pm$0.02	0.66 $\pm$0.02	0.63 $\pm$0.04
Spot 2 x position[a]	x_2	-	0.38 $^{+0.02}_{-0.007}$	-
Spot 2 y position[a]	y_2	-	0.69 $^{+0.02}_{-0.03}$	-
Spot 2 radius ratio[a]	r_2/R_*	-	0.07 $^{+0.08}_{-0.006}$	-
Spot 2 contrast	ρ_2	0.17 $\pm$0.07	0.19 $\pm$0.07	0.30 $\pm$0.13
Facula model				
Time of mid-transit (MJD)[a]	T_c	-	56178.12740 $\pm$0.00004	-
Scaled semi-major axis[a]	a/R_*	-	7.23 $\pm$0.05	-
Impact parameter[a]	b	-	0.593 $^{+0.008}_{-0.009}$	-
Radius ratio	R_P/R_*	0.1632 $\pm$0.0003	0.1641 $\pm$ 0.0003	0.1657 $\pm$0.0006
Limb darkening coeff.[b]	u_1	0.643	0.529	0.381
Limb darkening coeff.	u_2	0.161 $\pm$0.009	0.239 $\pm$0.009	0.29 $\pm$ 0.01
Facula x position[a]	x_1	-	0.123 $^{+0.007}_{-0.008}$	-
Facula y position[a]	y_1	-	0.62 $\pm$0.03	-
Facula radius ratio[a]	r/R_*	-	0.20 $^{+0.01}_{-0.02}$	-
Facula contrast	ρ	1.119 $\pm$0.007	1.089 $\pm$0.008	1.085 $\pm$ 0.014

[a] Parameter is shared and fit across channels with the result quoted in the g' column only and is held fixed at this value for the second MCMC run. The errors on all other parameters are those after the second MCMC run.
[b] Parameter not fit by model.
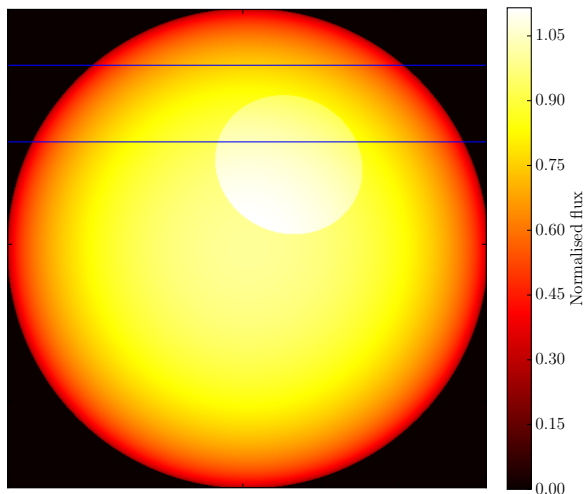
Figure 4.12: Schematic illustration of how the stellar surface may appear with the active region derived from the fitting of the facula model and contrasts consistent with those in the g' band. The blue lines indicate the planet's transit chord.

4.4 Discussion

4.4.1 Spots or faculae?

It is difficult to distinguish between the spot and facula models of Sections 4.3.3 and 4.3.4 using the quality of the fits alone. Application of the Bayesian information criterion (see Chapter 3) favours the facula model because of the similar χ^2 and six fewer free parameters, with a value of > 10 lower than that of the spot model. However, our prior knowledge of transit light curves tends to favour the spot model because spots are more commonly detected (e.g. Pont et al., 2007; Sing et al., 2011a; Mancini et al., 2013; Tregloan-Reed et al., 2013) than faculae or bright regions (Mohler-Fischer et al., 2013).

A second approach to distinguish between these models is to compare my fitted planet radius with that from independent studies. It can be seen in the model light curves of Figs 4.8 and 4.11 that the spot and facula models imply different underlying transit depths. In the spot model the anomaly is treated as two bumps (bottom panel of Fig. 4.8), and so the underlying transit is deeper than in the facula model, where the anomaly is treated as a single dip (Fig. 4.11). This difference can be seen in the fitted planet/star radius ratios (Table 4.1) where the spot model implies a significantly larger planet radius than the value from the facula model. In WASP-52b's discovery paper, Hébrard et al. (2013) measured the planet/star radius ratio using six individual transits, and found it to be 0.1646 ± 0.0012. This is

consistent with my value from the facula model (within $1\,\sigma$), and inconsistent with my value from the spot model ($> 5\,\sigma$ discrepancy).

Since my study of WASP-52b, Mancini et al. (2017) published a further eight transits of this planet. Spots were clearly detected in five of these light curves and their derived system parameters accounted for these spots. Again, I found that the planet/star radius ratio was consistent with my facula model (within $1\,\sigma$) and inconsistent with my spot model ($> 7\,\sigma$ discrepancy). Furthermore, both Louden et al. (2017) and Chen et al. (2017) present R_P/R_* values within 1 and $2\,\sigma$ of my facula model, and > 3 and $> 6\,\sigma$ discrepant with my spot model respectively.

I also considered the possibility that unoccluded spots could be the cause of the discrepant R_P/R_* values I derived from the spot model but find this would require a total dimming of the star of 12%, which I considered unlikely. This is discussed in more detail in Section 4.4.3.

I concluded that the comparison of my planet/star radius ratios with independent studies strongly favours the presence of a facula in my observed transit. Nevertheless, in the sections following, I discuss the implications of both models.

Spot properties

The best-fitting spot model indicates two large regions of stellar activity (Fig. 4.9) which is consistent with the majority of the in-transit data being activity affected (Fig. 4.8).

Although spot modelling codes, including SPOTROD, fit spots as circular areas on the stellar surface, in reality these may be areas of several small spots which could be arranged in more complex configurations on the stellar surface. If the features in my light curves really were single spots with temperatures consistent with those expected for K stars, I would expect to see large amplitude bumps and a strong colour dependence. When considering the spots resulting from my modelling, however, I believe that the spots are more accurately interpreted as active regions, with spots and pristine photosphere contained within, instead of two large individual spots (Fig. 4.9).

When considering these complexes of smaller star spots, it is useful to consider the area ratio of spots to uncontaminated photosphere within each of these regions, which I define here as the filling factor, f. This quantity can be related to the spot's contrast, ρ, through

$$\rho = \frac{f F_\bullet(\lambda) + (1 - f) F_*(\lambda)}{F_*(\lambda)} \tag{4.2}$$

where $F_\bullet$ is the flux of the spot and F_* is the flux of the star.

Berdyugina (2005) plotted the observed temperature differences of spots for several different stellar effective temperatures. For K stars, the spot temperature differences (ΔT) lie in the range $1250 - 1500\,K$. To calculate $F_\bullet$ and F_*, I used ATLAS9 (Kurucz, 1993) stellar atmosphere models of 3500 and $5000\,K$ respectively, and convolved these with the filter responses shown in Fig. 4.1.

Because of the degeneracy between spot size and contrast, the spot contrasts have large error bars in Fig. 4.10. However, by fixing the spot sizes and positions at the best-fitting values from the first MCMC run, the second run gave the relative uncertainty in the spot contrasts across the bands rather than the absolute uncertainty, as with the planetary radii. When considering the relative spot contrasts (Fig. 4.10, solid error bars) it can be seen that the contrasts and colour dependences of these active regions are consistent with a ΔT of $1500\,K$, given a single filling factor f.

For the smaller spot (Fig. 4.10, red error bars) I infer a higher filling factor (approximately 0.85), and for the larger spot (Fig. 4.10, blue error bars) a lower filling factor (approximately 0.4). The larger spot can be understood as a large active region on the stellar surface with a relatively low density of smaller dark spots. The smaller spot, however, has a high density of smaller dark spots leading to the greater contrast.

Furthermore, since my study, Louden et al. (2017) found that the spot in their transit light curve of WASP-52b was also more complex than a single feature and was instead likely caused by an active region, much like I see here.

Facula properties

As discussed in this section, the occultation of a bright region, analogous to a solar facula, is my favoured interpretation for these data. Unlike spots, the flux contrasts of faculae increase at the stellar limb as they display limb brightening behaviour (Unruh et al., 1999) and so the assumption that the modelled spots follow the same limb darkening as the star does not hold in this case. Therefore equation 4.2 cannot be used to calculate the filling factor of the facula. However, it is likely that what we are seeing in this scenario is a cluster of bright regions along the stellar surface rather than the single large feature depicted in Fig. 4.12.

The flux contrasts of the facula are given in Table 4.1, which indicate the hotspot is $\sim 10\%$ brighter in all three bands. Solar faculae are also able to display such contrasts, however, such contrasts are only observed at the solar limb with a high viewing angle (e.g. Unruh et al., 1999; Ahern & Chapman, 2000). At the

viewing angle of the area of facula occulted by the planet (Fig. 4.12), we would only expect to see a $\sim 5\%$ contrast if this were a Solar facula (e.g. Ahern & Chapman, 2000). However, since my study of WASP-52, Norris et al. (2017) presented faculae contrasts for early G type stars with a range of magnetic field strengths. They found that the contrasts of faculae could be as large as 40%, indicating that faculae contrasts can differ from the solar case. Therefore my derived faculae contrasts seem plausible (Table 4.1).

4.4.2   Measuring the residual red noise

In order to measure the systematic red noise in the residuals, I observed how the root mean square (RMS) of the residuals varied with binning of the data. Fig. 4.13 shows the RMS for a number of different bin sizes. The line overlaid corresponds to pure, Gaussian, white noise with a gradient equal to $1/\sqrt{N}$, where N is the number of points per bin. The factor β quantifies how the actual binning gradient deviates from pure white noise (Winn et al., 2008). The β factors found here show that the residuals do deviate from pure white noise but with amplitudes of only 0.5 mmag.

Note that while the noise on a timescale of one hour is $\sim 1.5 \times 10^{-3}$ in the blue channel, as compared to $\sim 0.8 \times 10^{-3}$ and 0.7×10^{-3} in the red and green channels respectively (Fig. 4.13), the error in R_{P}/R_* is actually greatest in the red channel (Fig. 4.16 and Table 4.1). The errors in R_{P}/R_* are determined by the uncertainty in the transit depth, which correlates with the spot contrast and the limb darkening. I chose to place priors on $u2$ to prevent the limb darkening modelling the effects of the spots. In the Na I filter, the priors were overly constraining, leading to larger errors in the spot contrast in the red channel than the other two channels (Table 4.1), as the spot contrast tried to compensate for the prior on $u2$. As a result, this led to the larger error in R_{P}/R_* in the red channel, which could be overestimated due to the prior on $u2$. This highlights the effect that limb darkening priors can have on the transit parameters. Therefore in Chapters 5 and 6 I run fits with both constrained and unconstrained limb darkening.

4.4.3   Unocculted spots

Just as the effects of occulted spots or bright regions must be taken into account when calculating the planetary radius, so must the effects of those regions which are not seen in the transit. As the photometric modulation of Sun-like stars with high magnetic activity is dominated by dark spots (e.g. Shapiro et al., 2014), I only consider the effect of unocculted spots on my derived transit depths. Unocculted
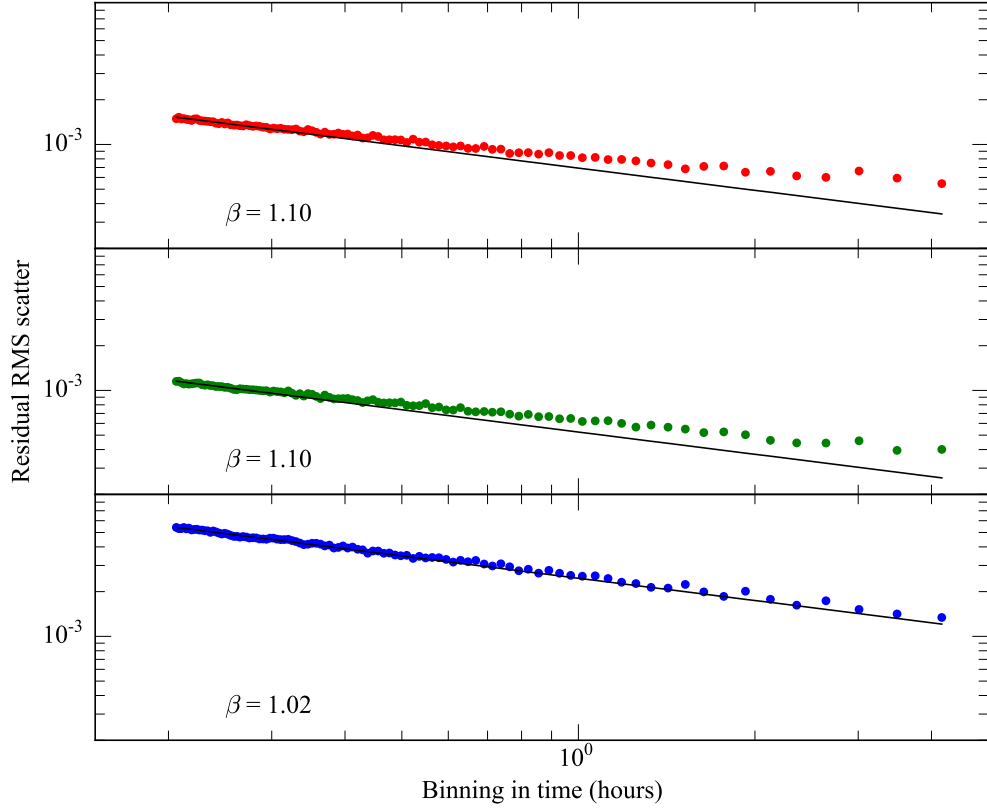
Figure 4.13: The residual rms scatter once the spot model has been subtracted for Na I (top panel), g' (middle panel) and u' (bottom panel). On a timescale of one hour, the rms values are $\sim 0.8 \times 10^{-3}$, 0.7×10^{-3} and 1.5×10^{-3} for the red, green and blue channels respectively. The error bars on the data points were rescaled to give a $\chi^2_\nu = 1$. β quantifies how the measured noise deviates from pure Gaussian white noise (solid black lines).

spots cause the observed stellar flux to decrease, increasing the transit depth as compared to a spot-free surface. This effect needs to be taken into account with a wavelength-dependent depth correction.

Following the formalism of Sing et al. (2011a), the variation in transit depth, $\Delta d/d$, can be related to the fractional decrease of the stellar flux due to unocculted spots at a reference wavelength λ_0, $\Delta f(\lambda_0, t)$, the flux of the pristine stellar disc, $F_\lambda^{T_{\text{star}}}$, and the flux of the spot, $F_\lambda^{T_{\text{spot}}}$, through

$$\frac{\Delta d}{d} = \Delta f(\lambda_0, t) \left(1 - \frac{F_\lambda^{T_{\text{spot}}}}{F_\lambda^{T_{\text{star}}}}\right) / \left(1 - \frac{F_{\lambda_0}^{T_{\text{spot}}}}{F_{\lambda_0}^{T_{\text{star}}}}\right) \tag{4.3}$$

leading to a variation in the ratio of planet radius to stellar radius of

$$\Delta(R_{\text{P}}/R_*) \approx \frac{1}{2}\frac{\Delta d}{d}(R_{\text{P}}/R_*) \tag{4.4}$$

I used stellar atmosphere models (ATLAS9; Kurucz, 1993) to generate synthetic spectra of the star and a spot with the maximum temperature contrast of $1500\,\text{K}$ cooler than the surrounding photosphere (Berdyugina, 2005).

If we rearrange the above equations for $\Delta f(\lambda_0, t)$, we can estimate the percentage of total dimming required to bring my derived R_{P}/R_* values into agreement with those of Hébrard et al. (2013). I find that the total dimming must be $> 12\%$ at a reference wavelength of $6000\,\text{Å}$, considerably higher than the 1% amplitude measured by Hébrard et al. (2013). As a result of this, I do not believe that unocculted spots are the cause of the discrepancy in the R_{P}/R_* values resulting from the spot model.

To confirm the absence of such variation, I downloaded all available ASAS-SN (Shappee et al., 2014; Kochanek et al., 2017) photometry for WASP-52, spanning the date range January 2012 – July 2018. I then calculated a Lomb-Scargle periodogram (Lomb, 1976; Scargle, 1982) for these data using code written by James McCormac, which makes use of ASTROPY's `LombScargle` function within PYTHON. This periodogram is shown in Fig. 4.14 (top panel), which peaks at a period of $17.85\,\text{d}$, consistent with the period found in the WASP photometry by Louden et al. (2017). I then phase-folded the ASAS-SN photometry on this period and fitted a sine curve to the variation, finding a peak-to-trough amplitude of 3% (Fig. 4.14, bottom panel). This is far smaller than the 12% necessary to explain the R_{P}/R_* values resulting from my spot modelling.

As a further check, I plotted the cumulative distribution function of all the ASAS-SN photometry (Fig. 4.15). This showed that the maximum fractional devi-
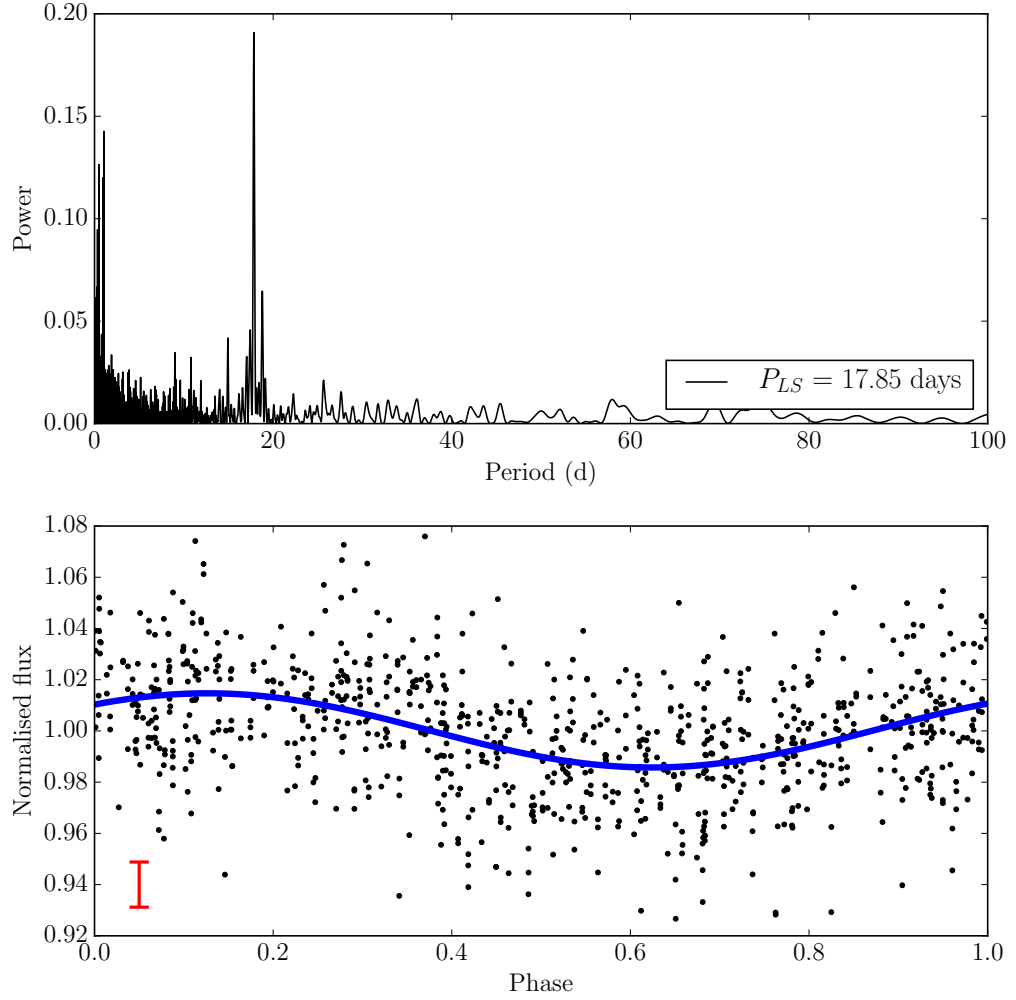
Figure 4.14: Top panel: the Lomb-Scargle periodogram (Lomb, 1976; Scargle, 1982) computed for 6 years of ASAS-SN photometry of WASP-52 (Shappee et al., 2014; Kochanek et al., 2017), showing a peak at a period of 17.85 d, consistent with that found in WASP photometry by Louden et al. (2017). Bottom panel: ASAS-SN photometry of WASP-52 phase-folded on a period of 17.85 d (black points). A sine curve fitted to the variation is shown in blue, which has a peak-to-trough amplitude of 3%. The red errorbar shows the typical uncertainty on the data points (= 0.009).
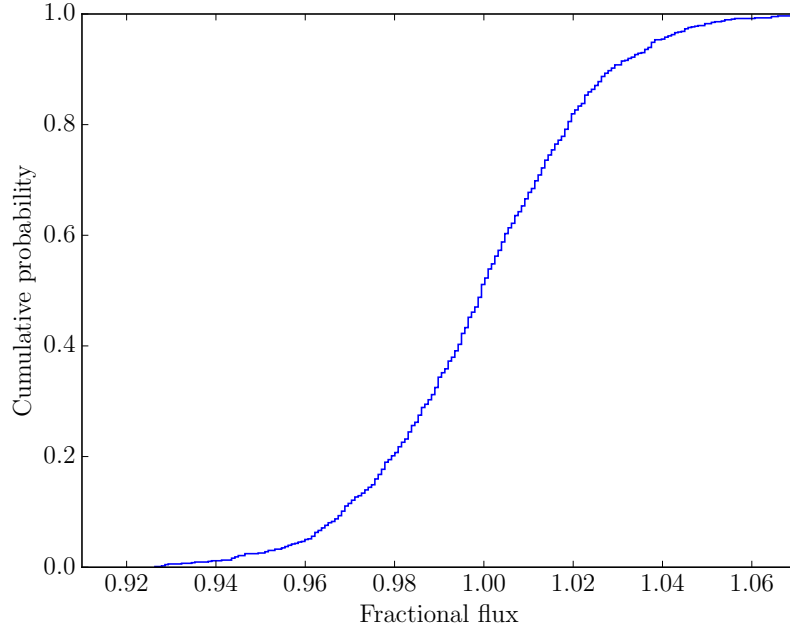
Figure 4.15: A plot of the normalised, cumulative probability of ASAS-SN photometry of WASP-52. The x-axis corresponds to all fluxes divided by the median flux and so gives the fractional deviation from the median.

ation from the median is 0.07. Given that Hébrard et al. (2013) observed transits of WASP-52 at 6 different epochs, including over a year of Super-WASP data, it is reasonable to assume that the R_P/R_* value they report corresponds to the median brightness level of WASP-52. Fig. 4.15 shows that the brightness of WASP-52 recorded by ASAS-SN never drops to 12% fainter than the median, which supports the argument that unocculted spots cannot be the cause of the discrepant R_P/R_* values.

To calculate what effect unocculted spots have on the shape of my transmission spectra, I used a total dimming of 1% at a reference wavelength of 6000 Å to calculate $\Delta(R_P/R_*)$ in each of my three light curves. The difference in $\Delta(R_P/R_*)$ between the red and the blue light curves is within the 1σ error bars and so unocculted spots do not affect the shape of my transmission spectrum (Fig. 4.16) or the conclusions I draw from it.

4.4.4 Transmission spectrum

Fig. 4.16 displays the derived transmission spectrum for WASP-52b. On this plot are the results from modelling the in-transit anomaly both as spots and as a single facula. The errors in the transmission spectra are those after fitting with the wavelength-

independent parameters held fixed at the best fitting values from the first MCMC run (b, a/R_*, T_c, spot/facula sizes and positions). This resulted in errors of less than one atmospheric scale height.

As noted in Chapter 1, in the presence of Rayleigh scattering, the expected slope of the planetary radius as a function of wavelength is given by (Lecavelier Des Etangs et al., 2008a):

$$\frac{dR_{\mathrm{P}}}{d\ln \lambda} = \frac{k}{\mu g}\alpha T,\tag{4.5}$$

where μ is the mean molecular mass of an atmospheric particle (taken to be 2.3 times the mass of a proton), k is the Boltzmann constant, g is the planet's surface gravity, $\alpha = -4$ as expected for Rayleigh scattering, and T I take as the equilibrium temperature.

Using my derived values for the planet's surface gravity (Table 4.2) I am able to rule out Rayleigh scattering in this atmosphere with $> 3\sigma$ confidence (Fig. 4.16). The spot model gives a χ^2 of 14.52 for Rayleigh scattering and 2.00 for a flat line, each with 2 degrees of freedom. Therefore the spot model strongly favours a flat transmission spectrum. For the case of the facula model, the transmission spectrum shows an increase in planetary radius towards the red which is not well fit by either a flat line or Rayleigh scattering. The implied increased radius in the red could indicate the detection of the broad wings of sodium.

An absence of Rayleigh scattering has been seen in several planets to date, such as HAT-P-32b (Gibson et al., 2013b) and HAT-P-19b (Mallonn et al., 2015a). In theses examples, the absence of Rayleigh scattering has been attributed to clouds in the upper atmosphere of the planet. I therefore interpret the lack of broad spectral features in the transmission spectrum of WASP-52b as being caused by clouds in its upper atmosphere.

Since my study, Chen et al. (2017), Louden et al. (2017), and Mancini et al. (2017) have all presented transmission spectra of WASP-52b. All these studies found flat transmission spectra suggesting the presence of clouds. However, Chen et al. (2017) additionally detected the narrow core of the sodium feature in their GTC/OSIRIS transmission spectrum.

4.4.5 Updated system parameters

As outlined in Section 4.4.1, I compared my results to those of Hébrard et al. (2013) (Table 4.2). Prior to my study, Baluev et al. (2015) also presented results for WASP-52b in looking for transit timing variations but the results of the parameters that
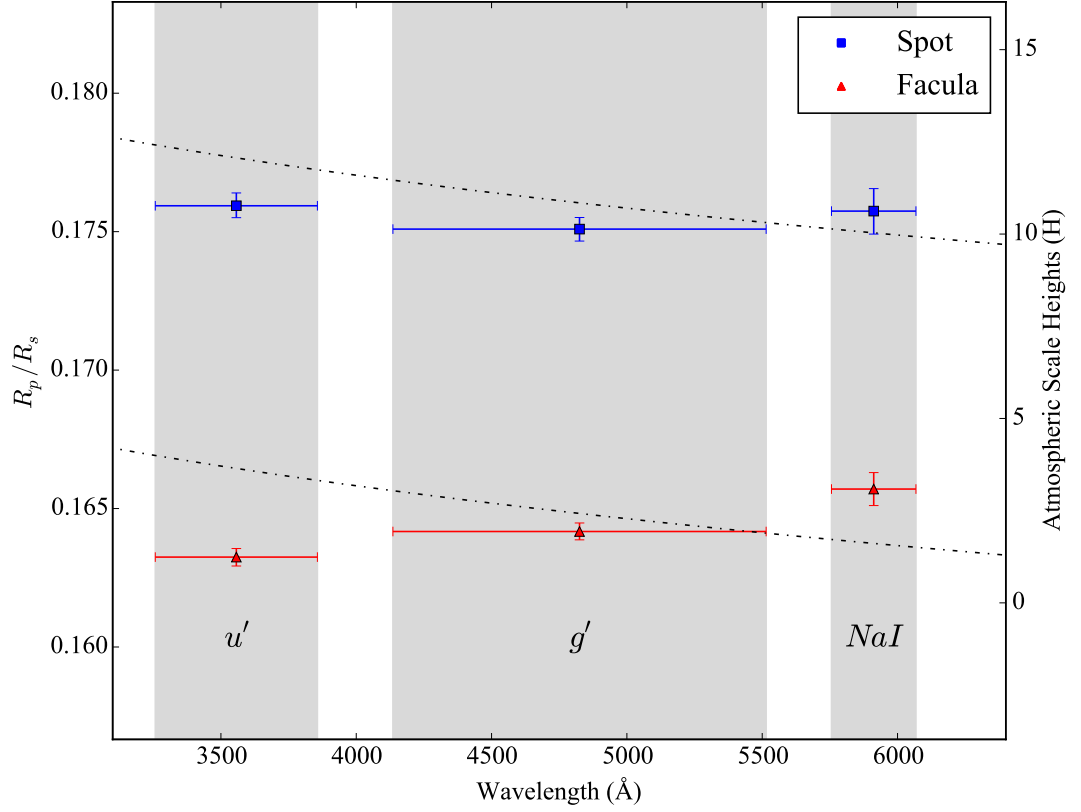
Figure 4.16: Transmission spectrum of WASP-52b. The errors in the x-axis are quoted as the FWHM of the filters used; u', g' and Na I. The results from both the spot (blue squares) and facula (red triangles) models are included. Over plotted is a Rayleigh scattering slope (black dot-dashed line) which is ruled out with $> 3\sigma$ significance. The scale height axis is calculated using the value for WASP-52b's surface gravity from Hébrard et al. (2013). I discuss the uncertainties on these data points in section 4.4.2.

they fit for, R_{P}/R_* and b, are within the 1σ errors of Hébrard et al. (2013).

All the parameters in Table 4.2 marked with [b] were derived from the g'-band transit light curve alone, following the prescription of Seager & Mallén-Ornelas (2003). The planet surface gravity was calculated using the equation from Southworth et al. (2007) which can be derived from the transit light curve but with the addition of the stellar reflex velocity for which I used the value from Hébrard et al. (2013). The parameters that I held fixed during fitting (semi-major axis and period) and those which I do not derive are left blank in Table 4.2. The parameters marked with [c] were derived assuming a stellar radius equal to that of Hébrard et al. (2013).

As discussed in Section 4.4.1, the differing implied transit depths of the spot and facula models result in significantly different system parameters (Table 4.2). When modelled as two spots, I find a transit depth $> 5\sigma$ larger than that of Hébrard et al. (2013), whereas the facula model results in a transit depth consistent with that of Hébrard et al. (2013).

Table 4.2: Comparison between derived system parameters from this work with WASP-52b's discovery paper (Hébrard et al., 2013).

Parameter (units)	Symbol	This work (spots)	This work (facula)	Hébrard et al. (2013)
Orbital period (d)	P	-	-	1.7497798 ± 0.0000012
Semi major axis (au)	a	-	-	0.0272 ± 0.0003
Orbital eccentricity	e	-	-	0 (fixed)
Planetary equilibrium temperature (K)	$T_{\rm P}$	-	-	1315 ± 35
Transit epoch (HJD − 2450000.0) (d)[a]	T_c	6178.62742 ± 0.00005	6178.62740 ± 0.00004	5793.68143 ± 0.00009
Scaled stellar radius[a]	R_*/a	0.1431 ± 0.0009	0.1383 ± 0.0008	0.1355 ± 0.0020
Impact parameter (R_*)[a]	b	$0.656\ ^{+0.006}_{-0.007}$	$0.593^{+0.008}_{-0.009}$	0.60 ± 0.02
Orbital inclination (°)[a]	i	84.62 ± 0.07	85.30 ± 0.08	85.35 ± 0.20
Transit duration (d)[b]	$t_{\rm T}$	0.0777 ± 0.0006	0.0772 ± 0.0006	0.0754 ± 0.0005
Planet/star area ratio[b]	$(R_{\rm p}/R_*)^2$	0.0306 ± 0.0005	0.0270 ± 0.0002	0.0271 ± 0.0004
Stellar density $(\rho_\odot)$[b]	ρ_*	1.50 ± 0.03	1.66 ± 0.03	1.76 ± 0.08
Stellar radius $({\rm R}_\odot)$	R_*	-	-	0.79 ± 0.02
Planet radius $(R_{\rm J})$[c]	$R_{\rm p}$	1.34 ± 0.04	1.26 ± 0.03	1.27 ± 0.03
Stellar reflex velocity (km s^{-1})	K_1	-	-	0.0843 ± 0.0030
Planet surface gravity (cgs)	$\log g_{\rm p}$	2.75 ± 0.02	2.83 ± 0.02	2.81 ± 0.03
Planet density $(\rho_{\rm J})$[c]	$\rho_{\rm p}$	0.383 ± 0.024	0.204 ± 0.009	0.22 ± 0.02
Planet mass $(M_{\rm J})$[c]	$M_{\rm p}$	0.38 ± 0.02	0.41 ± 0.03	0.46 ± 0.02

[a] fitted parameter.

[b] derived from the transit light curve alone.

[c] derived using stellar radius from Hébrard et al. (2013).

4.5 Conclusions

In this chapter I measured the optical transmission spectrum of the highly inflated hot Jupiter WASP-52b.

My multi-wavelength light curves, observed with WHT/ULTRACAM, have demonstrated that 0.5 mmag precision can be achieved with high quality ground-based photometry. This leads to errors in my planetary radii of less than one planetary atmospheric scale height, comparable to *HST* observations (e.g. Nikolov et al., 2014; Sing et al., 2015; Fischer et al., 2016).

My transit light curves revealed regions of stellar activity, which were modelled using SPOTROD (Béky et al., 2014). Although modelled as discrete events, it is likely that the planet in fact transits across latitudes of high stellar activity which may be near-continuous across the transit chord.

I find this activity can be most simply modelled as a bright region akin to solar faculae, which results in system parameters consistent with independent studies. The light curves can also be fit with two regions of dark spots but this requires a planet/star radius ratio inconsistent with these studies. As a result, the occultation of a bright region on the host star is the favoured interpretation of the feature seen in the light curves.

I find that Rayleigh scattering is not the dominant source of opacity within the planetary atmosphere. When modelling the in-transit anomalies as spots, I derive a transmission spectrum consistent with wavelength-independent absorption by clouds and find no evidence for any broad sodium absorption, although I cannot rule out the presence of the narrow line core. For the favoured facula model, I find an increasing planetary radius towards the red optical which could be interpreted as evidence for the broad wings of sodium.

This work demonstrates how high precision photometry can be used to study exoplanet atmospheres and analyse the effects of stellar activity, in the form of spots and bright regions, on the derived transit parameters.

Chapter 5

Transmission spectroscopy of HAT-P-18b

In this chapter, I present the transmission spectrum of the hot Jupiter HAT-P-18b acquired with the ACAM instrument on the William Herschel Telescope (WHT). Differential spectroscopy over an entire night was carried out at a resolution of $R \approx 400$ using a nearby comparison star. I detect a bluewards slope extending across my optical transmission spectrum which runs from $4750\,\text{Å}$ to $9250\,\text{Å}$. The slope is consistent with Rayleigh scattering at the equilibrium temperature of the planet ($852\,\text{K}$). I do not detect enhanced sodium absorption, which indicates that a high-altitude haze is masking the feature and giving rise to the Rayleigh slope. This is only the second discovery of a Rayleigh scattering slope in a hot Jupiter atmosphere from the ground, and my study illustrates how ground-based low-resolution spectroscopy can provide transmission spectra with precision comparable to the *Hubble Space Telescope*.

HAT-P-18b was the first target I observed using ACAM and the first result from the LRG-BEASTS programme, which was described in Chapter 2. The results presented in this chapter were published by Kirk et al. 2017, MNRAS, 468, 3907.

5.1 Introduction

As discussed in Chapter 1, transmission spectroscopy has revealed a continuum in hot Jupiter atmospheres from clear atmospheres showing strong alkali metal absorption to cloud- and haze-dominated atmospheres muting, and sometimes masking entirely, the absorption lines of the alkali metals (e.g. Sing et al., 2016).

Ground-based detections of spectral features within hot Jupiter atmospheres

have predominantly been of the narrow line cores of sodium (e.g. Snellen et al., 2008; Zhou & Bayliss, 2012; Nikolov et al., 2016) and potassium (e.g. Sing et al., 2011b; Wilson et al., 2015; Sedaghati et al., 2016). Whilst there have been a couple of detections of bluewards scattering slopes from the ground (e.g. Jordán et al., 2013; Di Gloria et al., 2015), ground-based measurements of hot Jupiters have often revealed featureless transmission spectra dominated by clouds (e.g. Kirk et al., 2016; Mallonn et al., 2016). The diversity of these results has emphasised the need to increase the current pool of studied gas giant atmospheres to better understand the processes and parameters governing the presence or absence of clouds and hazes.

In this chapter, I present the low-resolution ground-based transmission spectrum of the Saturn-mass planet HAT-P-18b (Hartman et al., 2011). HAT-P-18b's equilibrium temperature of $852 \, \text{K}$ is relatively cool for a hot Jupiter owing to its comparatively long period of $5.5 \, \text{d}$. HAT-P-18b's inflated radius ($0.995 \, R_{\text{J}}$) and relatively low mass ($0.197 \, M_{\text{J}}$) lead to a large atmospheric scale height of $540 \, \text{km}$. The combination of its large scale height with the small radius of the host star ($0.749 \, \text{R}_\odot$) makes HAT-P-18b well suited to transmission spectroscopy.

Referring to the ranking of all exoplanets by the absorption from one atmospheric scale height that was introduced in Chapter 2, HAT-P-18b was in the top 20 best planets for transmission spectroscopy at the time of my observations. The additional transit depth corresponding to one atmospheric scale height is $297 \, \text{ppm}$. Since my study, HAT-P-18b has not been studied again at optical wavelengths but has been the subject of HST/WFC3 transmission spectroscopy at near-infrared wavelengths (Tsiaras et al., 2017). At these wavelengths, Tsiaras et al. (2017) made a statistically significant detection of water.

5.2 Observations

HAT-P-18 was observed on the night of 2016 April 28 with the ACAM instrument (Benn et al., 2008) on the 4.2-m William Herschel Telescope (WHT), La Palma (see description of this setup in Chapter 3). For these observations, ACAM was used in spectroscopy mode, providing low resolution spectroscopy over the entire optical range with a throughput ranging between 0.5 and 0.8 (Benn et al., 2008). These observations were taken by Peter Wheatley and myself as part of WHT programme W/2016A/32 (PI: Wheatley).

ACAM was used unbinned in fast readout mode with a smaller than standard detector window to reduce the overhead to $\sim 11 \, \text{s}$ with exposure times of $60 \, \text{s}$. This integration time was used to limit the peak counts of the comparison star to the

range of the CCD response characterised as linear[1]. A custom made 27 arcsec slit was used to perform these observations to avoid the harmful effects of differential slit losses between the target and its comparison star which can lead to systematics in the derived transmission spectra. To further avoid slit losses, the locations of the spectral traces were monitored throughout the night (using the method described in Chapter 3), and manual guiding corrections made to keep the traces within a couple of pixels in the x (spatial) and y (dispersion) directions respectively (Fig. 5.3). A slight defocus was used to maintain the full width at half-maximum at around 1 arcsec. A total of 320 spectra were observed over the course of the night with airmass varying from 1.93 to 1.00 to 1.08 and a moon illumination of 62 per cent. Biases, sky flats and spectra of the CuAr plus CuNe arcs were taken at the start and end of the night.

A comparison star with a similar magnitude to HAT-P-18 was used to perform differential spectroscopy in order to account for telluric variability. The comparison star used was TYC 2594-731-1 at a distance of 3.4 arcmin from HAT-P-18 with a V magnitude of 11.2 and $B - V$ colour of 1.3. HAT-P-18 has a V magnitude of 12.7 and $B - V$ colour of 1.0. The separation between the stars was comfortably within ACAM's slit length of 6.8 arcmin. The comparison star is not known to be a variable star.

5.3   Data reduction

To reduce the data, I used the custom built PYTHON scripts described in Chapter 3.

To make a master bias, 51 bias frames were median-combined. I tried several methods to flat field but found the least red noise in the white light curve resulted from no flat-fielding. As discussed in Chapter 3, lamp flats did not contain enough blue photons while the sky flats contained sky emission lines. This study was performed before I experimented with using median and Gaussian filters to remove the sky lines (as described in detail in Chapter 3), however, through empirical tests on other nights' data, it is unlikely that smoothed sky flats would lead to a significant improvement in the photometry. I tried slitless sky flats but these were contaminated by a sector of the zero-order image, as depicted in Chapter 3. Since the spectral traces were spread across a large number of pixels, the lack of flat-fielding did not form a significant source of error. A more recent comparison between a reduction of this night's data with and without flat-fielding is presented in Chapter 7, which shows the difference is $< 100\,\mathrm{ppm}$ in the white light curve.

---

[1]http://www.ing.iac.es/Engineering/detectors/auxcam_fast.jpg

To extract the spectra, polynomials were fitted along each of the two traces in the dispersion direction and apertures placed around these. Fig. 5.1 shows an example frame with the aperture and background regions overplotted and Fig. 5.2 shows an example extracted spectrum.

To fit the background I iteratively fitted cubic polynomials row by row across the combined width of the extraction and background apertures. The iterative nature of this meant that the spectral traces themselves were masked from the resulting polynomial, which instead was only modelling the background variation in the spatial direction. In more recent versions of the reduction pipeline I mask the region within the extraction aperture from the background fit, which improves the performance of the code. However, both methods lead to the same result.

After subtraction of the polynomial, normal extraction of the counts within the aperture was performed. A number of different extraction and background aperture widths were experimented with to minimise the scatter in the resulting light curves. This resulted in an optimal aperture width of 30 pixels. The background was calculated from two 50 pixel-wide regions, either side of the aperture, at distances of 30 pixels from the edges of the aperture. The pixel scale of ACAM is 0.25 arcsec pixel^{-1}. The errors in the data points were a combination of the photon noise and the readnoise of ACAM, which, in fast readout mode, is 7 electrons pixel^{-1} with a gain of 1.86 electrons count^{-1}, as listed in Chapter 3.

After extraction of the spectral traces, cosmic rays were removed following the method laid out in Chapter 3. This was done by first dividing each spectral frame by a reference frame clean of cosmics. This self-division of each spectral trace resulted in spikes where cosmic rays were located. The flux values of the pixels affected by cosmic rays were then replaced by interpolating between the two nearest neighbouring pixels unaffected by the cosmic ray.

The differential white light curve of the target is plotted in Fig. 5.3 along with airmass and diagnostics of the traces over the course of the observation. This figure shows the small rotation of the traces on the CCD resulting from a combination of instrument flexure (Benn et al., 2008) and atmospheric refraction as there was no atmospheric dispersion corrector. Due to this rotation, the spectra were aligned in wavelength so that accurate differential spectroscopy could be performed. This was done by cross-correlating regions of each spectral frame with a reference frame around significant absorption features (as demonstrated in Chapter 3). This was repeated across the spectrum and allowed for the shift relative to the reference to be calculated as a function of location on the chip. A third order polynomial was then fitted to these shifts and the individual spectra resampled onto the grid of the
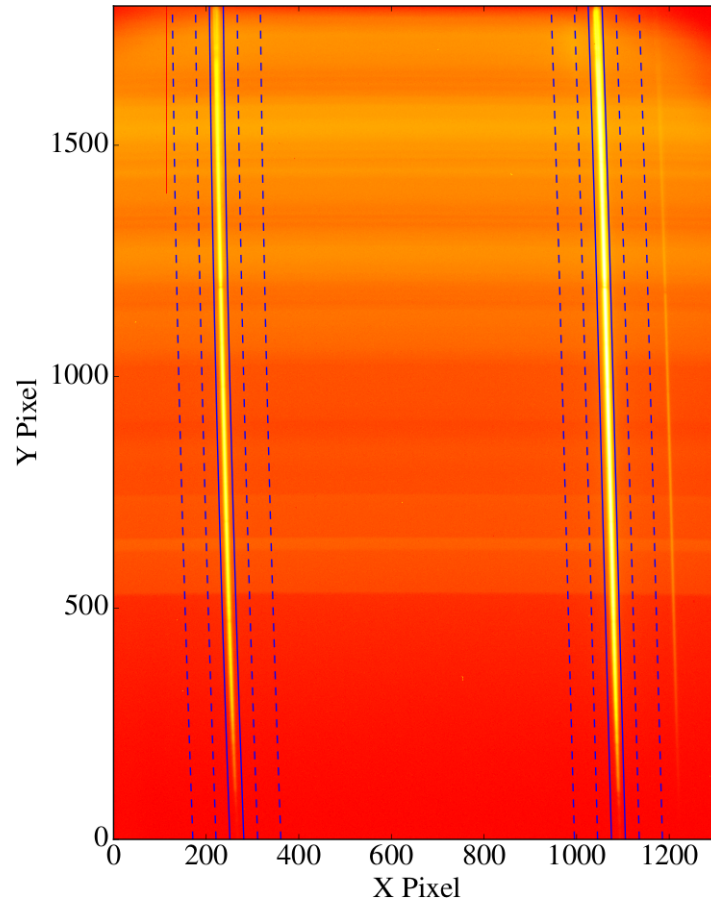
Figure 5.1: Example frame with extraction regions for the target (left-hand trace) and comparison (right-hand trace). The solid lines indicate the extraction apertures with the dotted lines indicating the sky background regions.

reference spectrum. The shifts for each star's spectra were calculated individually before resampling onto the same grid. This resulted in the spectra of both stars being well aligned for all frames.

With the spectra aligned in pixel space, the wavelength solution was calculated using strong telluric and stellar absorption lines. A synthetic spectrum of telluric and spectral lines with a resolution of $R = 1000$ was generated by Amanda Doyle using UCLSYN (Smith & Dworetsky, 1988; Smith, 1992; Smalley et al., 2001) for a star of the same temperature, surface gravity and metallicity as HAT-P-18, which I used to identify absorption lines in the observed spectrum. These absorption lines were then fitted with Gaussians (as described in Chapter 3). This gave the wavelength as a function of pixel position for a number of strong features. A second-order polynomial was then fitted to convert from pixel position into wavelength. I also calculated the arc solution using arcs taken through a 1 arcsec slit and found this was offset by 150 Å to the blue. This offset was because the centre of the 27 arcsec slit is 50 pixels lower than the 1 arcsec slit. Since the dispersion scale is ~ 3Å pixel^{-1}, this led to the offset seen. However, once correcting for this offset, the two solutions were consistent to within a couple of angstroms (1 per cent of the bin width).

For further analysis the very-blue and very-red ends of the spectra were excluded due to low signal to noise. This resulted in the spectra spanning a wavelength range between 4750 Å and 9250 Å.

5.4 Data Analysis

5.4.1 Light curve fitting with free limb darkening

With the data reduced and spectra extracted, I binned the data into 250 Å-wide wavelength bins running from 4750 Å to 9250 Å. This bin width was chosen as a compromise between resolution and signal-to-noise ratio. The bin containing the strong telluric oxygen feature at ~ 7600 Å showed a significant level of red noise and so this bin was masked from further analysis. This resulted in seventeen 250 Å-wide bins, which are shown in Fig. 5.2 with the masked region shaded grey.

I fitted the light curves with analytic transit light curves using a quadratic limb darkening law (Mandel & Agol, 2002). A long time-scale trend was simultaneously fitted to model the small overall trend which was less than a mmag in amplitude. This trend was most likely a second-order colour extinction effect in the 250 Å-wide bins. I experimented with different functions to fit this trend, including: a quadratic in time polynomial, a cubic in time polynomial, a linear in time extinc-
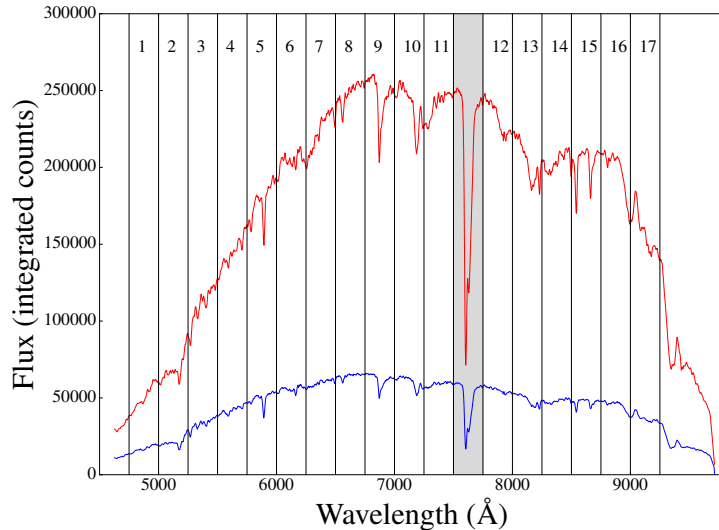
Figure 5.2: Locations of the seventeen 250 Å-wide bins used in the wavelength bin fitting. The target's spectrum is shown in blue and the comparison's in red, in raw counts for a single exposure. The grey region shows the bin containing the strong telluric feature that was excluded from further analysis.

tion coefficient and a quadratic in time extinction coefficient. All functions resulted in transmission spectra with blueward slopes, and the cubic in time polynomial was adopted for my final results due to the better Bayesian information criterion (BIC; Schwarz et al., 1978). The parameters defining the model were the radius ratio of planet to star R_P/R_*, the scaled stellar radius a/R_*, the inclination i, the quadratic limb darkening coefficients $u1$ and $u2$, the time of mid-transit T_c, and the four parameters defining the long time-scale trend.

The light curve models were initially fitted using SCIPY's OPTIMIZE package within PYTHON (Jones et al., 2001) using a Nelder-Mead algorithm to perform the minimisation (Nelder & Mead, 1965). With the results from this fit a Markov chain Monte Carlo (MCMC) was initiated at these values and was implemented using the PYTHON package EMCEE (Foreman-Mackey et al., 2013).

The system parameters (a/R_*, i and T_c) were held fixed to the results from a fit to the white light curve (Fig. 5.3, Table 5.2). This was created simply by summing the seventeen individual wavelength-binned light curves. I did this as I am interested in the relative error in the planetary radius between wavelengths and not in the absolute error in the radius. By fixing these system parameters shared between the wavelength bins, they are removed as sources of error within the relative radii.

I initially fitted the light curves with the quadratic limb darkening coefficients
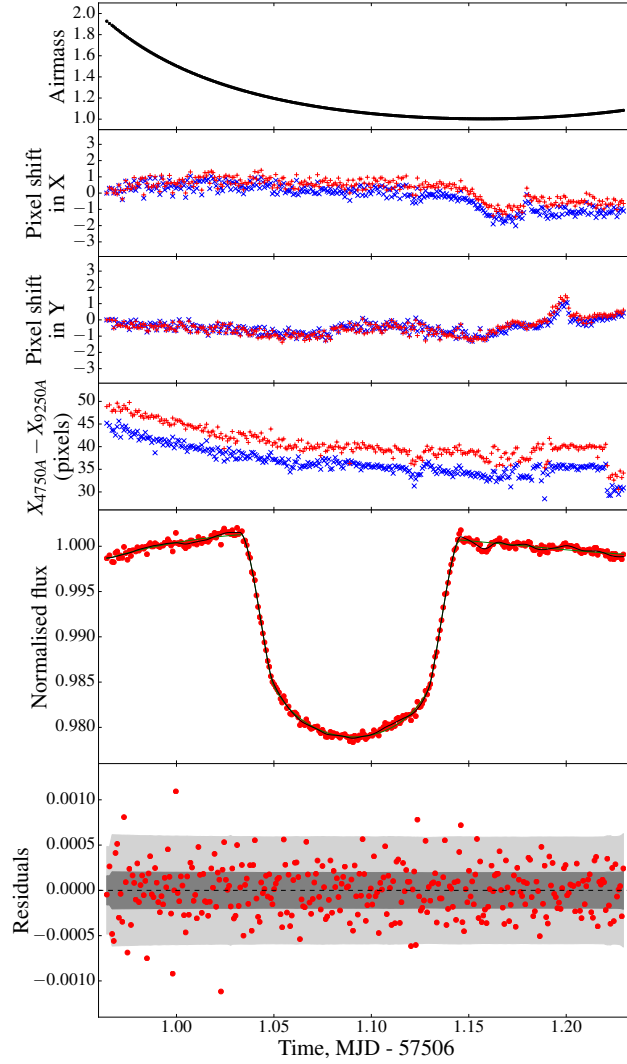
Figure 5.3: Top panel: Variation of airmass across the night. Second panel: Shift in the pixel positions in x (spatial direction) for the target trace (blue crosses) and comparison trace (red pluses). Third panel: Shift in the pixel positions in y (dispersion direction) for the target trace (blue crosses) and comparison trace (red pluses). Fourth panel: The rotation of the target trace (blue crosses) and comparison trace (red pluses) shown as the difference in x position at the bottom and top of the trace. Fifth panel: The normalised, differential, white light curve shown by red data points and generated from summing up the bins shown in Fig. 5.2. Over plotted are the fits from the analytic transit light curve with a cubic in time polynomial (green line, section 5.4.1) and a Gaussian process (GP, black line, section 5.4.3). Sixth panel: Residuals to the GP fit are given by the red points with the dark grey and light grey shaded regions indicating the 1σ and 3σ confidence intervals, respectively.

Table 5.1: Results from the fitting of the wavelength binned light curves shown in Fig. 5.4.

Bin	R_P/R_*	$u1$	$u2$	Residual rms (mmag)
$4750 - 4999$ Å	$0.14048^{+0.00108}_{-0.00110}$	$0.92^{+0.11}_{-0.12}$	$-0.142^{+0.186}_{-0.177}$	1.44
$5000 - 5249$ Å	$0.13970^{+0.00085}_{-0.00087}$	1.00 ± 0.10	$-0.382^{+0.155}_{-0.154}$	1.18
$5250 - 5499$ Å	0.13999 ± 0.00067	0.81 ± 0.08	$-0.132^{+0.121}_{-0.123}$	0.91
$5500 - 5749$ Å	$0.13898^{+0.00064}_{-0.00061}$	0.55 ± 0.07	$0.198^{+0.111}_{-0.107}$	0.78
$5750 - 5999$ Å	$0.14023^{+0.00054}_{-0.00053}$	0.63 ± 0.07	$-0.035^{+0.100}_{-0.103}$	0.70
$6000 - 6249$ Å	$0.13771^{+0.00056}_{-0.00055}$	0.48 ± 0.07	$0.203^{+0.109}_{-0.104}$	0.69
$6250 - 6499$ Å	0.13862 ± 0.00056	0.61 ± 0.07	$-0.030^{+0.107}_{-0.104}$	0.72
$6500 - 6749$ Å	$0.13870^{+0.00054}_{-0.00055}$	0.50 ± 0.07	0.104 ± 0.104	0.69
$6750 - 6999$ Å	$0.13817^{+0.00056}_{-0.00057}$	$0.41^{+0.07}_{-0.08}$	$0.206^{+0.112}_{-0.110}$	0.70
$7000 - 7249$ Å	0.13787 ± 0.00051	0.48 ± 0.07	$0.067^{+0.099}_{-0.102}$	0.64
$7250 - 7499$ Å	0.13863 ± 0.00056	$0.48^{+0.07}_{-0.08}$	$0.018^{+0.112}_{-0.109}$	0.70
$7750 - 7999$ Å	$0.13676^{+0.00058}_{-0.00059}$	0.43 ± 0.08	$0.077^{+0.118}_{-0.116}$	0.73
$8000 - 8249$ Å	$0.13769^{+0.00058}_{-0.00059}$	$0.60^{+0.07}_{-0.08}$	$-0.138^{+0.113}_{-0.111}$	0.75
$8250 - 8499$ Å	0.13895 ± 0.00062	0.58 ± 0.08	$-0.159^{+0.115}_{-0.112}$	0.77
$8500 - 8749$ Å	$0.13858^{+0.00091}_{-0.00095}$	$0.47^{+0.12}_{-0.13}$	$-0.031^{+0.187}_{-0.185}$	1.20
$8750 - 8999$ Å	$0.13722^{+0.00067}_{-0.00063}$	0.44 ± 0.09	$0.090^{+0.133}_{-0.130}$	0.84
$9000 - 9249$ Å	$0.13680^{+0.00077}_{-0.00078}$	$0.27^{+0.10}_{-0.11}$	$0.328^{+0.165}_{-0.153}$	0.96

$u1$ and $u2$ as free parameters. I used a uniform prior such that $u1 + u2 \leq 1$, enforcing the intensity to always be positive (e.g. Burke et al., 2007), and placed no priors on the other parameters. With the starting values resulting from the Nelder-Mead minimisation, an initial MCMC to each wavelength-binned light curve was run for 1000 iterations with 140 walkers ($20 \times n_p$, where n_p is the number of parameters). The first 500 steps were discarded as the chains were burning in. The best-fitting model resulting from this initial run was then used to rescale the error bars to give $\chi^2_\nu = 1$ for each wavelength bin. Following this, a second MCMC was run with 140 walkers and for 1000 iterations.

The results of the fits to the wavelength-binned light curves are shown in Fig. 5.4 and Table 5.1 with the rms of the residuals between 20 per cent and 80 per cent of the expected photon precision. The resulting transmission spectrum displayed a gradient running from blue to red (Fig. 5.5, top panel).

5.4.2 Light curve fitting with constrained limb darkening

When fitting for both limb darkening coefficients, I found the resulting values were consistent with predicted values for this star (Fig. 5.6). To generate the predicted
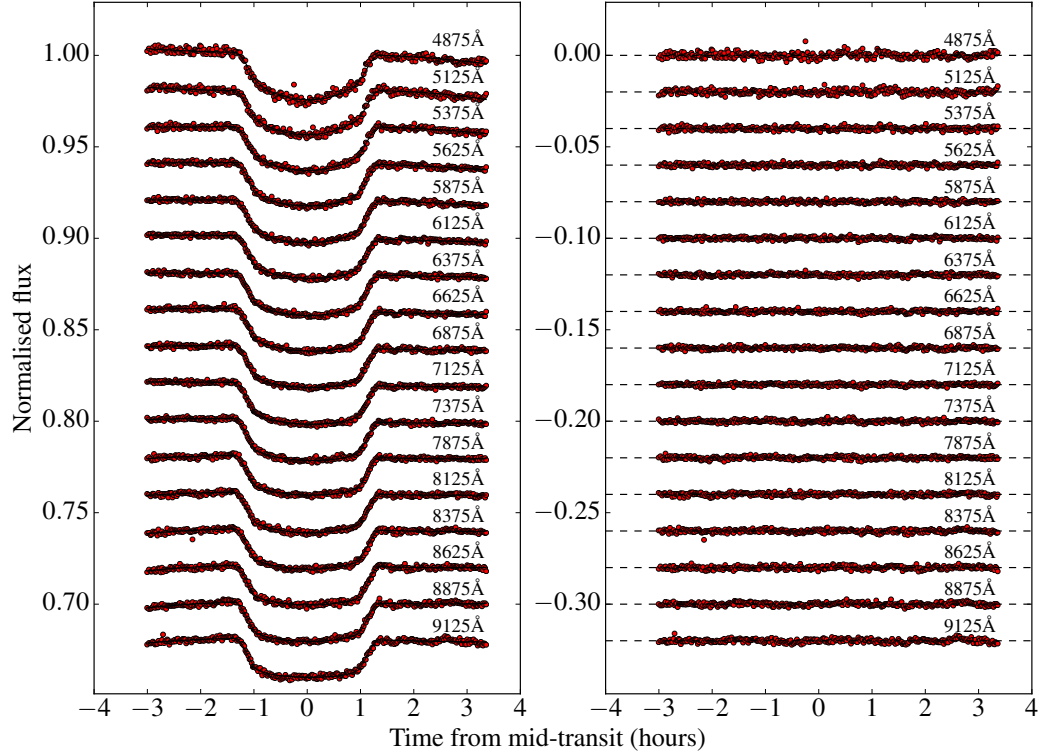
Figure 5.4: Left-hand panel: My fits to the wavelength binned light curves of HAT-P-18b, going from blue (top) to red (bottom), which are offset for clarity. The black lines are the fitted models with both limb darkening coefficients as fit parameters and using the cubic in time polynomial, giving 7 free parameters per wavelength bin. The red points show the normalised differential flux without any detrending. The quoted wavelengths are those of the centre of each bin. Right-hand panel: the residuals to the fits in the left-hand panel after subtracting the best-fitting model.
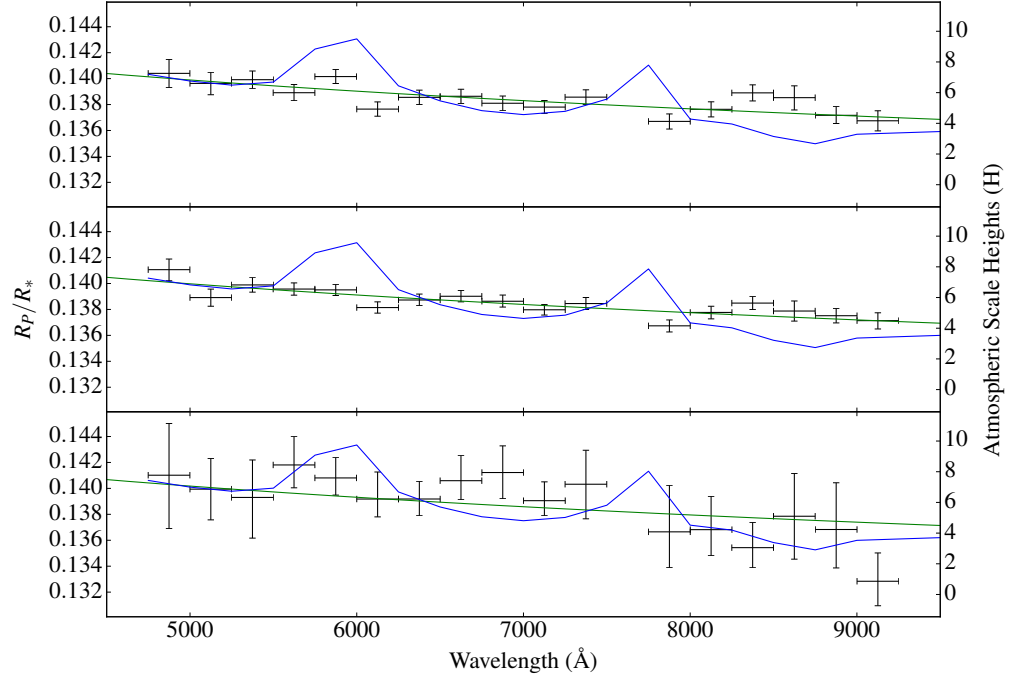
Figure 5.5: Transmission spectrum of HAT-P-18b resulting from different treatments of the limb darkening and long time scale trend. Top panel: analytic transit light curves with a cubic in time polynomial. In this fit, both $u1$ and $u2$ were free parameters (section 5.4.1). Second panel: analytic transit light curves with a cubic in time polynomial. In this fit $u1$ was fixed with $u2$ described by a polynomial (section 5.4.2). Bottom panel: analytic transit light curves with a Gaussian process (section 5.4.3). In this fit $u1$ was fixed with $u2$ described by a polynomial. In each panel, the black data points are the resulting R_P/R_* values. The green line shows a Rayleigh scattering slope at the equilibrium temperature of the planet (852 K). The blue line shows a clear atmosphere model, binned to the size of the data points, highlighting the absence of the broad wings of the Na and K features, indicating the presence of a high-altitude haze.

values, I made use of the *Limb Darkening Tool Kit* (LDTK, Parviainen & Aigrain, 2015) which uses PHOENIX models (Husser et al., 2013) to calculate $u1$ and $u2$ with errors propagated from the errors in the stellar parameters. Fig. 5.6 also shows the anticorrelation in the scatter of $u1$ and $u2$ around their expected values. Because of this degeneracy, it is common to hold one coefficient fixed and fit for the other (e.g. Southworth, 2008; Kirk et al., 2016) and so I experimented with this also. Since I am interested in the relative radii between wavelengths, fixing one of the limb darkening coefficients removes this as a source of error within R_P/R_*. I therefore ran another MCMC with $u1$ fixed and $u2$ as a free parameter.

I held $u1$ fixed to a second-order polynomial fitted to the predicted values of $u1$ generated by LDTK (Fig. 5.6, lower panel). This removed $u1$ as a fit parameter, leaving R_P/R_*, $u2$, and the four parameters defining the long time-scale trend as fit parameters. I ran an initial MCMC to each wavelength binned light curve for 1000 iterations with 120 walkers ($20 \times n_p$), and discarded the first 500 steps as the chains were burning-in. The best fitting models from this run were then used to rescale the error bars to give $\chi^2_\nu = 1$ for each wavelength bin. Following this initial run of the MCMC, I fitted a second-order polynomial to the resulting values for $u2$ to describe $u2$ as a function of wavelength (Fig. 5.6, lower panel). A second MCMC was then run with 100 walkers ($20 \times n_p$) for 1000 steps with this function describing $u2$, thus removing it as a fit parameter.

This treatment of the limb darkening led to the same slope in the transmission spectrum as the fit with the limb darkening coefficients free, but with slightly smaller errors in R_P/R_* (Fig. 5.5, middle panel). Fig. 5.6 (lower panel) shows the fitted values of $u2$ compared with those generated by LDTK. While the fitted values for $u2$ did differ to those from LDTK, importantly, this did not affect the slope in the transmission spectrum. The difference between the predicted and fitted limb darkening coefficients resulting from this method is discussed in section 5.5.2.

5.4.3   Light curve fitting with Gaussian process detrending

Whilst the models using the cubic in time polynomial provided good fits to the data (Fig. 5.4), I wanted to confirm that the transmission spectra were independent of the function used to model the long time-scale trend. To do this, I used a Gaussian process (GP) which is a non-parametrised method to model covariance in data (see Chapter 3). This offered a robust test of whether the results from the use of the cubic in time polynomial were independent of the choice of function used. GPs have been demonstrated to be powerful in modelling correlated noise in exoplanet transit light curves by, for example, Gibson et al. (2012a, 2013b) and Evans et al. (2015).
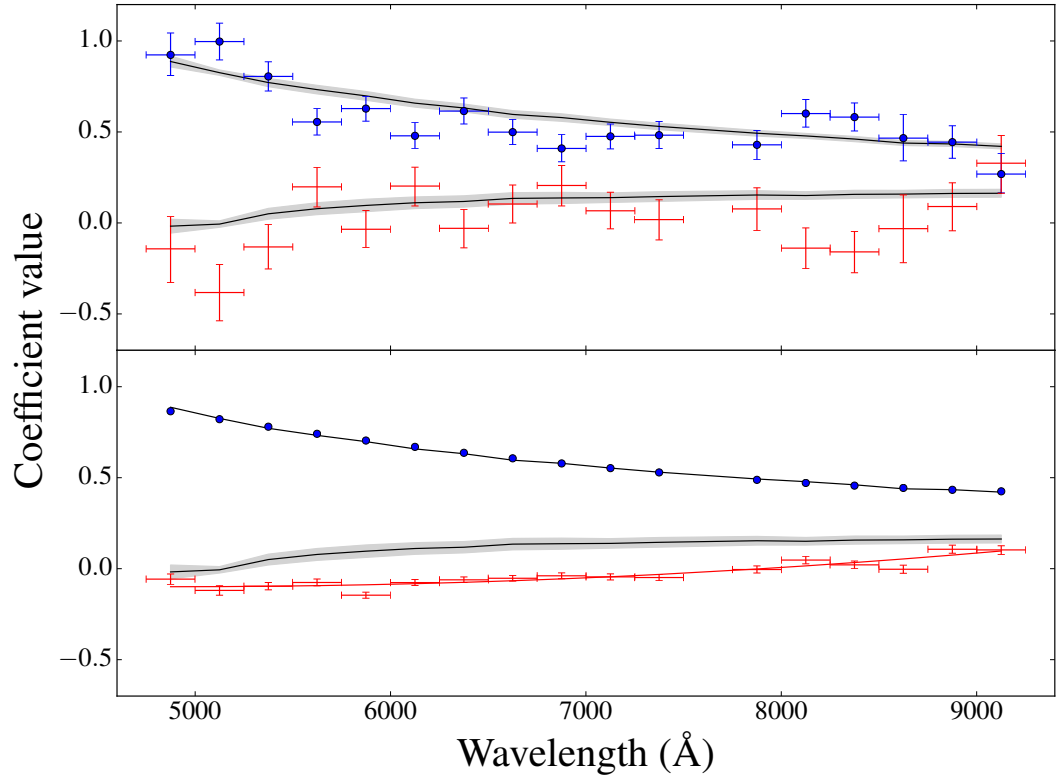
Figure 5.6: Top panel: plot of the limb darkening coefficients when both are fitted as free parameters. The blue error bars show the fitted values for $u1$ and the red error bars show those for $u2$. The solid black lines show the predicted values for $u1$ and $u2$ generated by LDTK while the grey regions show the 1σ confidence regions. Bottom panel: results when $u1$ is held fixed to a polynomial fitted to the predicted values of $u1$. The blue data points indicate the values that $u1$ is fixed to for each bin. The red error bars indicate the fitted values for $u2$ with the solid red line showing the polynomial fitted to these values of $u2$. The black line and grey regions are the same as in the top panel.

140

The GP was implemented through the GEORGE PYTHON package (Ambikasaran et al., 2014). The mean function of the GP in each wavelength bin was the quadratically limb darkened analytic transit light curve (Mandel & Agol, 2002). I treated the limb darkening in the same way as described in section 5.4.2. I used a Matérn 3/2 kernel to model correlations in the data, defined by the hyper-parameters τ (the time-scale) and a (the amplitude), in addition to a white-noise kernel defined by the variance σ^2. The fit parameters were therefore R_P/R_*, $u2$, a, τ and σ^2, with $u1$ described by the same polynomial as before. I used loose, uniform priors to encourage convergence.

An initial MCMC was run for 2000 steps with 100 walkers ($20 \times n_p$) with the first 1000 steps discarded as burn-in. As before, I then ran a second MCMC with a function describing the fitted values of $u2$ resulting from the first run, thus removing $u2$ as a fit parameter. The second run was 2000 steps long with 80 walkers ($20 \times n_p$).

When fitting with the GP, I held the system parameters (i, a/R_* and T_c) fixed to values obtained from fitting the white light curve with a GP (Fig. 5.3, Table 5.2).

5.4.4 Transmission spectrum

The transmission spectrum resulting from each of the three fitting methods is displayed in Fig. 5.5, revealing a strong blueward slope resembling a Rayleigh scattering signature. The transmission spectra were consistent between all methods, confirming that the presence of the slope was independent of the choice of function used to fit the long time scale trend and whether I fixed or fitted the limb darkening coefficients. The GP resulted in larger errors in the transmission spectrum due to its ability to model a much wider set of detrending functions.

A Rayleigh scattering slope is plotted on this figure at the equilibrium temperature of the planet (852 K; Hartman et al., 2011), with a slope given by Lecavelier Des Etangs et al. (2008a) as

$$\frac{dR_p}{d\ln\lambda} = \frac{k}{\mu g}\alpha T \tag{5.1}$$

where μ is the mean molecular mass of an atmospheric particle taken to be 2.3 times the mass of a proton, k is the Boltzmann constant, g is the planet's surface gravity, $\alpha = -4$ as expected for Rayleigh scattering, and T I take as the equilibrium temperature.

The transmission spectra resulting from the three different modelling tech-

niques were all well fitted by a Rayleigh slope at the equilibrium temperature of the planet (852 K).

When using the cubic in time polynomial with both $u1$ and $u2$ as free parameters (Fig. 5.5, top panel) the ΔBIC between a Rayleigh slope at the equilibrium temperature and a flat transmission spectrum was 22.9. This very strongly favoured the Rayleigh slope (cf. Chapter 3, Kass & Raftery, 1995). A fit to this transmission spectrum with the temperature as a free parameter resulted in a temperature of 784 ± 194 K, consistent with the equilibrium temperature.

When fixing $u1$ and using the cubic in time polynomial (Fig. 5.5, middle panel), the ΔBIC between the Rayleigh slope at the equilibrium temperature and a flat transmission spectrum was 37.1, which again very strongly favoured the Rayleigh slope. The temperature resulting from a Rayleigh slope fitted to the transmission spectrum was 798 ± 150 K, which was again consistent with the equilibrium temperature.

The GP resulted in larger errors in the transmission spectrum (Fig. 5.5, bottom panel) and consequently lower values of the ΔBIC between models. However, a Rayleigh slope at the equilibrium temperature was still strongly favoured over a flat transmission spectrum with a ΔBIC of 8.6. When fitting this transmission spectrum with a free temperature, the preferred gradient was at a significantly higher temperature of 2023 ± 393 K. However, the ΔBIC between the fitted temperature and equilibrium temperature was less than 2, which is insignificant, and so I do not conclude this as evidence for a temperature inversion.

I also plot on Fig. 5.5 a clear, solar-metallicity, atmosphere model (in blue) as resulting from the NEMESIS radiative transfer code (Irwin et al., 2008), binned to the resolution of the data. The clear atmosphere model does not provide as good a fit as the Rayleigh scattering slope, as I do not detect the broad wings of the sodium and potassium features. This indicates a condensate haze is masking the wings as in, for example, HD 189733b (Pont et al., 2008; Sing et al., 2011a; Huitson et al., 2012; Pont et al., 2013), WASP-31b (Sing et al., 2015) and WASP-6b (Jordán et al., 2013; Nikolov et al., 2015).

Although Rayleigh scattering has been seen in a number of planets to date by *HST* (e.g. Pont et al., 2013; Fischer et al., 2016), it had only been seen in two hot Jupiter atmospheres from the ground prior to this study (Jordán et al., 2013; Di Gloria et al., 2015). Of these two ground-based detections of Rayleigh scattering slopes, one was the discovery of the slope in WASP-6b using transmission spectroscopy (Jordán et al., 2013) and the other was a re-detection of the slope in HD 189733b using the chromatic Rossiter-McLaughlin effect (Di Gloria et al., 2015).

Whilst there had been other detections of blueward slopes from the ground, these have been considerably steeper than the expected Rayleigh scattering slopes and require further explanation (e.g. Southworth et al., 2015; Parviainen et al., 2016; Southworth & Evans, 2016). The detection in Fig. 5.5 therefore represents only the second discovery of a Rayleigh scattering slope in a hot Jupiter atmosphere from the ground.

5.4.5   Targeted sodium search

While clouds and hazes can mask the broad wings of sodium, it is possible for the narrow line core to be visible as it originates at higher altitudes. This is the case in HD 189733b (Pont et al., 2008; Sing et al., 2011a; Huitson et al., 2012; Pont et al., 2013; Louden & Wheatley, 2015) and HD 209458b (Charbonneau et al., 2002; Sing et al., 2008a,b; Snellen et al., 2008; Langland-Shula et al., 2009; Vidal-Madjar et al., 2011; Deming et al., 2013). I performed a separate targeted search for the narrow feature of sodium using eight 50 Å-wide bins running from 5670 to 6070 Å, with one bin centred on the sodium doublet. This revealed significant red noise in the bin containing the Na feature, which can be seen in Fig. 5.7. While the cause of this variability is unknown it could possibly be instrumental or stellar in origin. To properly account for this red noise, and to obtain robust errors for the transit depth, I again used a GP, which was implemented in the same way as when fitting the 250 Å-wide bins.

The wavelength binned fits with the GP are shown in Fig. 5.7, which displays the power of the GP to model the red noise in bin 5 whilst not overfitting the data. The transmission spectrum around the sodium feature is shown in Fig. 5.8, which shows I do not detect sodium in the atmosphere at a resolution of 50 Å.

I was unable to place constraints on the presence of potassium due to its proximity to the deep telluric oxygen feature at 7600Å (Fig. 5.2).

5.4.6   Unoccultedspots

Although I do not observe any occultations of star spots within the transit of HAT-P-18b, I do need to take into account the possibility of star spots that may be present on the surface of the star but are not along the transit chord. Seeliger et al. (2015) monitored the activity of HAT-P-18 over a timespan of 12 months. They found a mean variation in the R-band brightness of HAT-P-18 of only ~ 0.9 mmag, which they find is consistent with no variation when taking their error bars into account. Therefore HAT-P-18 does not show strong activity induced photometric variation
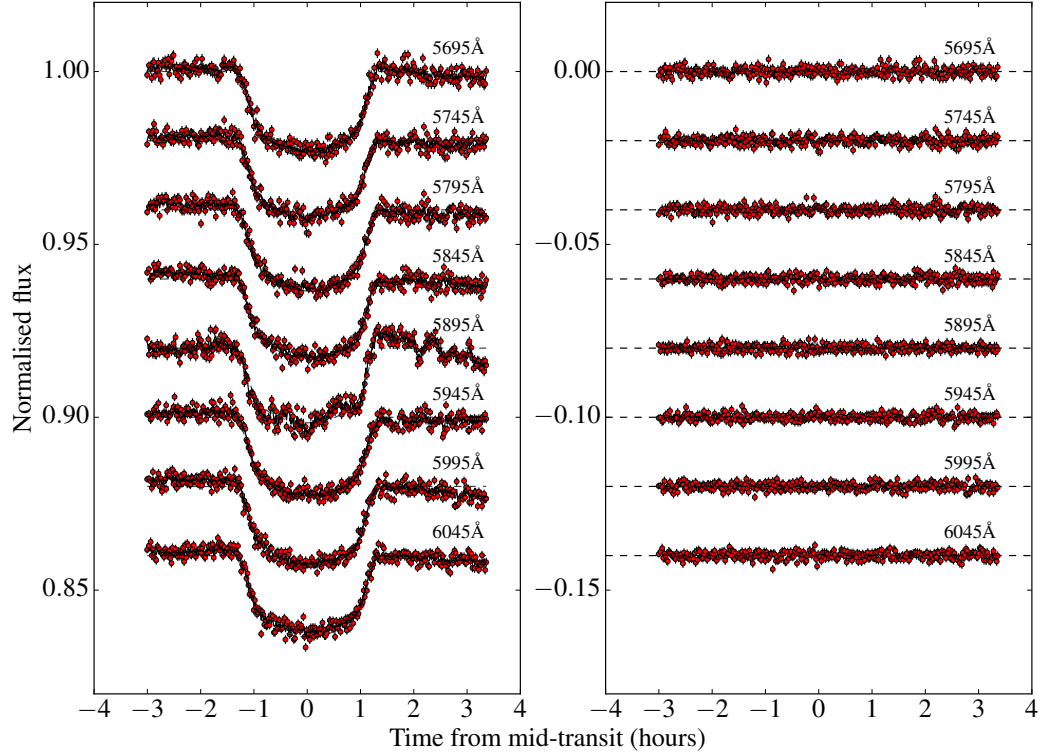
Figure 5.7: Left-hand panel: fits to the 50 Å-wide wavelength bins around the sodium feature. I use a GP to model these light curves due to the significant correlated noise visible at 5895 Å. The dashed lines indicate the underlying Mandel & Agol (2002) transit model defining the mean function of the GP. Right-hand panel: the residuals to the fits in the left-hand panel after subtracting the best fitting model.
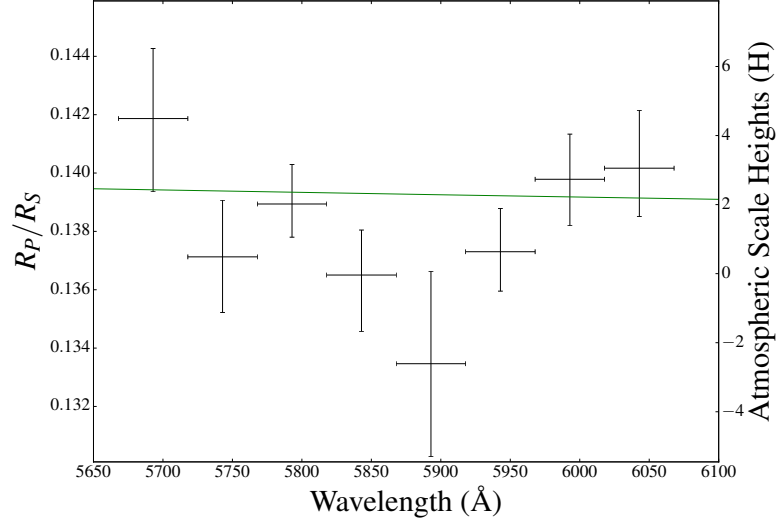
Figure 5.8: The transmission spectrum centred around the sodium feature at a resolution of 50 Å, which is not detected. The green line indicates a Rayleigh slope at the equilibrium temperature of the planet.

and so I do not expect unocculted activity regions to have a noticeable effect on the transmission spectrum. Nevertheless, I quantify their effects below.

Unocculted spots and plages can have the effect of inducing a slope mimicking Rayleigh scattering within the transmission spectrum and so their effect needs to be considered (McCullough et al., 2014; Oshagh et al., 2014). These unocculted spots can be accounted for by a wavelength-dependent depth correction. As with Kirk et al. (2016) (Chapter 4), I follow the formalism of Sing et al. (2011a) to make this correction.

I use ATLAS9 stellar models (Kurucz, 1993) of a star with a temperature of 4750 K and spot temperatures with a temperature difference ΔT ranging from $250 - 1250$ K. Using this variation I assume a total dimming of 0.1 per cent at a reference wavelength of 6000 Å and use equations (4) and (5) of Sing et al. (2011a) to find the correction in R_P/R_* across a wavelength range spanning $4500 - 9500$Å. I find that the effect of spots on this star is minimal as they lead to a correction in R_P/R_* of between 0.00005 and 0.0001 which is a small fraction of the 1σ error bars of my R_P/R_* values. Unocculted spots cannot therefore be the cause of the blueward slope seen in the transmission spectrum (Fig. 5.5).

145

5.4.7 Comparison of system parameters

With the high-quality light curves, I was able to compare the results from my white light curve fitting to those published in the literature (Table 5.2), which I find are consistent with previous studies.

Table 5.2: Comparison of system parameters to previous studies, resulting from the fits of a cubic in time polynomial and a GP to the white light curve with the limb darkening coefficients as free parameters. Although these values are consistent, the GP fit produces slightly larger errors, as expected, and so should be adopted as my conservative final values, which are shown in boldface.

Parameter	This work	Seeliger et al. (2015)	Esposito et al. (2014)	Hartman et al. (2011)
Cubic polynomial				
R_P/R_*	$0.1385^{+0.0010}_{-0.0011}$	0.1362 ± 0.0011	0.136 ± 0.011	0.1365 ± 0.0015
i (deg.)	$88.63^{+0.12}_{-0.10}$	88.79 ± 0.21	88.79 ± 0.25	88.8 ± 0.3
a/R_*	$16.71^{+0.18}_{-0.16}$	17.09 ± 0.71	16.76 ± 0.82	16.04 ± 0.75
T_c (BJD, days)	$2457507.59300891 \pm 0.000060$	$2454715.02254 \pm 0.00039$	2455706.7 ± 0.7	$2454715.02174 \pm 0.00020$
Gaussian process				
R_P/R_*	$\mathbf{0.1356^{+0.0028}_{-0.0024}}$	0.1362 ± 0.0011	0.136 ± 0.011	0.1365 ± 0.0015
i (deg.)	$\mathbf{88.53^{+0.16}_{-0.13}}$	88.79 ± 0.21	88.79 ± 0.25	88.8 ± 0.3
a/R_*	$\mathbf{16.39^{+0.24}_{-0.23}}$	17.09 ± 0.71	16.76 ± 0.82	16.04 ± 0.75
T_c (BJD, days)	$\mathbf{2457507.59219566^{+0.000194}_{-0.000181}}$	$2454715.02254 \pm 0.00039$	2455706.7 ± 0.7	$2454715.02174 \pm 0.00020$

5.5 Discussion

5.5.1 Transmission spectrum

The transmission spectrum of HAT-P-18b displays a gradient rising towards the blue (Fig. 5.5), which is consistent with Rayleigh scattering at the equilibrium temperature of the planet (852 K). Whilst I detect the Rayleigh slope, I do not detect either the broad wings of sodium or the line core in a 50 Å-wide bin (Fig. 5.8). This suggests that a high-altitude haze is masking the sodium feature in the atmosphere and giving rise to the Rayleigh slope, similar to, for example, WASP-12b (Sing et al., 2013), WASP-6b (Jordán et al., 2013; Nikolov et al., 2014) and HAT-P-12b (Sing et al., 2016). A haze is also preferable due to the better fit than the clear atmosphere model in Fig. 5.5.

HAT-P-18b, with an equilibrium temperature of 852 K (Hartman et al., 2011), is cooler than any of the planets studied by Sing et al. (2016) in their survey of 10 hot Jupiters. The nearest comparison object is HAT-P-12b (Hartman et al., 2009) with an equilibrium temperature of 963 K. For this object, Sing et al. (2016) found a haze layer leading to a Rayleigh scattering slope extending across the optical spectrum. They also detected potassium absorption although did not detect sodium. Whilst this might indicate some correlation of haze with temperature, Sing et al. (2016) found that the presence of clouds and hazes is not strongly dependent on temperature in their sample.

The difficulty in pinning down the relation between the presence of clouds and hazes and the planetary parameters was further highlighted in the discussion of Fischer et al. (2016) concerning the clear atmosphere of WASP-39b, which is the only hot Jupiter studied to date in which the broad wings of both sodium and potassium are visible. In this paper, the authors consider, among other parameters, the role of surface gravity and metallicity on the planet's atmosphere. If surface gravity were the dominant factor, leading to differences in the settling rates of condensates, we would also expect to see condensates in the atmosphere of WASP-39b as it has a lower surface gravity ($\log g = 2.61$, Faedi et al., 2011) than both HAT-P-12b ($\log g = 2.75$, Hartman et al., 2009) and HAT-P-18b ($\log g = 2.69$, Hartman et al., 2011). However, this is not seen in the data. If we consider metallicity, HAT-P-18 has a higher [Fe/H] of 0.1 than both HAT-P-12 (-0.29) and WASP-39 (-0.12). If we assume that the stellar metallicity is a good proxy for the planet metallicity, the presence of condensates in the atmosphere of HAT-P-18b might be due to its higher [Fe/H]. The absence of condensates in the atmosphere of WASP-39b could therefore be related to its lower metallicity. This argument however breaks down

when considering HAT-P-12b, which has a metal-poor host yet a hazy atmosphere.

However, since my study of HAT-P-18b, progress has been made in searching for correlations between planetary parameters and the presence of clouds and hazes, with mounting evidence that temperature is a key factor (Heng, 2016; Stevenson, 2016; Crossfield & Kreidberg, 2017; Fu et al., 2017). In this scenario, hotter planets are more likely to be cloud-free as it is too hot for condensibles and there are fewer soot precursors (e.g. Morley et al., 2015). As I find that HAT-P-18b is hazy, this is compatible with this argument. HAT-P-18b's equilibrium temperature of $852\,\mathrm{K}$ puts it at the boundary of planets likely to be hazy found by Crossfield & Kreidberg (2017) ($850\,\mathrm{K}$), although these authors were considering Neptunes not Jupiters, and close to the boundary of planets likely to be hazy of Stevenson (2016) ($700\,\mathrm{K}$). Its temperature also means that CH_4 is expected to be the dominant carbon-bearing molecule (cf. Chapter 1 and Crossfield, 2015) meaning that soot-precursors can form through its photolysis (Morley et al., 2015). Furthermore, alkali chlorides condense at temperatures near HAT-P-18b's equilibrium temperature which could also be contributing to the aerosols in the upper atmosphere (e.g. Pinhas & Madhusudhan, 2017).

My detection of aerosols in the atmosphere of HAT-P-18b supports the arguments above. However, in a more recent study by Tsiaras et al. (2017), a statistically significant detection of water was made in this planet's atmosphere. These authors found the volume mixing ratio of water to be -2.63 ± 1.18 dex, consistent with the abundance of H_2O in a solar composition planet at this equilibrium temperature (~ -3.0 dex, e.g. Madhusudhan et al., 2014). Given I detect haze in the optical, we might expect to see muted water features in the infrared, such as in the case of HAT-P-12b (Line et al., 2013; Sing et al., 2016). However, this doesn't seem to be the case and instead the combined optical and near-infrared transmission spectrum is more similar to HD 189733b, where haze is observed in the optical and a solar-metallicity water feature is seen in the near-infrared (Pont et al., 2008; Sing et al., 2011a; McCullough et al., 2014). As discussed in Chapter 1, Heng (2016) suggested that clouds in the optical did not necessarily imply clouds in the infrared. My work highlights the importance of combined optical and infrared transmission spectra for interpreting the strength of the water absorption and building a self-consistent picture of the whole planetary atmosphere.

5.5.2 Limb darkening parameters

When fitting my light curves with the limb darkening coefficients as free parameters, I found good agreement between the fitted and predicted values but with large error
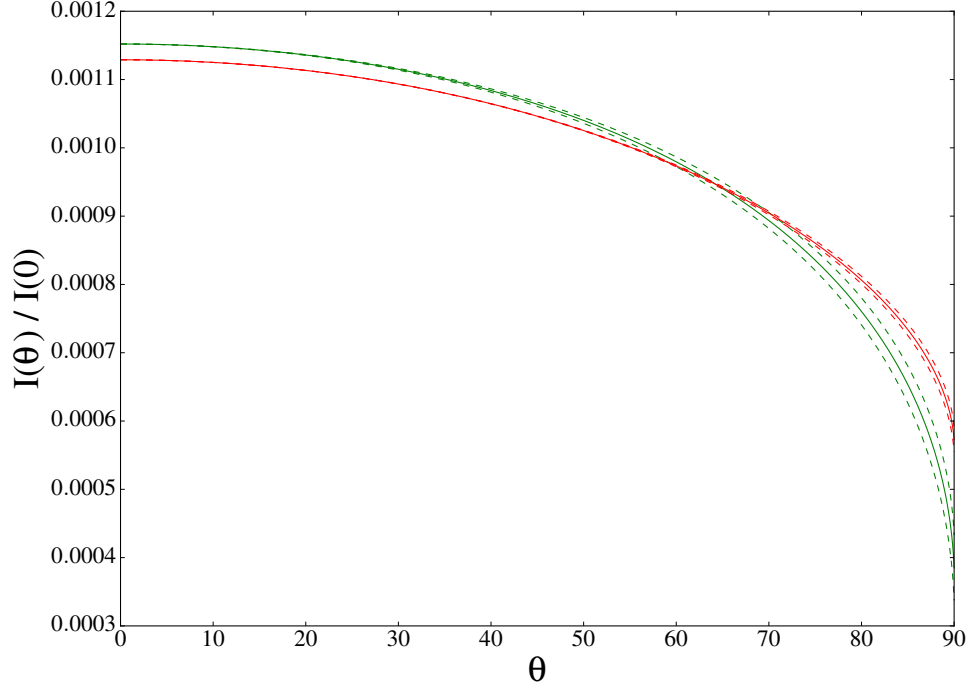
Figure 5.9: Comparison between the stellar limb darkening profile using the predicted values for $u1$ and $u2$ (green line) with those from holding $u1$ fixed and fitting for $u2$ (red line) for wavelength bin 10. The dashed lines show the 1σ upper and lower bounds for each. Both profiles have been normalised such that the total luminosity is equal to unity.

bars due to the degeneracy between $u1$ and $u2$ (Fig. 5.6, upper panel). This figure also demonstrates how fixing $u1$ leads to much more precise values for $u2$. However, fixing $u1$ at the predicted values led to $u2$ values that deviated significantly from the model (Fig. 5.6, lower panel). Fig. 5.9 shows the stellar brightness profile for one wavelength bin resulting from the use of the predicted limb darkening coefficients and the profile resulting from fixing $u1$ and fitting for $u2$. This shows that the data favours a profile with a brighter limb than the profile using the predicted limb darkening coefficients. This could indicate that I am measuring a more realistic profile of the star than the 1D model atmospheres predict. Indeed, it has been shown that limb darkening profiles resulting from 1D and 3D models do differ and 3D profiles can lead to brighter limbs (e.g. Hayek et al., 2012; Magic et al., 2015). My results highlight the value of exoplanet transits in probing stellar atmospheres in detail.

5.6 Conclusions

In this chapter I presented the transmission spectrum of the hot Jupiter HAT-P-18b using the low-resolution grism spectrograph ACAM on the WHT. I find a strong blueward scattering slope, extending across my transmission spectrum from 4750 to 9250 Å, consistent with Rayleigh scattering at the equilibrium temperature of the planet. I do not detect enhanced absorption around the sodium doublet, suggesting a high altitude haze is masking this feature whilst giving rise to the Rayleigh slope. I consider the effect of unocculted spots and find that the slope cannot be explained by these. This is only the second discovery of a Rayleigh scattering slope in a hot Jupiter atmosphere from the ground.

Given HAT-P-18b's relatively cool equilibrium temperature of 852 K (Hartman et al., 2011), the fact that I detect a haze in its atmosphere appears to be consistent with recent studies that have found cooler planets are more likely to be hazy.

My results for HAT-P-18b demonstrate that ground-based observations are capable of detecting atmospheric opacity sources such as Rayleigh scattering and highlight the importance of optical data in order to get a more complete understanding of atmospheric clouds and hazes. This work has also shown that optical data are important when interpreting the strength of water absorption detected with HST.

Chapter 6

Transmission spectroscopy of WASP-80b

In this chapter I present ground-based transmission spectroscopy of the hot Jupiter orbiting the cool dwarf WASP-80. These observations were taken with the ACAM instrument on the William Herschel Telescope (WHT) as part of the LRG-BEASTS programme. We observed two transits of the planet and constructed transmission spectra spanning a wavelength range of 4640 – 8840 Å. My transmission spectrum is inconsistent with a previously claimed detection of potassium in WASP-80b's atmosphere, and is instead most consistent with a haze. I also do not see evidence for sodium absorption at a resolution of 100 Å.

This study constituted the third paper of the LRG-BEASTS survey following the detection of a haze-induced Rayleigh scattering slope in the atmosphere of HAT-P-18b (Chapter 5 and Kirk et al., 2017) and clouds in the atmosphere of WASP-52b (Louden et al., 2017).

The results presented in this chapter were published by Kirk et al. 2018, MNRAS, 474, 876.

6.1   Introduction

Transmission spectroscopy is revealing a diverse array of exoplanet atmospheres, from clear (e.g. WASP-39b; Fischer et al., 2016; Nikolov et al., 2016; Sing et al., 2016) to hazy (e.g. HAT-P-18b; Kirk et al., 2017) to cloudy (e.g. WASP-52b; Kirk et al., 2016; Chen et al., 2017; Louden et al., 2017). There is increasing evidence that hotter exoplanets are more likely to be cloud free, however, these studies have been performed on relatively small sample sizes (Heng, 2016; Stevenson, 2016; Crossfield

& Kreidberg, 2017; Fu et al., 2017). This makes additional studies necessary to test such correlations, and LRG-BEASTS, along with other ground-based surveys such as the Gran Telescopio Canarias (GTC) exoplanet transit spectroscopy survey (e.g. Parviainen et al., 2016, Chen et al., 2017), ACCESS (Rackham et al., 2017), the Gemini/GMOS Transmission Spectral Survey (Huitson et al., 2017) and the VLT/FORS2 survey (e.g. Nikolov et al., 2016; Gibson et al., 2017) will expand the sample of studied planets.

While low-resolution transmission spectroscopy can probe the deeper, pressure-broadened features in the atmosphere and reveal clouds and hazes, high-resolution transmission spectroscopy can reveal narrow line features at lower pressures. Recent high-resolution results at optical wavelengths include the measurements of the wind speed on HD 189733b (Louden & Wheatley, 2015; Brogi et al., 2016), abnormal variability in the stellar line profiles of H_α near the transit of HD 189733b (Cauley et al., 2017), detections of exoplanetary sodium (e.g. Wyttenbach et al., 2015; Khalafine-jad et al., 2017; Wyttenbach et al., 2017) and analysing the affect of centre-to-limb variation on transmission spectra (Yan et al., 2017).

WASP-80b is a gas giant, with a radius of $0.986\,R_J$, orbiting a cool dwarf, with a radius of $0.593\,R_\odot$ (Triaud et al., 2013; Mancini et al., 2014). This puts it in a rare class of objects, and its large transit depth of 2.9% makes it a good candidate for transmission spectroscopy, with the extra absorption from one atmospheric scale height corresponding to 168 ppm. For this reason, there have been previous atmospheric studies of WASP-80b.

Transit photometry of WASP-80b has suggested a hazy atmosphere with no large variation with planetary radius (Fukui et al., 2014; Mancini et al., 2014; Triaud et al., 2015; Turner et al., 2017). However, Sedaghati et al. (2017a), using the VLT/FORS2, reported a detection of pressure-broadened potassium absorption. This suggests a clear and low-metallicity atmosphere, as clouds or hazes would act to mask the wings of this feature. In contrast, Parviainen et al. (2017) recently published a transmission spectrum of WASP-80b, using the GTC, which was best represented by a flat line and showed no evidence for potassium absorption.

In this chapter, I present a low-resolution transmission spectrum of WASP-80b with ACAM, as part of the LRG-BEASTS programme, which is inconsistent with broad potassium absorption.

6.2 Observations

Two transits of WASP-80b were observed on the nights of 2016 August 18 and 2016 August 21 using the ACAM instrument (Benn et al., 2008) on the William Herschel Telescope (WHT). This is the same instrument and setup I used in Chapter 5 and Kirk et al. (2017). These observations were taken by Tom Louden as part of WHT programme W/2016B/28 (PI: Kirk).

The observations were taken in fast readout mode with a smaller-than-standard window to reduce the overheads to 10 s, with exposure times of 50 s. For these transits we used a 40 arcsec wide slit as the 27 arcsec slit used in my study of HAT-P-18b was broken. I chose to use this wide slit to avoid differential slit losses between the target and comparison star. Due to the relatively sparse field, this wide slit did not cause problems with contaminating stars (Fig. 6.1, top panel). As with my study of HAT-P-18b (Chapter 5 and Kirk et al., 2017), I chose not to use an order blocking filter due to concerns this may introduce unwanted systematics. Biases, flat-fields and arc spectra were taken at the start and end of both nights.

On the first night we observed the target from airmass 1.41 to 1.16 to 1.30, with a moon illumination of 99 % at 33 degrees from the target. On the second night we observed the target from airmass 1.20 to 1.16 to 1.70, with a moon illumination of 83 % at 73 degrees from the target. The second night was affected by clouds passing overhead before and during ingress but these cleared by the time of mid-transit (Fig. 6.3).

To perform differential spectroscopy, we simultaneously observed a comparison star with a similar magnitude to that of WASP-80. The colour difference between the two stars was larger than desirable as we were limited by the length of the ACAM slit, which is 7.6 arcmin, and the fact that WASP-80 has a very close companion. This companion needed to be oriented such that its light was not blended with that of the target (Fig. 6.1). The chosen comparison star is 4.3 arcmin from the target and has a V magnitude of 12.6 and $B - V$ colour of 0.97. This compared to WASP-80's V magnitude of 11.9 and $B - V$ colour of 1.38.

6.3 Data Reduction

To reduce the data, I used the same custom PYTHON scripts as described in Chapter 3 and used in my study of HAT-P-18b (Chapter 5 and Kirk et al., 2017). For the first night, 102 bias frames were median combined to create a master bias and for the second night, 111 bias frames were median combined. To flat-field the science

images, I used a spectral sky flat normalised with a running median (as described in Chapter 3), taken at twilight with the same 40 arcsec wide slit that was used for the science images. To create the sky flat, 66 flat frames were median combined for night 1 and 86 flat frames for night 2.

To extract the spectra of both stars, I followed the process as described in Chapter 3. This involved fitting a polynomial to the locations of the traces, fitting the sky background across two regions either side of each trace, and performing normal extraction with a fixed aperture (Fig. 6.1). I experimented with the choice of extraction aperture width, background offset, background width and the order of the polynomial fitted across the background. In each case, I fitted an analytic transit light curve with a cubic polynomial to the resulting white light curve and used the extraction parameters that produced the lowest rms in the residuals. I found that a 25 pixel-wide aperture produced the lowest scatter. The pixel scale of ACAM is 0.25 arcsec pixel^{-1}.

The background was estimated by fitting a quadratic polynomial across two regions either side of the target and comparison traces after masking contaminating stars from these regions (Fig. 6.1). For the target these regions were 100 pixels wide, and offset by 20 pixels from the target aperture. For the comparison, these were again offset by 20 pixels from the comparison aperture but 150 pixels wide. The extra width was needed for the comparison as several contaminating stars had to be masked from these regions, while for the target a narrower region was used to avoid getting too close to the left-hand edge of the window (Fig. 6.1). I found that the combination of these background widths with the quadratic polynomial modelled the background variation well (Fig. 6.1). Cosmic rays were removed from the background regions by masking pixels that deviated by greater than 3σ from the median. Cosmic rays falling within the target aperture were removed in the same way as described in Chapter 3, by dividing each spectrum by a frame clean of cosmics and removing deviant points. These reductions resulted in 230 spectra for night 1 and 237 spectra for night 2.

Diagnostics of the extraction of the data for night 1 and night 2 are shown in Figs. 6.2 and 6.3. The clouds in night 2 are clearly seen as drops in transmission in the raw light curves of the target and comparison (Fig. 6.3, fifth panel). Using the raw light curves, I removed frames corresponding to drops in the transmission before further analysis. Fig. 6.6 shows the white light curve for night 2 once the cloud-affected frames had been removed, leaving 174 frames.

With spectra extracted for each frame, I aligned the spectra in pixel space and wavelength calibrated them following the procedure described in Chapter 3.
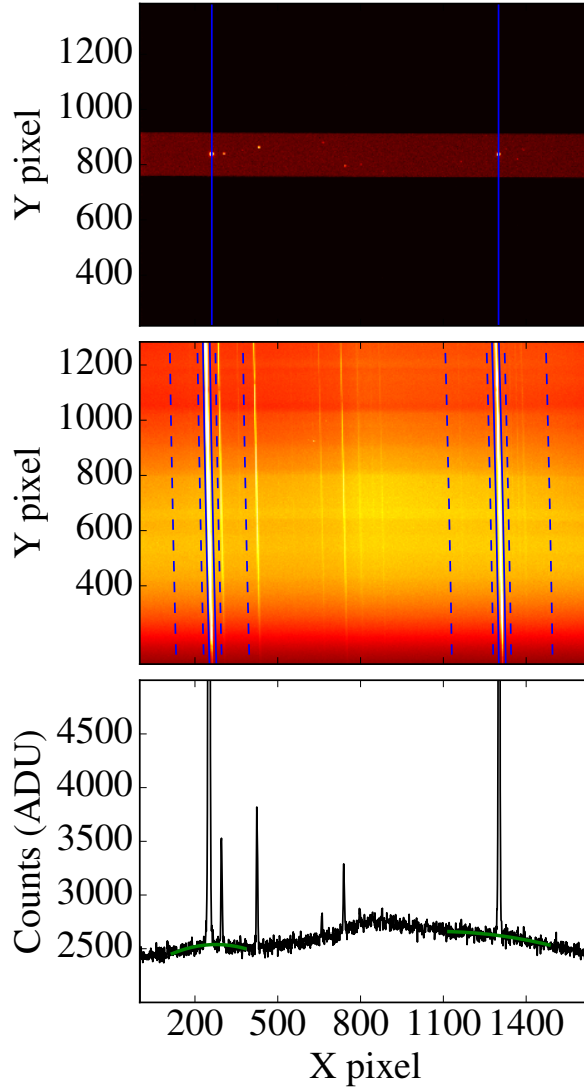
Figure 6.1: Top panel: a through-slit image of the field, taken with a 40 arcsec wide slit. The target is the star intersected by the left vertical line and the comparison is the star intersected by the right vertical line. The image has been cropped in the vertical direction for clarity. Middle panel: example science frame following bias and flat-field corrections, with blue wavelengths at the bottom of the CCD and red wavelengths at the top. The extraction apertures are shown by the solid blue lines for the target (left) and comparison (right). The dashed lines indicate the regions in which the sky background was estimated. Contaminating stars falling within these regions were masked from the background fit. The close companion to the target as noted in the text can be seen as the spectral trace immediately to the right of the target trace. This was masked from the background estimation in this region. Bottom panel: A cut along the spatial direction at a y-pixel of 740. The background polynomial fits to each star are shown by the green lines. The y-axis has been cropped for clarity.
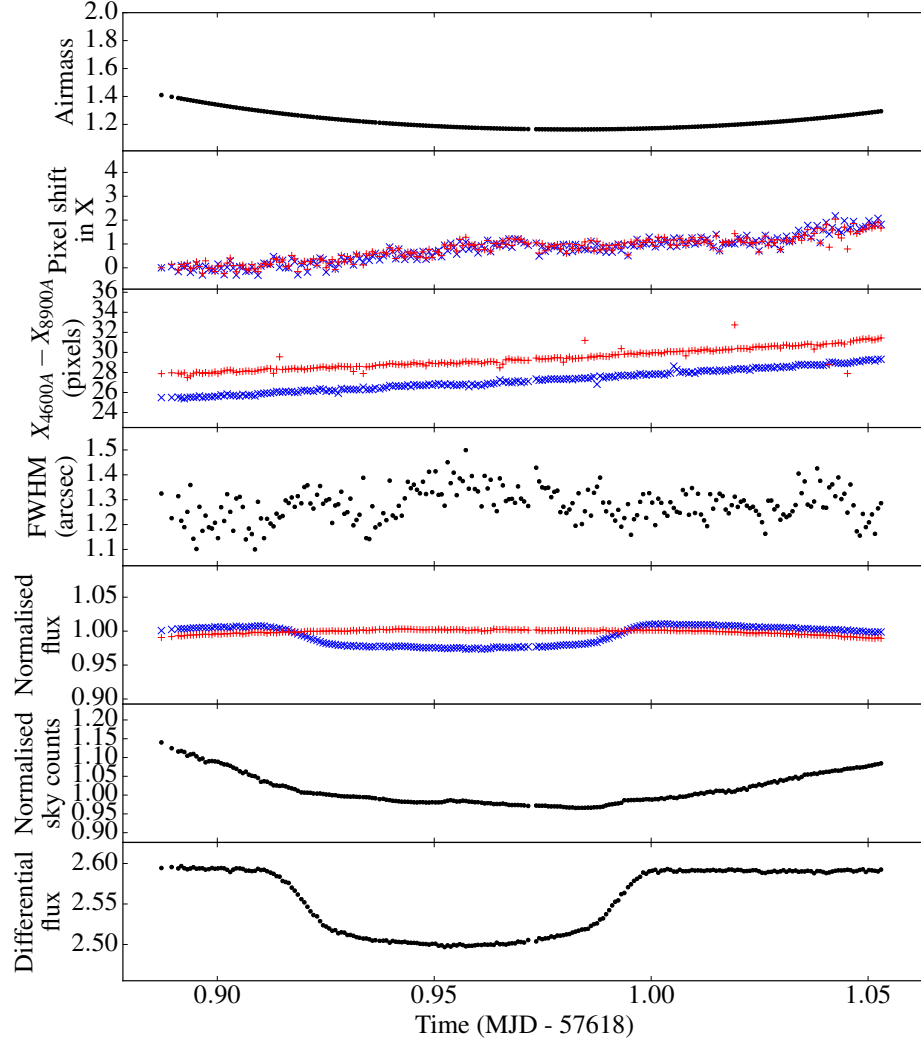
Figure 6.2: Diagnostics of night 1. Top panel: the variation of airmass across the night. Second panel: the pixel shift in x of the target (blue crosses) and comparison (red pluses) traces. Third panel: Rotation of the target (blue crosses) and comparison (red pluses) traces shown as the difference in x position at opposite ends of the traces. Fourth panel: the variation of the full width at half-maximum of the traces across the night. Fifth panel: the normalised light curves for the target (blue crosses) and comparison (red pluses). Sixth panel: the sky background across the night. Bottom panel: the differential white light curve.

157
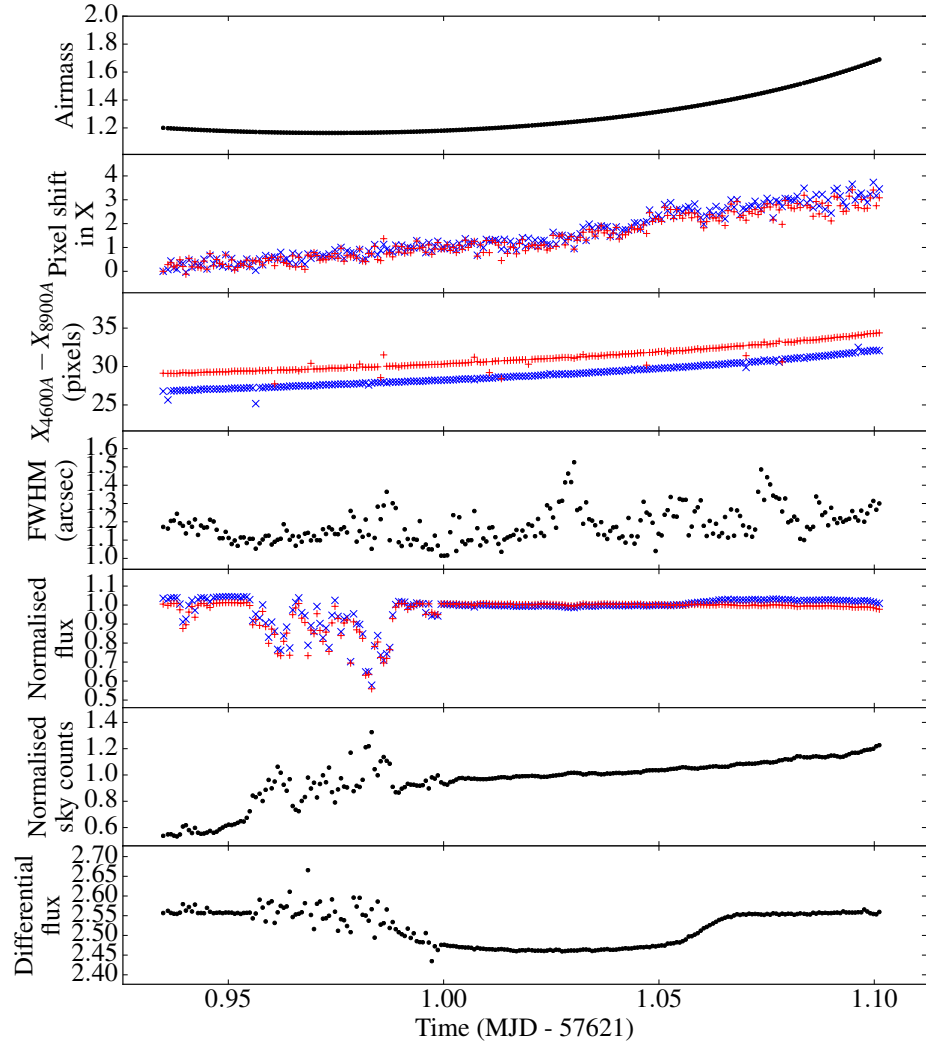
Figure 6.3: The same plots as in Fig. 6.2 but for night 2's data.

The pixel alignment was done by cross-correlating strong features in each spectral frame with a reference frame individually for the target and comparison. For each frame this resulted in pixel shift as a function of position on the CCD for both the target and comparison, which I fitted with a third-order polynomial. Using this polynomial, I then resampled the target and comparison spectra on to the grid of the reference frame, using PYSYNPHOT[1] which conserves flux. To calibrate wavelength, I constructed an arc solution for both the target and comparison. However, the arc frames were taken with a different slit to the science images and produced a wavelength calibration that was offset from the wider slit. Although this could have been corrected for in principle, the final wavelength calibration was performed using stellar and telluric lines in the spectra of the target and comparison.

Following the wavelength calibration, I divided the spectra into 35 wavelength bins (Fig. 6.4). The bluest four bins were made wider than the rest to increase the signal-to-noise ratio in these bins, given the red spectrum of the star and red sensitivity of ACAM. This led to two bins 300 Å wide, two bins 200 Å wide and 31 bins 100 Å wide. These widths were chosen as a compromise between signal to noise and resolution.

Similarly to Gibson et al. (2017), and unlike in my study of HAT-P-18b (Chapter 5), I used a Tukey window to smooth the edges of the bins to avoid potential problems arising from sharp bin edges (Fig. 6.4). The implementation of the Tukey window is shown in Fig. 6.5 and was implemented using the `tukey` function within the `scipy.signal` module (Jones et al., 2001) in PYTHON. To define the Tukey window, one needs to define the number of points falling within the tempered wings (see the example Tukey window in Fig. 6.5). For the case here, I chose to taper the 2 pixels (equivalent to ~ 6Å) at either side of the 100Å bins (Fig. 6.5). This number of points was selected so only the very edges of the windows were down-weighted.

I chose to ignore the region containing the strong telluric oxygen feature at ~ 7600 Å due to increased noise in this binned light curve. While there were other telluric features present in the spectra, these were weaker and the light curves containing these features were not significantly worse than neighbouring bins.

---

[1]https://pysynphot.readthedocs.io/en/latest/
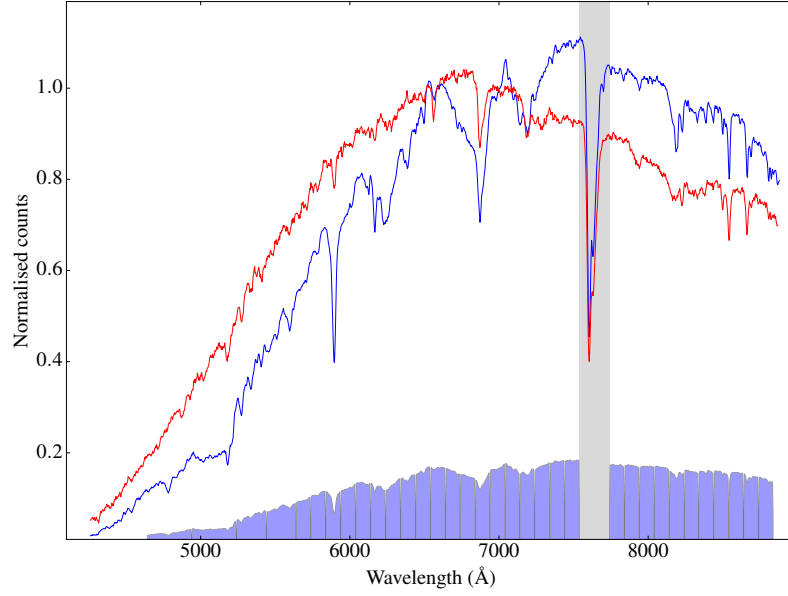
Figure 6.4: Example extracted spectra of the target (blue) and comparison (red), normalised such that $F(7000\,\text{Å}) = 1.0$. The lower filled boxes indicate the passbands used to create my wavelength-binned light curves. The grey region was excluded from my analysis.
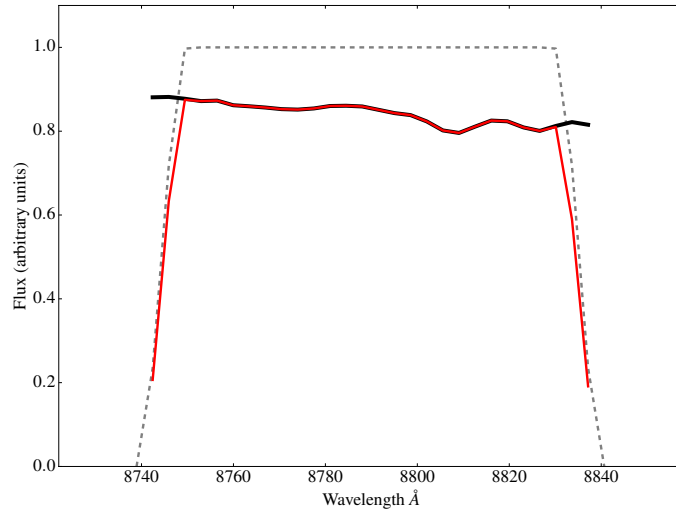
Figure 6.5: Demonstration of a Tukey window. A 100Å portion of the target's spectrum is shown in black running from $\sim 8740 - 8840\,\text{Å}$ and rescaled in y for clarity. The Tukey window is shown by the grey dashed lines. The spectrum convolved with the Tukey window is shown by the red line.

6.4 Data Analysis

6.4.1 Fitting the white light curves

Following extraction of the spectra, I created normalised, differential, white light curves for both nights by integrating the spectra for each frame. I then fitted the white light curves from both nights together with analytic transit light curves (Mandel & Agol, 2002) with a quadratic limb darkening law. Similar to my study of HAT-P-18b (Chapter 5 and Kirk et al., 2017), I chose to include a Gaussian process (GP) to model the red noise in the data and obtain robust errors for each of my fitted parameters. This was performed using the GEORGE PYTHON package (Ambikasaran et al., 2014).

GPs have been used many times in the literature and have been shown to be powerful in modelling correlated noise in data (e.g. Gibson et al., 2012a; Evans et al., 2015; Kirk et al., 2017; Louden et al., 2017). To model the red noise in the light curves, I chose to use a Matérn 3/2 kernel, defined by the length-scale τ and the amplitude a. I also included a white noise kernel, defined by the variance σ^2.

The parameters defining the model fitted to both white light curves were the inclination i, the ratio of semimajor axis to stellar radius a/R_*, the time of mid-transit T_c, the planet-to-star radius ratio R_P/R_*, the quadratic limb darkening coefficients $u1$ and $u2$ and the three parameters defining the GP, a, τ and σ^2. i and a/R_* were shared between both nights whereas the remaining parameters were not. This resulted in a total of 16 free parameters in the white light curve fits.

A prior was placed on the limb-darkening coefficients such that $u1 + u2 \leq 1$, with no priors placed on the other transit parameters. Loose, uniform priors were placed on the GP hyperparameters to encourage convergence.

To generate the starting values for the limb darkening coefficients, I made use of the limb darkening toolkit (LDTK; Parviainen & Aigrain, 2015), which uses PHOENIX (Husser et al., 2013) models to derive limb darkening coefficients and errors. For the stellar parameters, I used the values of Triaud et al. (2015). For the starting values of the GP hyperparameters, I optimised these to the out of transit data, using a Nelder-Mead algorithm (Nelder & Mead, 1965).

With these starting values, I ran a Markov chain Monte Carlo (MCMC) to the white light curves of both nights together, using the EMCEE PYTHON package (Foreman-Mackey et al., 2013). An initial run was performed with 2000 steps and 160 walkers. Following this the walkers were resampled around the result from the first run and run again for 2000 steps with 160 walkers, with the first 1000 steps discarded as burn in. The resulting fit is shown in Fig. 6.6 with the parameter values

161
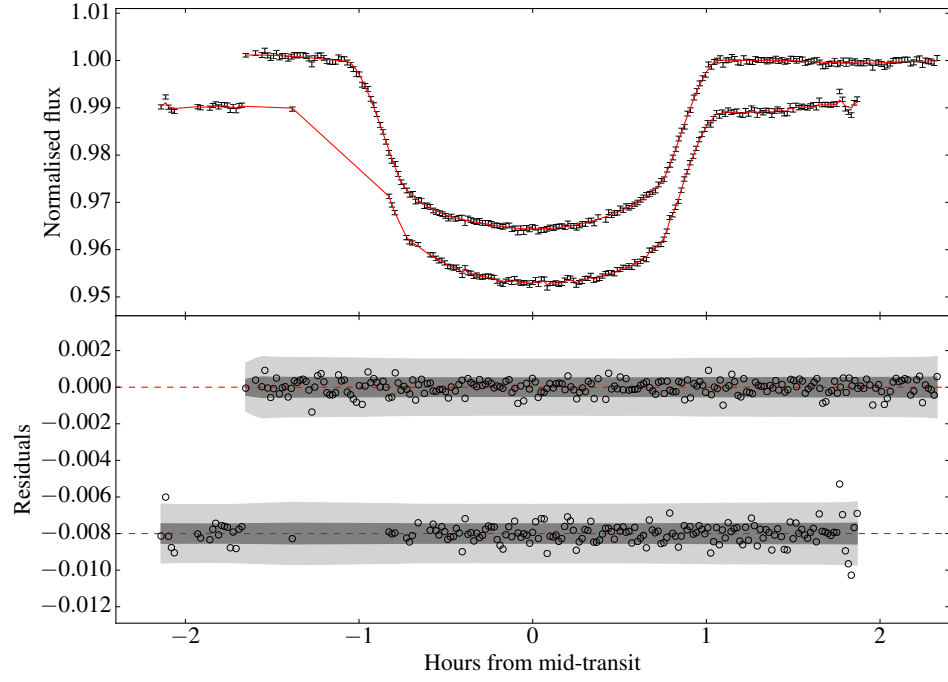
Figure 6.6: Combined fit to the white light curves of nights 1 and 2 using a GP. Top panel: the fitted model to nights 1 and 2 is shown by the red line, with an offset in y of -0.01 applied for clarity. The gaps in night 2's light curve indicate the points that were clipped due to clouds. Bottom panel: the residuals to the fits in the top panel, offset by -0.008 for clarity. The dark grey and light grey regions indicate the 1σ and 3σ confidence intervals of the GP, respectively.

in Table 6.1.

Table 6.1: Results from the fitting of the white light curves of nights 1 and 2 using a GP. Those parameters without an entry in the 'Night 2' column were shared between the nights. I also include the results from Sedaghati et al. (2017a) and the discovery paper (Triaud et al., 2013).

Parameter	Night 1	Night 2	Sedaghati et al., 2017a	Triaud et al., 2013
a/R_*	$12.66^{+0.12}_{-0.11}$	-	12.0647 ± 0.0099	12.99 ± 0.03
i (°)	$89.10^{+0.31}_{-0.22}$	-	88.90 ± 0.06	$88.92^{+0.07}_{-0.12}$
T_c (BJD)	$2457619.461954^{+0.000096}_{-0.000102}$	$2457622.529914^{+0.000135}_{-0.000131}$	$2456459.809578 \pm 0.000073$	$2456125.417512^{+0.000067}_{-0.000052}$
R_P/R_*	$0.16975^{+0.00179}_{-0.00178}$	$0.17302^{+0.00198}_{-0.00235}$	0.17386 ± 0.00030	$0.17126^{+0.00017}_{-0.00026}$
$u1$	$0.47^{+0.09}_{-0.10}$	0.48 ± 0.09	0.491 ± 0.238^a	-
$u2$	0.23 ± 0.17	$0.21^{+0.18}_{-0.17}$	0.485 ± 0.198^a	-

[a]Note that these limb darkening coefficients are calculated over a different wavelength range to the study presented here.

6.4.2 Wavelength-binned light curve fitting

The fitting of the white light curves allowed me to derive a common set of system parameters and independent common noise models from each of the nights' light curves. With these I was able to fit wavelength-binned light curves for each night, after removing the common mode systematics model from each light curve and holding the system parameters (a/R_* and i) and the time of mid-transit T_c fixed to the results from the white light curves.

I again used a GP defined by a Matérn 3/2 kernel and a white noise kernel in order to account for the noise in the data. The remaining parameters defining the wavelength-dependent models were the same as for the white light curve, resulting in six free parameters per light curve (R_P/R_*, $u1$, $u2$, a, τ and σ^2). For the limb darkening parameters, I used the log-likelihood evaluation of LDTK (Parviainen & Aigrain, 2015) in my fitting, and recovered values that were consistent with those that I measured from the white light curve.

I began by fitting an analytic transit light curve model with a cubic-in-time polynomial to remove any points that deviated by greater than 4σ from the fit. This typically clipped at most 1-2 points per light curve. As with the white light curves, I then optimised the GP hyperparameters to the out of transit data using a Nelder-Mead algorithm. Following these steps, I marginalised over all six free parameters with an MCMC. This was run for each of the light curves, with 2000 steps and 60 walkers. A second run was then performed for each, with a small perturbation added to the values resulting from the first run and again for 2000 steps with 60 walkers.

The fitted light curves are shown in Figs 6.7 and 6.8 with the resulting R_P/R_* values listed in Table 6.2 and transmission spectrum plotted in Fig. 6.9. In Table 6.2, I also include the error in the combined R_P/R_* in terms of the atmospheric scale height H, which I calculate to be $206\,\mathrm{km}$, given an equilibrium temperature of $825\,\mathrm{K}$ (Triaud et al., 2015), a planetary surface gravity of $14.34\,\mathrm{ms^{-1}}$ (Mancini et al., 2014) and assuming the mean molecular mass to be 2.3 times the mass of a proton. My results have a mean error across the transmission spectrum of 2.5 scale heights ($0.00125\,R_P/R_*$) which represents a good precision given WASP-80b's relatively small scale height.
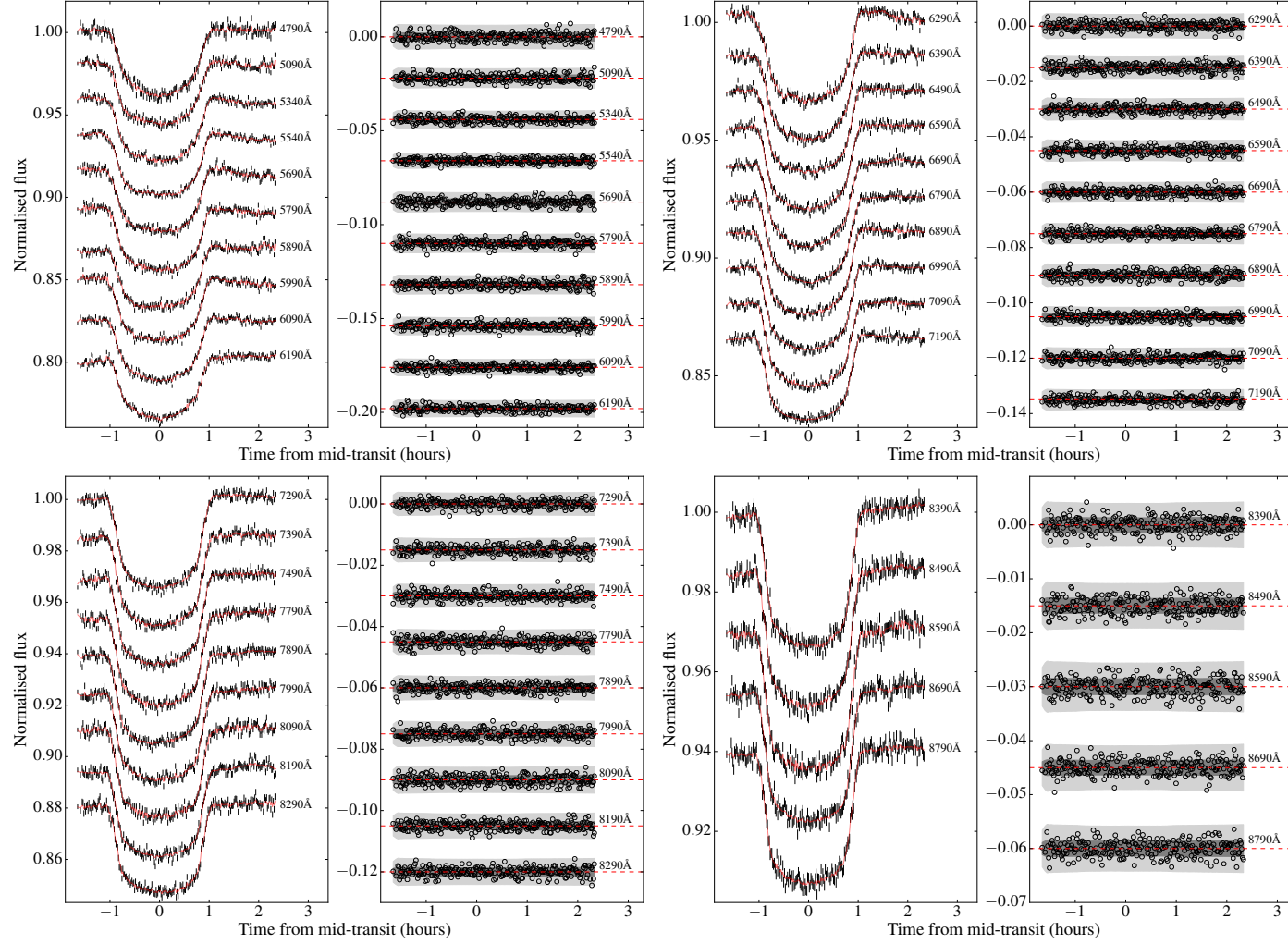
Figure 6.7: Best fitting models to the wavelength-binned light curves of night 1, which are offset for clarity. In each of the four plots, the left hand panel shows the data as black errorbars with the red line showing the best-fitting model that included both an analytic transit light curve and a GP. The right hand panels show the residuals to the best-fitting model with the dark grey and light grey shaded regions indicating the 1σ and 3σ confidence intervals.
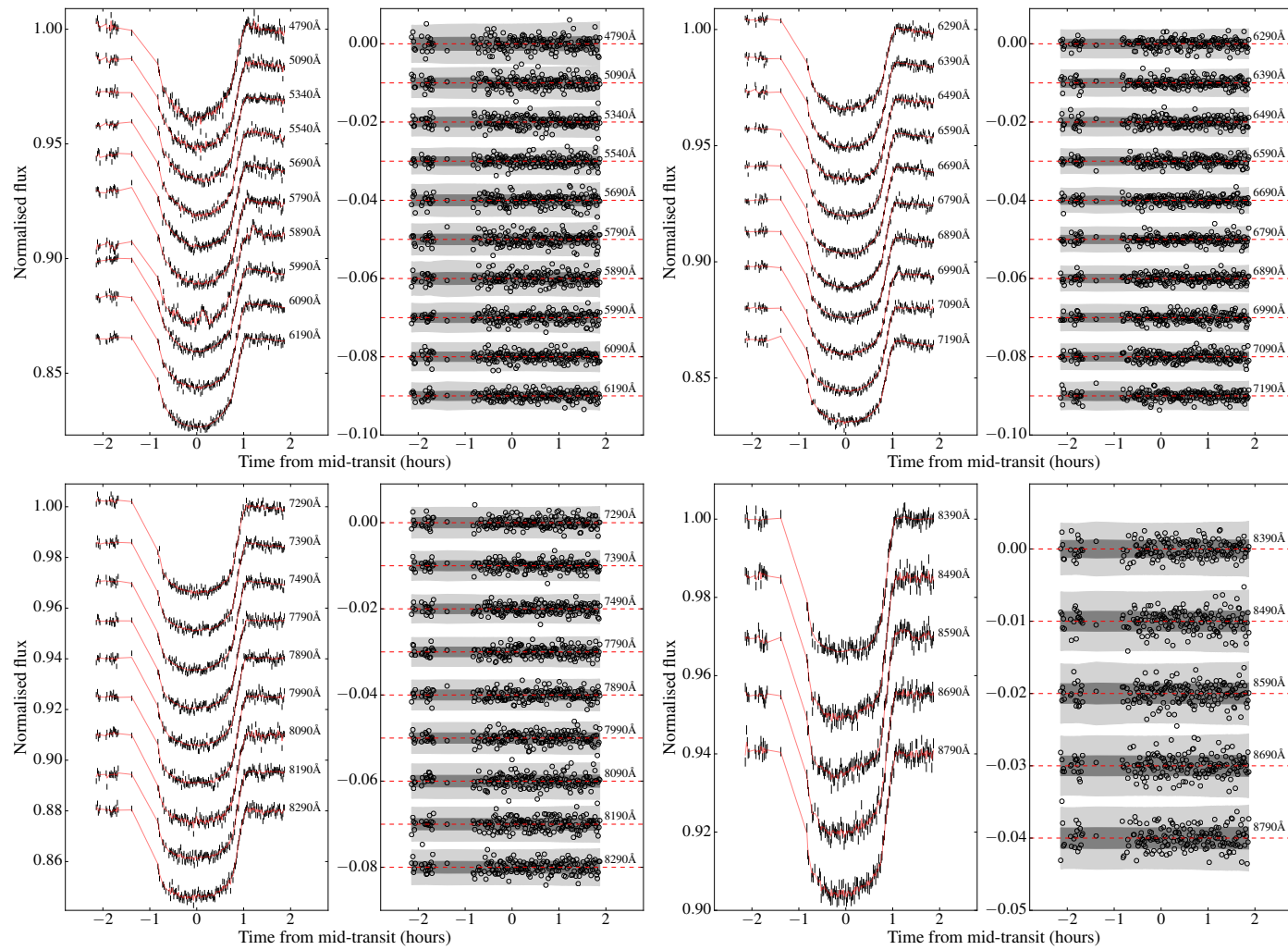
Figure 6.8: The same plots as in Fig. 6.7 but for night 2's data.

6.4.3 Transmission spectrum

Following the fitting of the wavelength binned light curves of both nights, I constructed individual transmission spectra for each night and a combined transmission spectrum from the weighted mean (Fig. 6.9). I have renormalised the transmission spectrum for the first night, given the difference in R_P/R_* from the two nights. I note that the results for night 1 in Table 6.2 are with the offset applied.

In Fig. 6.9, I also plot my combined transmission spectrum along with a clear atmosphere model, a Rayleigh scattering slope, and a flat line indicating a grey opacity source such as clouds. The clear atmosphere model was generated using EXO-TRANSMIT (Kempton et al., 2017) and assuming an isothermal temperature pressure profile of 800 K, with a metallicity of 0.1× solar.

Table 6.3 shows the goodness of fit of the model atmospheres considered. I find that while none of these models provide a satisfactory fit to the entire data set, a Rayleigh scattering slope at the equilibrium temperature of the planet (825 K) is preferred.

In Fig. 6.10 I plot my combined transmission spectrum against the spectra of both Sedaghati et al. (2017a) and Parviainen et al. (2017). My results clearly favour a transmission spectrum without broad potassium absorption and show no evidence of sodium at a resolution of 100 Å. When considering just the region that overlaps with the study of Sedaghati et al. (2017a), my transmission spectrum strongly favours a flat line over all the other models when comparing the Bayesian information criterion (which is a comparison of χ^2 for the same degrees of freedom) for this region. However, the model atmospheres considered do not produce satisfactory fits in this region (Table 6.3).

Fig. 6.10 also provides a favourable comparison between the precision of my results using a 4-metre telescope with those from 8- and 10-metre telescopes (Parviainen et al., 2017; Sedaghati et al., 2017a). It is clear that the precision of ground-based studies is limited by systematics and not photon noise.

6.5 Discussion

6.5.1 Transmission spectrum

The transmission spectrum presented here is best represented by a Rayleigh scattering slope with no broad alkali metal absorption, suggesting the presence of a high altitude haze (Fig. 6.9, Table 6.3). It is inconsistent with the previously claimed detection of broad potassium absorption by Sedaghati et al. (2017a).

Table 6.2: Results of the wavelength bin fits shown in Figs. 6.7 and 6.8.

| Wavelength bin	Night 1 R_P/R_*	Night 2 R_P/R_*	Combined R_P/R_*	Error in scale heights
$4640 - 4940$Å	$0.17509^{+0.00084}_{-0.00085}$	$0.17119^{+0.00332}_{-0.00227}$	0.17474 ± 0.00081	1.6
$4940 - 5240$Å	$0.17126^{+0.00439}_{-0.00475}$	$0.17162^{+0.00228}_{-0.00195}$	0.17147 ± 0.00192	3.7
$5240 - 5440$Å	$0.17354^{+0.00399}_{-0.00496}$	$0.17324^{+0.00295}_{-0.00280}$	0.17339 ± 0.00242	4.7
$5440 - 5640$Å	$0.17082^{+0.00694}_{-0.00716}$	$0.17326^{+0.00289}_{-0.00325}$	0.17300 ± 0.00282	5.5
$5640 - 5740$Å	$0.17594^{+0.00470}_{-0.00570}$	$0.17267^{+0.00244}_{-0.00238}$	0.17329 ± 0.00219	4.3
$5740 - 5840$Å	$0.17606^{+0.00245}_{-0.00219}$	$0.17328^{+0.00208}_{-0.00197}$	0.17441 ± 0.00153	3.0
$5840 - 5940$Å	$0.17062^{+0.00397}_{-0.00287}$	$0.17144^{+0.00295}_{-0.00323}$	0.17093 ± 0.00230	4.5
$5940 - 6040$Å	$0.17726^{+0.00320}_{-0.00346}$	$0.17249^{+0.00279}_{-0.00278}$	0.17450 ± 0.00213	4.1
$6040 - 6140$Å	$0.17050^{+0.00325}_{-0.00312}$	$0.17315^{+0.00165}_{-0.00155}$	0.17257 ± 0.00143	2.8
$6140 - 6240$Å	$0.16903^{+0.00340}_{-0.00334}$	$0.17382^{+0.00133}_{-0.00122}$	0.17318 ± 0.00119	2.3
$6240 - 6340$Å	$0.16675^{+0.00516}_{-0.00484}$	$0.17074^{+0.00219}_{-0.00180}$	0.17004 ± 0.00186	3.6
$6340 - 6440$Å	$0.16993^{+0.00199}_{-0.00244}$	$0.17290^{+0.00173}_{-0.00180}$	0.17183 ± 0.00138	2.7
$6440 - 6540$Å	$0.17216^{+0.00165}_{-0.00129}$	$0.17167^{+0.00248}_{-0.00192}$	0.17184 ± 0.00122	2.4
$6540 - 6640$Å	$0.17171^{+0.00108}_{-0.00091}$	$0.17150^{+0.00208}_{-0.00280}$	0.17167 ± 0.00092	1.8
$6640 - 6740$Å	$0.16922^{+0.00178}_{-0.00226}$	$0.17267^{+0.00236}_{-0.00217}$	0.17082 ± 0.00151	2.9
$6740 - 6840$Å	$0.17020^{+0.00154}_{-0.00156}$	$0.17261^{+0.00185}_{-0.00178}$	0.17121 ± 0.00118	2.3
$6840 - 6940$Å	$0.17209^{+0.00149}_{-0.00152}$	$0.17206^{+0.00213}_{-0.00181}$	0.17204 ± 0.00120	2.3
$6940 - 7040$Å	$0.17241^{+0.00106}_{-0.00115}$	$0.17003^{+0.00163}_{-0.00197}$	0.17182 ± 0.00094	1.8
$7040 - 7140$Å	$0.17110^{+0.00097}_{-0.00189}$	$0.17059^{+0.00057}_{-0.00051}$	0.17067 ± 0.00051	1.0
$7140 - 7240$Å	$0.17496^{+0.00165}_{-0.00146}$	$0.17119^{+0.00244}_{-0.00281}$	0.17396 ± 0.00134	2.6
$7240 - 7340$Å	$0.17356^{+0.00143}_{-0.00155}$	$0.17016^{+0.00151}_{-0.00134}$	0.17177 ± 0.00103	2.0
$7340 - 7440$Å	$0.17186^{+0.00136}_{-0.00147}$	$0.16880^{+0.00246}_{-0.00122}$	0.17060 ± 0.00113	2.2
$7440 - 7540$Å	$0.17210^{+0.00168}_{-0.00114}$	$0.16949^{+0.00150}_{-0.00127}$	0.17061 ± 0.00099	1.9
$7740 - 7840$Å	$0.17179^{+0.00153}_{-0.00163}$	$0.17007^{+0.00056}_{-0.00051}$	0.17023 ± 0.00051	1.0
$7840 - 7940$Å	$0.17365^{+0.00133}_{-0.00150}$	$0.16950^{+0.00067}_{-0.00060}$	0.17018 ± 0.00058	1.1
$7940 - 8040$Å	$0.17140^{+0.00146}_{-0.00140}$	$0.16984^{+0.00060}_{-0.00054}$	0.17003 ± 0.00053	1.0
$8040 - 8140$Å	$0.17171^{+0.00149}_{-0.00155}$	$0.17107^{+0.00063}_{-0.00058}$	0.17114 ± 0.00056	1.1
$8140 - 8240$Å	$0.17087^{+0.00208}_{-0.00234}$	$0.16923^{+0.00101}_{-0.00091}$	0.16947 ± 0.00088	1.7
$8240 - 8340$Å	$0.17223^{+0.00152}_{-0.00160}$	$0.17070^{+0.00093}_{-0.00082}$	0.17104 ± 0.00076	1.5
$8340 - 8440$Å	$0.17239^{+0.00179}_{-0.00156}$	$0.17058^{+0.00088}_{-0.00088}$	0.17095 ± 0.00078	1.5
$8440 - 8540$Å	$0.17130^{+0.00152}_{-0.00161}$	$0.17460^{+0.00113}_{-0.00091}$	0.17356 ± 0.00086	1.7
$8540 - 8640$Å	$0.17451^{+0.00188}_{-0.00196}$	$0.17353^{+0.00175}_{-0.00152}$	0.17390 ± 0.00124	2.4
$8640 - 8740$Å	$0.17040^{+0.00154}_{-0.00156}$	$0.17326^{+0.00127}_{-0.00155}$	0.17202 ± 0.00104	2.0
$8740 - 8840$Å	$0.17094^{+0.00123}_{-0.00134}$	$0.17488^{+0.00129}_{-0.00145}$	0.17284 ± 0.00094	1.8
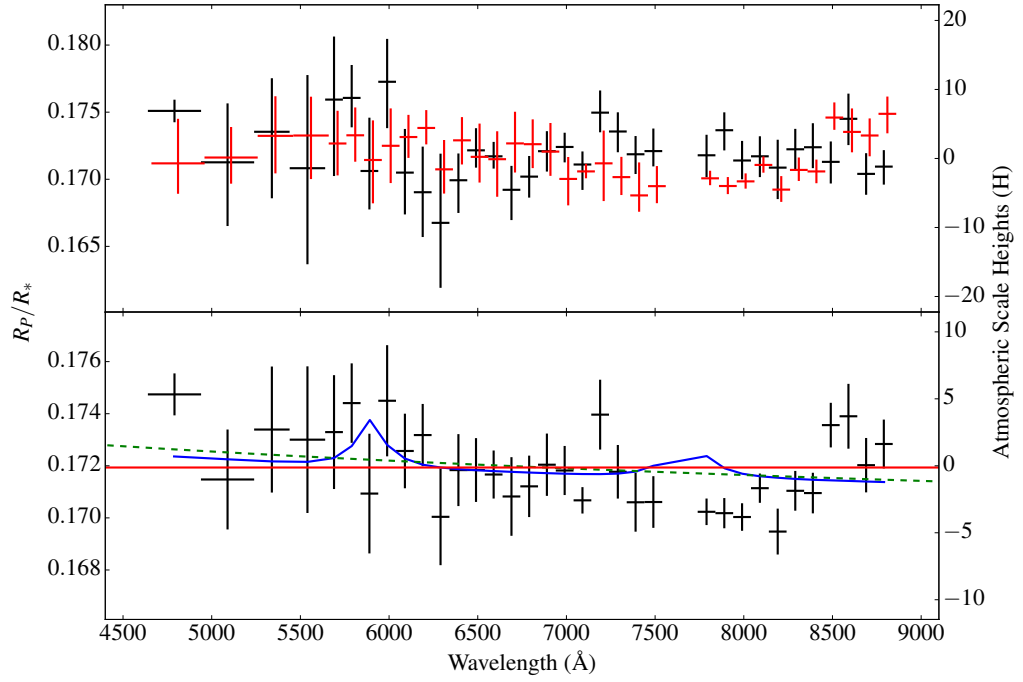
Figure 6.9: Top panel: the transmission spectra resulting from the wavelength-binned fits to the data from night 1 (black error bars) and night 2 (red error bars). The data from night 2 are offset by $+20\,\text{Å}$ for clarity. Bottom panel: the transmission spectrum resulting from the weighted mean of the two individual nights (black error bars), along with a clear, $0.1\times$ solar atmosphere (blue line) assuming an isothermal temperature pressure profile, a flat line indicating a cloudy atmosphere (red line), and a pure Rayleigh scattering model (green dashed line) at the equilibrium temperature of the planet ($825\,\text{K}$).

Table 6.3: Goodness of fit of different model atmospheres to both my whole transmission spectrum ($4640 - 8840\text{Å}$, with 33 degrees of freedom) and the region overlapping with the study of Sedaghati et al. (2017a) ($7440 - 8840\text{Å}$, with 11 degrees of freedom). The 95% confidence limit for a χ^2 distribution with 33 degrees of freedom is 47.4 and for 11 degrees of freedom is 19.7. The models preferred in each region are shown in boldface.

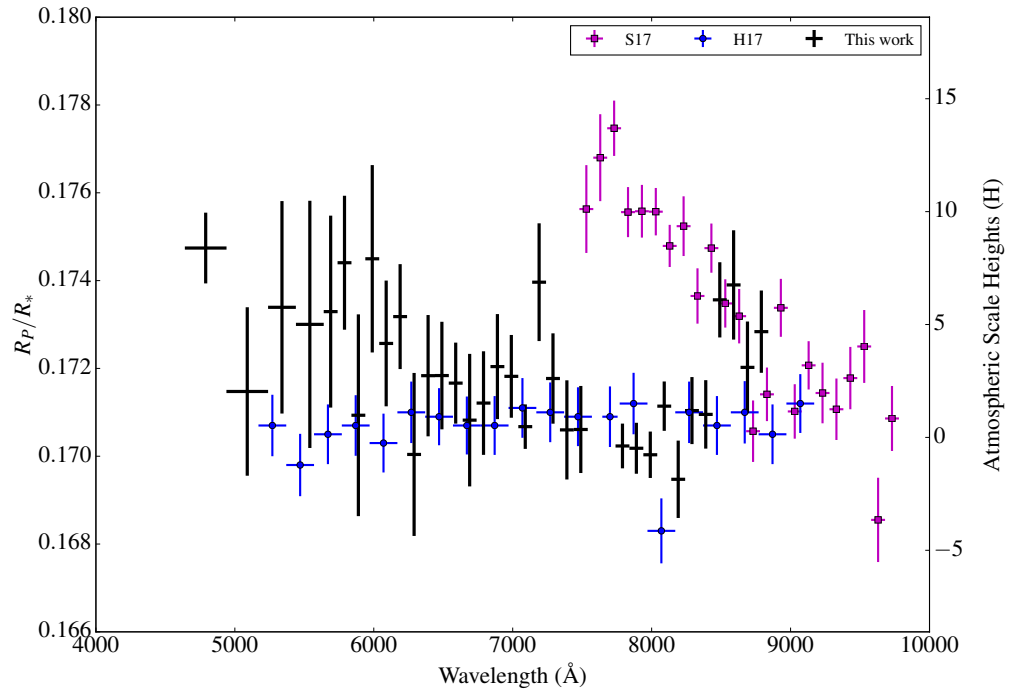|	$4640 - 8840\text{Å}$ χ^2 (χ^2_ν)	$7440 - 8840\text{Å}$ χ^2 (χ^2_ν)
Flat line	74.9 (2.27)	**32.1 (2.92)**
Rayleigh (T_{eq})	**68.5 (2.08)**	39.5 (3.59)
$0.1\times$ solar, clear	81.6 (2.47)	50.6 (4.60)
$1\times$ solar, clear	71.6 (2.17)	40.4 (3.67)

169

Figure 6.10: My two night combined transmission spectrum (black error bars) along with the transmission spectra of Sedaghati et al. (2017a) (magenta squares) and Parviainen et al. (2017) (blue circles).

In the study of Sedaghati et al. (2017a), the authors interpreted their transmission spectrum as evidence for the pressure-broadened wings of the potassium feature. This would suggest a clear atmosphere such as in WASP-39b (Fischer et al., 2016; Nikolov et al., 2016; Sing et al., 2016), as clouds and hazes have the effect of masking the wings of the alkali features (e.g. Mallonn et al., 2015b; Nikolov et al., 2015; Kirk et al., 2017; Louden et al., 2017). In this case, we might expect to see the broad wings of the sodium feature but I see no evidence for these either, nor do I detect the core at a resolution of 100 Å (Fig. 6.9).

Alkali chlorides are expected to condense at around the temperature of WASP-80b (825 K; Mancini et al., 2014), which reduces the sodium and potassium in the upper atmosphere and gives rise to Rayleigh scattering (Wakeford & Sing, 2015). This scenario is consistent with my transmission spectrum as I do not detect broad sodium or potassium absorption, and find an atmosphere best represented by Rayleigh scattering (Fig. 6.9, Table 6.3). However, we might expect alkali chlorides to produce a steeper slope than I observe (Wakeford & Sing, 2015).

Given that WASP-80b's equilibrium temperature (825 K) is very similar to HAT-P-18b's (852 K), it is interesting that in both cases I find an atmosphere most consistent with a Rayleigh scattering haze. Its temperature again implies that CH_4 is the predominant carbon-bearing species and therefore could again be hosting optically thick soot precursors in its atmosphere (Morley et al., 2015). My results for WASP-80b are therefore again in agreement with the findings of Heng (2016), Stevenson (2016), Crossfield & Kreidberg (2017) and Fu et al. (2017) that cooler planets are more likely hazy.

Despite the presence of a haze, it can still be possible to detect the narrow line core arising from higher altitudes in the planetary atmosphere such as has been done for WASP-52b (Chen et al., 2017), HD 189733b (Pont et al., 2008; Sing et al., 2011a; Huitson et al., 2012; Pont et al., 2013; Louden & Wheatley, 2015) and HD 209458b (Charbonneau et al., 2002; Sing et al., 2008a,b; Snellen et al., 2008; Langland-Shula et al., 2009; Vidal-Madjar et al., 2011; Deming et al., 2013); however, these detections are often made in much narrower bins than the 100 Å bins used here. For this reason, I cannot rule out the presence of the narrow core of the sodium feature despite the absence of the broad wings.

While I cannot be certain what the cause of the discrepancy is between my results and those of Sedaghati et al. (2017a), it is perhaps plausible that there were residual systematics in the VLT/FORS data of Sedaghati et al. (2017a) associated with the Longitudinal Atmospheric Dispersion Corrector (LADC; Avila et al., 1997) despite the authors' thorough attempts to account for these. The observations of

Sedaghati et al. (2017a) were taken before the replacement of the anti-reflective coating of the LADC which was known to cause systematic errors (Boffin et al., 2015). I note that more recent observations by these authors, also using VLT/FORS, resulted in a detection of TiO in the atmosphere of WASP-19b but with the LADC left in park mode (Sedaghati et al., 2017b). I am also encouraged that ACAM is optically simpler than FORS with no atmospheric dispersion corrector.

In addition to possible systematics in the VLT/FORS data of Sedaghati et al. (2017a), I consider the effects of stellar activity in the next section.

6.5.2 Stellar activity

Despite the lack of rotational modulation in its light curve, WASP-80 has a log R'_{HK} of -4.495, suggesting the star has a relatively high level of chromospheric activity (Triaud et al., 2013; Mancini et al., 2014). However, WASP-80 has a spectral type between K7V and M0V (Triaud et al., 2013) and when comparing this value to a sample of K stars with measured log R'_{HK} values, we see that WASP-80's log R'_{HK} is fairly typical (Fig. 6.11, values from Martínez-Arnáiz et al., 2010).

Given WASP-80's chromospheric activity, it is puzzling that no star spot crossings have been observed in any of the studies of WASP-80 to date. This would suggest that either WASP-80 is spot free or that there are a large number of spots continuously on its surface, at latitudes not occulted by the planet (as postulated by Mancini et al., 2014).

The orbit of WASP-80b is aligned with the stellar spin axis (Triaud et al., 2015) and therefore crosses the same latitudes of the host during each transit. I calculate WASP-80b to have an impact parameter of 0.2, therefore it is plausible that there are spots at higher latitudes that are not occulted during transit. High latitude spots have been observed in a number of stars (e.g. Strassmeier, 2009), although early M dwarfs have been observed to show spots more uniformly distributed in longitude and latitude (e.g. Barnes & Collier Cameron, 2001; Barnes et al., 2004), making the prospect of numerous unocculted spots less likely in this case.

Unocculted spots can induce blueward slopes in transmission spectra (McCullough et al., 2014), mimicking a Rayleigh scattering signature. Unocculted spots cannot be strongly affecting my transmission spectrum as I do not detect a clear scattering signature, and since we do not know whether the star has evenly distributed spots, or has no spots, I have chosen not to adjust the transmission spectra presented here. However, I consider whether they could cause the discrepancy between the transmission spectrum presented here and that of Sedaghati et al. (2017a).

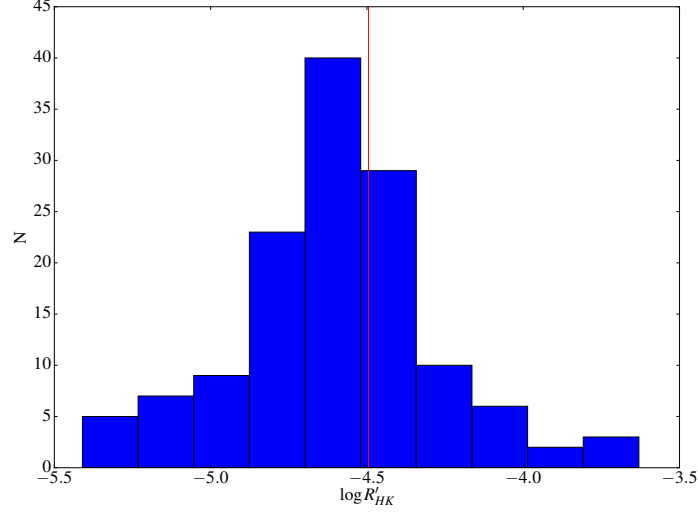Following the formalism of McCullough et al. (2014), who considered the

Figure 6.11: A plot of $\log R'_{\mathrm{HK}}$ values of K dwarfs displaying chromospheric features from the sample of Martínez-Arnáiz et al. (2010), who gathered chromospheric activity measurements of 371 F to K stars within 25 pc. The red line indicates WASP-80's $\log R'_{\mathrm{HK}}$.

effects of unocculted spots on the transmission spectrum of HD 189733b, it can be shown that the apparent depth $(\tilde{R}_P/\tilde{R}_*)^2$, can be related to the measured depth $(R_P/R_*)^2$, the filling factor of spots on the projected stellar surface f, and the fluxes of the star and a spot at a particular wavelength, $F_\lambda(spot)$ and $F_\lambda(star)$, through

$$\left(\frac{\tilde{R}_P}{\tilde{R}_*}\right)^2 = \frac{(R_P/R_*)^2}{1 - f(1 - F_\lambda(spot)/F_\lambda(star))}, \tag{6.1}$$

which can be rearranged to find the filling factor, f. Using ATLAS9 (Kurucz, 1993) models of a star with $T_{\mathrm{eff}} = 4000\,\mathrm{K}$, [Fe/H] $= 0.0$ and $\log g = 4.5$, and a spot with a temperature of $3500\,\mathrm{K}$ (as consistent with observations of spots on stars of similar spectral type; Berdyugina, 2005), I find that the difference in transit depth between my shortest and longest wavelength observations would require a filling factor of 2.6%. The slope found by Sedaghati et al. (2017a) would require unocculted spots with a filling factor of 16%. If these transmission spectra were due to unocculted spots, then the filling factor of such spots would have to change by at least 13%. Since photometric modulations of this order are ruled out by the WASP data I conclude that stellar activity cannot be the cause of the discrepancy between my results and those of Sedaghati et al. (2017a).

6.6 Conclusions

In this chapter I presented a ground-based optical transmission spectrum of the hot Jupiter WASP-80b. My transmission spectrum is best represented by a Rayleigh scattering slope, suggesting the presence of a high altitude haze in the atmosphere of WASP-80b. I see no evidence for the broad wings of the potassium feature as claimed previously by Sedaghati et al. (2017a) nor sodium at a resolution of 100 Å. Instead, my transmission spectrum is in better agreement with those of Fukui et al. (2014), Mancini et al. (2014), Triaud et al. (2015), Turner et al. (2017) and Parviainen et al. (2017).

Stellar activity cannot be the cause of the discrepancy between my results and those of Sedaghati et al. (2017a) due to WASP-80's lack of photometric modulation (Triaud et al., 2013). Instead, it is possible that there were residual systematics in the VLT/FORS data of Sedaghati et al. (2017a), perhaps related to the LADC, despite the authors' thorough attempts to account for these.

This was the third paper in the LRG-BEASTS programme following the detection of a Rayleigh scattering haze in HAT-P-18b (Chapter 5 and Kirk et al., 2017) and clouds in WASP-52b (Louden et al., 2017). LRG-BEASTS is demonstrating that 4-metre class telescopes can provide transmission spectra with precisions comparable to 8- and 10-metre class telescopes, and are capable of detecting and ruling out model atmospheres. This work also highlights the importance of independent and repeat studies of hot Jupiters. These studies as part of our larger survey will help to shed light on the prevalence, and physical origins, of clouds and hazes in hot Jupiter atmospheres by increasing the sample of studied hot Jupiters.

Chapter 7

Conclusions and future work

In this thesis I presented ground-based transmission spectra of WASP-52b, HAT-P-18b and WASP-80b. In all three cases I found atmospheres that demonstrate the presence of aerosols, highlighting that such features are very common in exoplanet atmospheres.

These results have formed part of the LRG-BEASTS survey which will characterise at least 15 exoplanets (see Chapter 2) with the goal of testing the correlations between aerosols and fundamental planetary parameters that have been proposed by a number of authors (Heng, 2016; Stevenson, 2016; Crossfield & Kreidberg, 2017; Fu et al., 2017). The combined results from the LRG-BEASTS survey are shown in Fig. 7.1, which are individually published by Kirk et al. (2016, 2017, 2018) and Louden et al. (2017).

In the case of WASP-52b the aerosols were grey, with no wavelength dependence (Chapter 4). In this case, the aerosols could be causing Mie scattering which is featureless over the wavelengths studied. This would suggest the presence of a condensate cloud as we might expect a photochemical haze to produce a scattering signature in the optical, much like in the well-studied atmosphere of HD 189733 b (Pont et al., 2008; Sing et al., 2011a; Pont et al., 2013). In my study of WASP-52b I also saw strong evidence for faculae on the host star that were occulted during transit. The contrasts I derived were consistent with predictions for solar type stars (Norris et al., 2017). This work highlighted the effects that stellar activity can have on the derivation of transit parameters, as well as how planet transits can be used to better understand stars.

For HAT-P-18b, I detected a slope extending across optical wavelengths and rising towards the blue (Chapter 5). This slope was very well matched by a Rayleigh scattering slope at the equilibrium temperature of the planet (852 K, Hartman et al.,
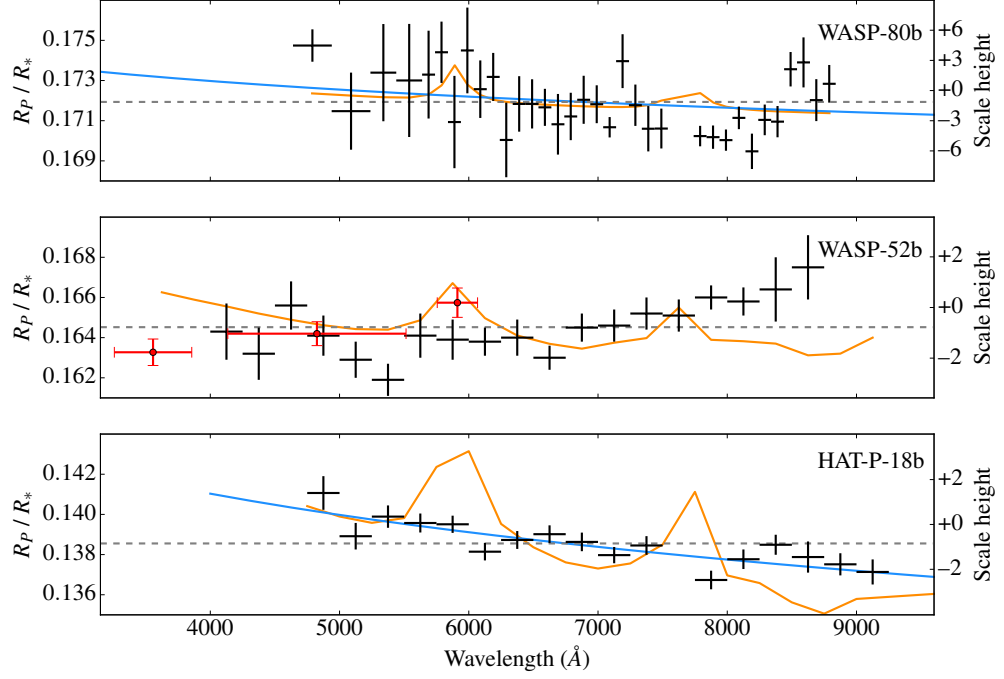
Figure 7.1: The current results from LRG-BEASTS, all using WHT/ACAM. I ruled out a previously claimed potassium detection in WASP-80b's atmosphere, instead finding a hazy atmosphere consistent with GTC results (Chapter 6, Kirk et al., 2018); found clouds in the atmosphere of WASP-52b using ULTRACAM (red points resulting from my facula model in Chapter 4, Kirk et al., 2016), which was consistent with our later ACAM study (black points, Louden et al., 2017); and detected Rayleigh scattering by a high altitude haze in HAT-P-18b's atmosphere (Chapter 5, Kirk et al., 2017), which was only the second discovery of a Rayleigh slope in a hot Jupiter atmosphere from the ground. This figure demonstrates our ability to obtain a precision of $1H$, which allows us to differentiate atmospheric models with significance.

2011). Since I did not detect the broad wings of the sodium and potassium features, I concluded that a high altitude haze was present. This would give rise to the slope while masking the pressure-broadened alkali features, again similar to the case of HD 189733 b (Pont et al., 2008; Sing et al., 2011a; Pont et al., 2013). While my conclusion of a haze implies the aerosols are photochemically produced, it is also possible that the aerosols could be condensates. Alkali chlorides condense at temperatures near HAT-P-18b's equilbrium temperature and can give rise to Rayleigh scattering slopes while also decreasing atomic sodium and potassium abundances (Wakeford & Sing, 2015). This is consistent with what I see (Chapter 5) although alkali chlorides might produce a steeper slope than I detect (Wakeford & Sing, 2015).

Similar to HAT-P-18b, the atmosphere of WASP-80b favoured a Rayleigh scattering slope over a flat transmission spectrum, although it did not provide a good fit to the entire data set (Chapter 6). However, I was able to rule out the presence of a previously reported potassium feature, highlighting the need for repeat studies of planets. Instead, I found an atmosphere most consistent with aerosols. As WASP-80b's equilbrium temperature is very similar to HAT-P-18b (825 K compared to 852 K), it is possible to make similar arguments for this planet's atmosphere. WASP-80b does, however, have a higher surface gravity than HAT-P-18b, with a $\log g$ of 3.178 (Triaud et al., 2013), compared to HAT-P-18b's $\log g$ of 2.69 (Hartman et al., 2011). Lower gravities lead to lower settling rates of aerosols and lower sedimentation efficiencies. This means that clouds can form at higher layers with a larger vertical extent (e.g. Gao et al., 2018), such as was inferred for Kepler-7b (Demory et al., 2013), which has a $\log g$ of 2.616 (Southworth et al., 2012). If surface gravity were the dominant factor here, we might expect WASP-80b to display less aerosols in its optical transmission spectrum, however, this is not the case and instead I conclude that temperature is dominating.

Indeed, in recent studies of samples of exoplanets, temperature does seem to be important, with several findings suggesting that hotter planets are more likely to be cloud-free (Heng, 2016; Stevenson, 2016; Crossfield & Kreidberg, 2017; Fu et al., 2017). This is consistent with what I find, as both HAT-P-18b and WASP-80b are near the regions designated as being more likely to be hazy by both Stevenson (2016) ($T_{eq} = 700$ K) and Crossfield & Kreidberg (2017) ($T_{eq} = 850$ K), although Fu et al. (2017) saw no evidence for such a cut-off. All these studies have noted the need for a larger sample size, with Fu et al. (2017) suggesting that more planets at the very hot end need to be observed. LRG-BEASTS will add to the number of studied planets, while covering a very broad range of parameter space ($825 \leq T_{eq} \leq 2489$ K, $0.26 \leq M_P/M_J \leq 1.47$, $2.50 \leq \log g \leq 3.24$). This broad coverage of parameter space

will allow us to test these correlations, while the hotter planets will also enable us to search for the presence of TiO and VO.

TiO and VO are important opacity sources in late-type stars and were predicted to cause temperature inversions in hot Jupiters (Fortney et al., 2008). Due to the lack of detection of these molecules for several years, the role of TiO and VO in hot Jupiter atmospheres was called into question. However, more recent studies have revealed TiO in transmission (WASP-19b, Sedaghati et al., 2017b; WASP-33b, Nugroho et al., 2017) and possibly even VO in emission (WASP-121b, Evans et al., 2017). The presence of TiO and VO has also been inferred in the relatively cool WASP-127b ($T_{eq} = 1400$ K, Palle et al., 2017). This has reopened the debate and our hotter targets will allow us to probe for the presence of these molecules and enable us to make predictions as to the thermal structure of their atmospheres.

7.1 Future work

My immediate goal is the continuation of the LRG-BEASTS survey (as presented in Chapter 2), which has acquired data for 15 planets to date. Table 7.1 includes a list of all planet transits observed to date with their associated parameters, as listed on TEPCAT[1] (Southworth, 2011), and the quality of the nights' data. The quality of the data is represented by the rms of the residuals of a fit to the white light curve for each night. The fits were performed with a quadratically limb-darkened analytic transit light curve combined with a cubic in time polynomial to model red noise. This is not necessarily the best model to use for each night and so I expect the rms of the residuals to improve with a fuller analysis and following the optimal choice of aperture width and background polynomial order. However, the rms values presented in Table 7.1 are useful to get an idea of the quality of the data.

I intend to use the methods described in Chapter 3 to perform a detailed reduction and analysis of these data in order to produce transmission spectra for every object observed. However, as can be seen in the tabulated rms values in Table 7.1, it might be difficult to reach the precision required for some of these nights. Where a single transit has been observed I aim to obtain a second transit that will help to decipher atmospheric signals from systematic noise.

---

[1]http://www.astro.keele.ac.uk/jkt/tepcat/

Table 7.1: The LRG-BEASTS sample of 15 exoplanets that have had at least one transit observed as of April 2018. The three planets that have been published are in boldface. Other than for values marked with an asterisk, the values in the 'rms of the residuals' column are following an initial reduction and fit and are likely to improve with a more detailed analysis. Those marked with an asterisk are following fitting with a Gaussian process. Objects with two transits observed have the rms for each night. The parameters for each planet are from TEPCAT (Southworth, 2011).

Target	No. transits	Expected trans. signal (ppm)	Equilibrium temp. (K)	Planet radius (R_J)	Planet mass (M_J)	Surface gravity ($cm\,s^{-2}$)	Host V mag.	rms of residuals (ppm)
HAT-P-18b	**1**	**297**	**841**	**0.947**	**0.196**	**5.42**	**12.8**	**280*,[a]**
HAT-P-44b	1	283	1108	1.242	0.352	5.62	13.21	549
HATS-46b	1	267	1054	0.903	0.173	5.1	13.63	438
WASP-4b	1	172	1673	1.364	1.249	16.64	12.46	797
WASP-21b	1	246	1333	1.162	0.276	5.07	11.59	416
WASP-23b	1	97	1152	1.067	0.917	19.7	12.68	694
WASP-25b	1	193	1210	1.247	0.598	9.54	11.85	[b]
WASP-39b	1	452	1116	1.27	0.28	4.26	12.1	362
WASP-41b	1	111	1242	1.178	0.977	17.45	11.64	636
WASP-52b	**2**	**405**	**1315**	**1.253**	**0.434**	**6.85**	**12.22**	**237, 367*,[c]**
WASP-69b	1	301	963	1.057	0.26	5.32	9.87	1226
WASP-74b	2	129	1910	1.56	0.95	8.91	9.76	485, 1161
WASP-80b	**2**	**168**	**825**	**0.986**	**0.562**	**14.34**	**11.87**	**418, 756*,[d]**
WASP-94Ab	1	314	1604	1.72	0.452	3.48	10.06	296
WASP-103b	2	149	2489	1.646	1.47	14.34	12.39	603, 908

[a] Fits presented in Chapter 5

[b] Transit requires thorough analysis to remove systematics

[c] Fits presented in Louden et al. (2017)

[d] Fits presented in Chapter 6

In the analyses of these data, I would like to better understand the ACAM and EFOSC2 instruments to improve the precision of my results. In particular, I aim to measure the effects of non-linearity and flat-fielding on my results. As mentioned in Chapter 3, ACAM is linear to within 0.7% over its full range. As a quick initial test, and more recently than the analysis of the results presented in Chapters 5 and 6, I re-reduced the HAT-P-18b data but with the inclusion of a non-linearity correction. To perform this correction, I multiplied each pixel by the linearity correction from the ACAM webpages, which takes the form

$$F_{corr} = ((-0.007/65000)F_{raw} + 1)F_{raw} \qquad (7.1)$$

where F_{corr} is the linearity-corrected flux and F_{raw} is the raw (recorded) flux. Using equation 7.1 I re-ran the reduction with the same parameters as in the analysis presented in Chapter 5. I also performed a reduction with a flat field using the median filtering technique presented in Chapter 3. The ratios of the white light curves, with and without flat fielding and with the linearity correction are shown in Fig. 7.2. This figure shows that the flat field has very little effect with an amplitude in the flux ratio of < 100 ppm but that the non-linearity effect is larger and correlates with the FWHM of the spectral trace.

As a quick test of the colour dependence of the effects of non-linearity and the flat field, I divided the spectra into bins ~ 250 Å-wide in a wavelength range of ~ 4750 to 8500 Å. This was done before resampling of the target and comparison's spectra, and before wavelength calibration, as it was designed as a quick test. The flux ratios between the light curves with and without the flat and with the linearity correction were then calculated. The colour-dependent effects of the flat are shown in Fig. 7.3 and the linearity correction in Fig. 7.4. These tests again showed the flat field made a small difference (Fig. 7.3), although the effect is larger towards the blue causing ramps that rise in some wavelength bins and drop in others. The effect of non-linearity is larger towards redder wavelengths, presumably as the counts are higher here (Fig. 7.4). These effects will be fully explored in future analyses by comparing transmission spectra resulting from each reduction. It is not obvious what effect, if any, these features will have on the transmission spectrum, especially because they may already be adequately removed with a Gaussian process. By better understanding these processes in future work I hope to be able to further improve the precision of my transmission spectroscopy measurements.

These initial tests, combined with the contribution to the error budget from photon noise, read noise and scintillation presented in Chapter 3, indicate that a small defocus is a good observational strategy. When defocussing, we require a
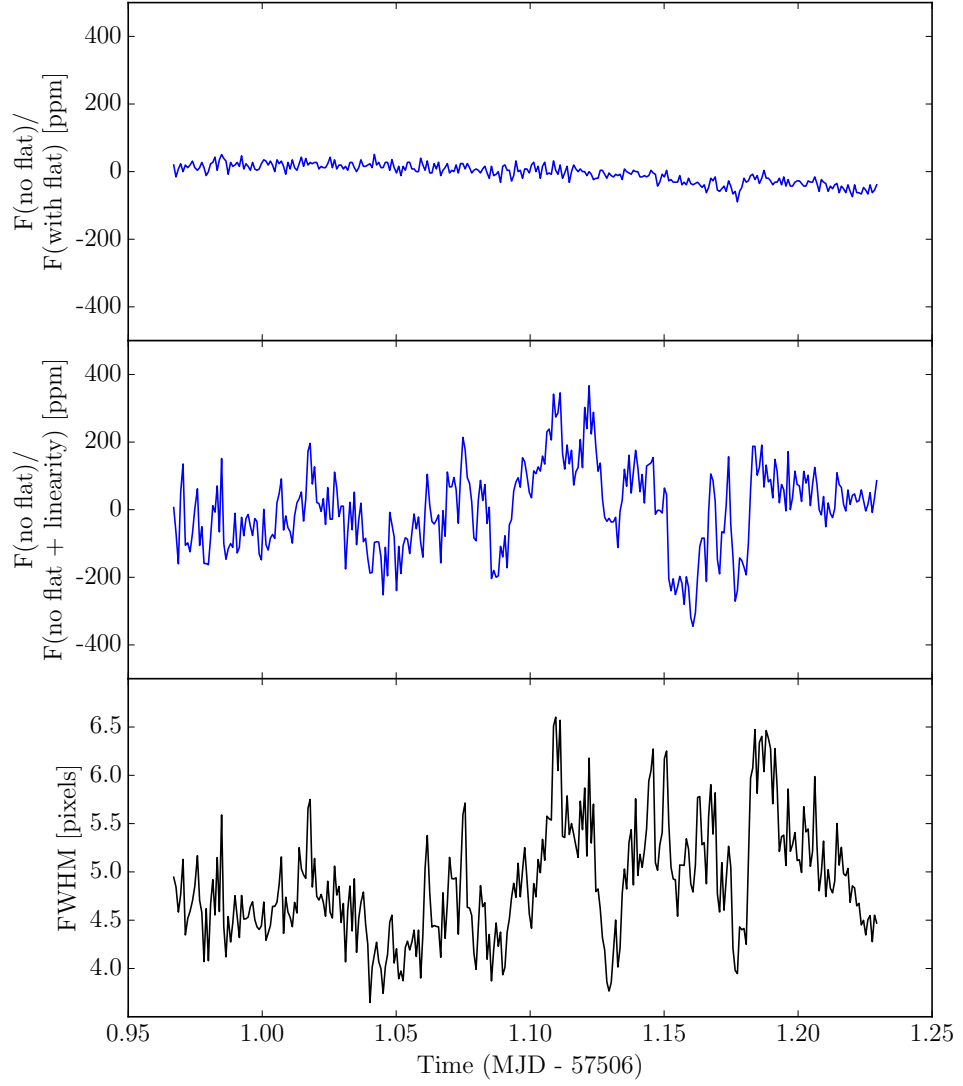
180

Figure 7.2: Top panel: the white light curve for HAT-P-18b without a flat field correction divided by the white light curve *with* a flat field correction. Middle panel: the white light curve for HAT-P-18b without a flat field correction divided by the white light curve without a flat field correction but *with* a non-linearity correction. Bottom panel: the measured variation in FWHM of the trace in pixels.
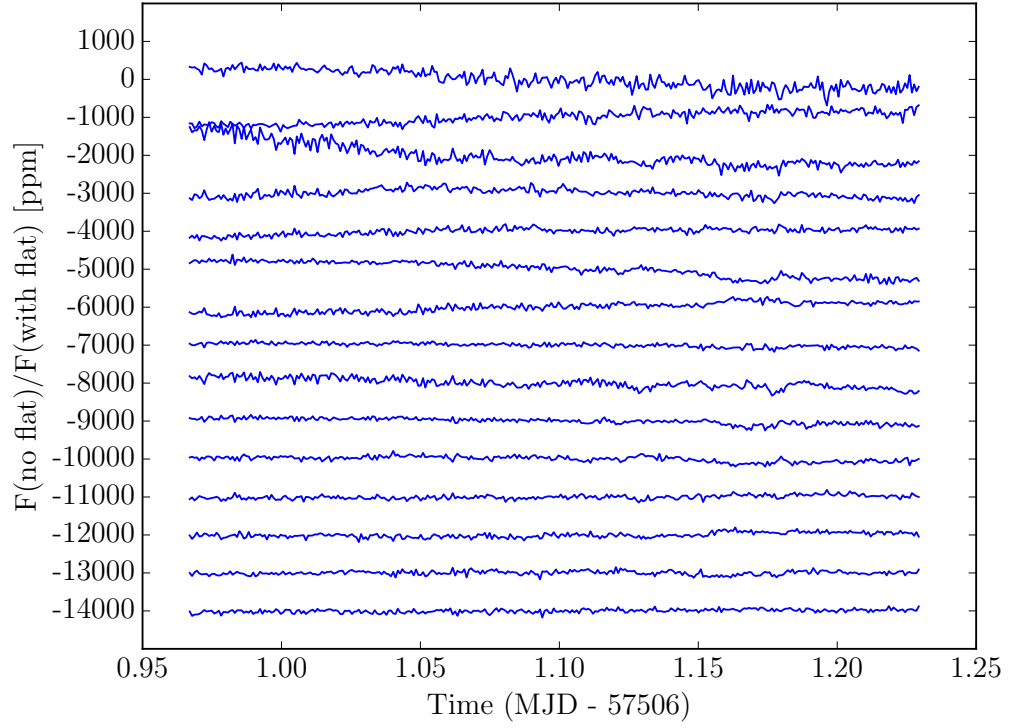
181

Figure 7.3: Wavelength binned light curves for HAT-P-18b without a flat field correction divided by wavelength binned light curves *with* a flat field correction. Each light curve is offset by 1000 ppm for clarity. The wavelength bins are $\sim$250 Å wide and run from $\sim 4750 - 8500$ Å.
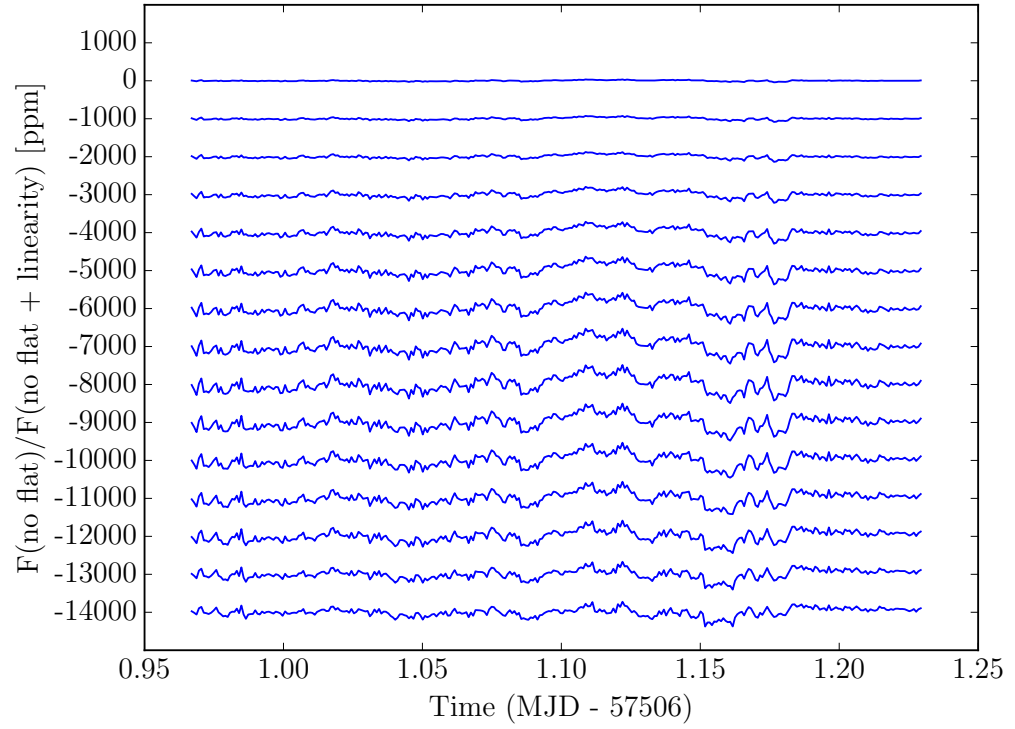
Figure 7.4: Wavelength binned light curves for HAT-P-18b without a flat field correction divided by wavelength binned light curves without a flat field correction but *with* a linearity correction. Each light curve is offset by 1000 ppm for clarity. The wavelength bins are ~250 Å wide and run from ~ 4750 − 8500 Å.

183

larger aperture, causing the photon noise from the sky and the read noise to increase. However, seeing variations change the peak count level, so a small defocus is useful to minimise the effects of non-linearity.

7.2   Looking ahead

The next few years promise to be incredibly exciting for the field of exoplanets, with the Transiting Exoplanet Survey Satellite (TESS, Ricker et al., 2015) launched on 18th April 2018, and James Webb Space Telescope (JWST, Gardner et al., 2006) scheduled for launch in 2020.

TESS is predicted to discover > 1800 sub-Neptunes and > 200 super-Earths orbiting bright stars (Barclay et al., 2018). This haul will offer fantastic opportunities for atmospheric characterisation using ground-based transmission spectroscopy. To date, only a handful of exoplanets of this size have been amenable to further study, but TESS will open up this parameter space by finding planets around bright stars. This will allow us to characterise smaller planets than currently possible without the need for giant telescopes. It is a goal of mine to capitalise on TESS's discoveries by obtaining ground-based transmission spectra of these targets. This will extend the LRG-BEASTS sample down to smaller planets, perhaps with an analogous survey of smaller planets called SML-BEASTS.

While TESS will open up a new era of exoplanet detection, JWST will revolutionise the field of exoplanet characterisation. This facility will be the premier observatory for near- and mid-infrared observations. Among other things, this will allow us to detect carbon species previously unobserved in transiting exoplanet spectra and place constraints on carbon to oxygen (C/O) ratios. This will enable us to study their formation histories (see Chapter 1). JWST will provide high-precision measurements of atmospheric metallicities, which can be compared to the mass-metallicity relation of solar system planets, and it will allow us to break degeneracies between cloudy atmospheres and high mean molecular weight atmospheres (Greene et al., 2016). However, it will not be able to observe at optical wavelengths, which can reveal key atmospheric opacity sources. These include Rayleigh scattering (e.g. Kirk et al., 2017), from which the cloud species can be inferred (Pinhas & Madhusudhan, 2017), in addition to titanium oxide (e.g. Sedaghati et al., 2017a) and the alkali metals (e.g. Nikolov et al., 2016). Therefore, optical measurements from the ground and with HST will continue to be essential for a full characterisation of an exoplanet's atmosphere.

Looking even further ahead, the 2020s will see the advent of the Giant Mag-

ellen Telescope (GMT, expected first light 2023), European Extremely Large Telescope (E-ELT, expected first light 2024) and Thirty Meter Telescope (TMT, expected first light 2026). This new generation of telescopes with primary mirrors of between 25 and 39 m in diameter will allow us to characterise the atmospheres of smaller planets. In addition to these facilities, ESA's PLATO mission is scheduled to launch in 2026 (Rauer et al., 2014). PLATO is predicted to find up to 2,200 planets with radii $< 2R_\oplus$ around host stars with $V < 11$ (ESA PLATO Definition Study Report, Red Book[2]), with the particular aim of finding habitable zone planets around Sun-like stars. PLATO's bright targets will further increase the opportunities for ground-based atmospheric characterisation.

Exoplanet atmospheres is a rapidly moving field with several significant and incredibly exciting results published during the course of my PhD. Upcoming space missions and extremely large ground-based telescopes will continue to increase the pace of discovery with ground-breaking results and I am very excited to be a part of this future.

---

[2]http://sci.esa.int/plato/59251-plato-definition-study-report-red-book/

Bibliography

Agol E., Fabrycky D. C., 2017, Transit-Timing and Duration Variations for the Discovery and Characterization of Exoplanets. Springer International Publishing, Cham, pp 1–20, doi:10.1007/978-3-319-30648-3˙7-1, `https://doi.org/10.1007/978-3-319-30648-3_7-1`

Agol E., Steffen J., Sari R., Clarkson W., 2005, MNRAS, 359, 567

Ahern S., Chapman G. A., 2000, Sol. Phys., 191, 71

Aigrain S., Parviainen H., Pope B. J. S., 2016, MNRAS, 459, 2408

Ambikasaran S., Foreman-Mackey D., Greengard L., Hogg D. W., O'Neil M., 2014, preprint, (`arXiv:1403.6015`)

Anglada-Escudé G., et al., 2016, Nature, 536, 437

Arcangeli J., et al., 2018, preprint, (`arXiv:1801.02489`)

Armitage P. J., Bonnell I. A., 2002, MNRAS, 330, L11

Armstrong D. J., de Mooij E., Barstow J., Osborn H. P., Blake J., Saniee N. F., 2016, Nature Astronomy, 1, 0004

Avila G., Rupprecht G., Beckers J. M., 1997, in Ardeberg A. L., ed., Proc. SPIEVol. 2871, Optical Telescopes of Today and Tomorrow. pp 1135–1143, doi:10.1117/12.269000

Baglin A., et al., 2006, in 36th COSPAR Scientific Assembly.

Bakos G., Noyes R. W., Kovács G., Stanek K. Z., Sasselov D. D., Domsa I., 2004, PASP, 116, 266

Ballard S., et al., 2011, ApJ, 743, 200

Baluev R. V., et al., 2015, MNRAS, 450, 3101

Barclay T., Pepper J., Quintana E. V., 2018, preprint, (`arXiv:1804.05050`)

Barnes J. R., Collier Cameron A., 2001, MNRAS, 326, 950

Barnes J. R., James D. J., Collier Cameron A., 2004, MNRAS, 352, 589

Baruteau C., et al., 2014, Protostars and Planets VI, pp 667–689

Batygin K., 2012, Nature, 491, 418

Batygin K., Stevenson D. J., 2010, ApJ, 714, L238

Baudoz P., 2017, Future Exoplanet Research: High-Contrast Imaging Techniques. Springer International Publishing, Cham, pp 1–16, doi:10.1007/978-3-319-30648-3˙160-1, `https://doi.org/10.1007/978-3-319-30648-3_160-1`

Béky B., Kipping D. M., Holman M. J., 2014, MNRAS, 442, 3686

Benn C., Dee K., Agócs T., 2008, in Ground-based and Airborne Instrumentation for Astronomy II. p. 70146X, doi:10.1117/12.788694

Bento J., et al., 2014, MNRAS, 437, 1511

Berdyugina S. V., 2005, LRSP, 2, 8

Beuzit J.-L., et al., 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. p. 18, doi:10.1117/12.790120

Birkby J. L., de Kok R. J., Brogi M., Schwarz H., Snellen I. A. G., 2017, AJ, 153, 138

Bodenheimer P., Lin D. N. C., Mardling R. A., 2001, ApJ, 548, 466

Bodenheimer P., D'Angelo G., Lissauer J. J., Fortney J. J., Saumon D., 2013, ApJ, 770, 120

Boffin H., et al., 2015, The Messenger, 159, 6

Boisse I., et al., 2012, A&A, 545, A55

Bond I. A., et al., 2001, MNRAS, 327, 868

Bond I. A., et al., 2004, ApJ, 606, L155

Borucki W. J., et al., 2010, Science, 327, 977

Borucki W. J., et al., 2011, ApJ, 736, 19

Boss A. P., 1997, Science, 276, 1836

Bouchy F., et al., 2005, A&A, 444, L15

Bowler B. P., 2016, Publications of the Astronomical Society of the Pacific, 128, 102001

Bowler B. P., Nielsen E. L., 2018, preprint, p. arXiv:1802.10132 (`arXiv:1802.10132`)

Brogi M., Snellen I. A. G., de Kok R. J., Albrecht S., Birkby J., de Mooij E. J. W., 2012, Nature, 486, 502

Brogi M., de Kok R. J., Birkby J. L., Schwarz H., Snellen I. A. G., 2014, A&A, 565, A124

Brogi M., de Kok R. J., Albrecht S., Snellen I. A. G., Birkby J. L., Schwarz H., 2016, ApJ, 817, 106

Brogi M., Giacobbe P., Guilluy G., de Kok R. J., Sozzetti A., Mancini L., Bonomo A. S., 2018, preprint, (`arXiv:1801.09569`)

Brown T. M., 2001, ApJ, 553, 1006

Burke C. J., et al., 2007, ApJ, 671, 2115

Burrows A., Hubeny I., Budaj J., Hubbard W. B., 2007, ApJ, 661, 502

Burton J. R., Watson C. A., Littlefair S. P., Dhillon V. S., Gibson N. P., Marsh T. R., Pollacco D., 2012, ApJS, 201, 36

Campbell B., Walker G. A. H., Yang S., 1988, ApJ, 331, 902

Cauley P. W., Redfield S., Jensen A. G., 2017, AJ, 153, 185

Chabrier G., Baraffe I., 2007, ApJ, 661, L81

Chaplin W. J., et al., 2013, ApJ, 766, 101

Charbonneau D., Noyes R. W., Korzennik S. G., Nisenson P., Jha S., Vogt S. S., Kibrick R. I., 1999, ApJL, 522, L145

Charbonneau D., Brown T. M., Latham D. W., Mayor M., 2000, ApJ, 529, L45

Charbonneau D., Brown T. M., Noyes R. W., Gilliland R. L., 2002, ApJ, 568, 377

Charbonneau D., et al., 2005, ApJ, 626, 523

Chauvin G., Lagrange A.-M., Dumas C., Zuckerman B., Mouillet D., Song I., Beuzit J.-L., Lowrance P., 2004, A&A, 425, L29

Chauvin G., et al., 2017, A&A, 605, L9

Chen G., Pallé E., Nortmann L., Murgas F., Parviainen H., Nowak G., 2017, A&A, 600, L11

Chen G., et al., 2018, preprint, p. arXiv:1805.11744 (`arXiv:1805.11744`)

Claret A., 2000, A&A, 363, 1081

Claret A., Bloemen S., 2011, A&A, 529, A75

Collier Cameron A., Horne K., Penny A., James D., 1999, Nature, 402, 751

Connelly J. N., Bizzarro M., Krot A. N., Nordlund Å., Wielandt D., Ivanova M. A., 2012, Science, 338, 651

Copperwheat C. M., et al., 2013, MNRAS, 434, 661

Crossfield I. J. M., 2015, PASP, 127, 941

Crossfield I. J. M., Kreidberg L., 2017, AJ, 154, 261

Cumming A., Butler R. P., Marcy G. W., Vogt S. S., Wright J. T., Fischer D. A., 2008, PASP, 120, 531

Dawson R. I., 2014, ApJL, 790, L31

Dawson R. I., Johnson J. A., 2018, preprint, (`arXiv:1801.06117`)

Delrez L., et al., 2018, MNRAS,

Deming D., Seager S., Richardson L. J., Harrington J., 2005, Nature, 434, 740

Deming D., et al., 2013, ApJ, 774, 95

Demory B.-O., Seager S., 2011, ApJS, 197, 12

Demory B.-O., et al., 2013, ApJ, 776, L25

Dhillon V. S., et al., 2007, MNRAS, 378, 825

Di Gloria E., Snellen I., Albrecht S., 2015, A&A, 580, A84

Dittmann J. A., Close L. M., Green E. M., Fenwick M., 2009, ApJ, 701, 756

Dodson-Robinson S. E., Veras D., Ford E. B., Beichman C. A., 2009, ApJ, 707, 79

Dragomir D., Benneke B., Pearson K. A., Crossfield I. J. M., Eastman J., Barman T., Biddle L. I., 2015, ApJ, 814, 102

Dravins D., Lindegren L., Mezey E., Young A. T., 1998, PASP, 110, 610

Dressing C. D., Charbonneau D., 2015, ApJ, 807, 45

Durisen R. H., Boss A. P., Mayer L., Nelson A. F., Quinn T., Rice W. K. M., 2007, Protostars and Planets V, pp 607–622

Ehrenreich D., et al., 2015, Nature, 522, 459

Einstein A., 1936, Science, 84, 506

Espinoza N., Jordán A., 2015, ArXiv e-prints, 1503.07020,

Espinoza N., Jordán A., 2016, MNRAS, 457, 3573

Esposito M., et al., 2014, A&A, 564, L13

Evans T. M., Aigrain S., Gibson N., Barstow J. K., Amundsen D. S., Tremblin P., Mourier P., 2015, MNRAS, 451, 680

Evans T. M., et al., 2017, Nature, 548, 58

Faedi F., et al., 2011, A&A, 531, A40

Faedi F., et al., 2013, A&A, 551, A73

Fischer D. A., Valenti J., 2005, ApJ, 622, 1102

Fischer P. D., et al., 2016, ApJ, 827, 19

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Fortney J. J., Marley M. S., Barnes J. W., 2007, ApJ, 659, 1661

Fortney J. J., Lodders K., Marley M. S., Freedman R. S., 2008, ApJ, 678, 1419

Fortney J. J., Mordasini C., Nettelmann N., Kempton E. M.-R., Greene T. P., Zahnle K., 2013, ApJ, 775, 80

Fraine J., et al., 2014, Nature, 513, 526

Fu G., Deming D., Knutson H., Madhusudhan N., Mandell A., Fraine J., 2017, ApJ, 847, L22

Fukui A., et al., 2014, ApJ, 790, 108

Fulton B. J., et al., 2017, AJ, 154, 109

Gaia Collaboration Brown A. G. A., Vallenari A., Prusti T., de Bruijne J. H. J., Babusiaux C., Bailer-Jones C. A. L., 2018, preprint, p. arXiv:1804.09365 (`arXiv:1804.09365`)

Gao P., Marley M. S., Ackerman A. S., 2018, preprint, (`arXiv:1802.06241`)

Gardner J. P., et al., 2006, Space Sci. Rev., 123, 485

Gaudi B. S., Winn J. N., 2007, ApJ, 655, 550

Gelman A., Roberts G. O., Gilks W. R., 1996, Bayesian statistics, 5, 42

Gibson N. P., Aigrain S., Roberts S., Evans T. M., Osborne M., Pont F., 2012a, MNRAS, 419, 2683

Gibson N. P., et al., 2012b, MNRAS, 422, 753

Gibson N. P., Aigrain S., Barstow J. K., Evans T. M., Fletcher L. N., Irwin P. G. J., 2013a, MNRAS, 428, 3680

Gibson N. P., Aigrain S., Barstow J. K., Evans T. M., Fletcher L. N., Irwin P. G. J., 2013b, MNRAS, 436, 2974

Gibson N. P., Nikolov N., Sing D. K., Barstow J. K., Evans T. M., Kataria T., Wilson P. A., 2017, MNRAS, 467, 4591

Gillon M., et al., 2016, Nature, 533, 221

Gillon M., et al., 2017, Nature, 542, 456

Gizon L., Solanki S. K., 2003, ApJ, 589, 1009

Goldreich P., Tremaine S., 1980, ApJ, 241, 425

Gould A., et al., 2014, Science, 345, 46

Greene T. P., Line M. R., Montero C., Fortney J. J., Lustig-Yaeger J., Luther K., 2016, ApJ, 817, 17

Grether D., Lineweaver C. H., 2006, ApJ, 640, 1051

Guillot T., Showman A. P., 2002, A&A, 385, 156

Hartman J. D., et al., 2009, ApJ, 706, 785

Hartman J. D., et al., 2011, ApJ, 726, 52

Hastings W. K., 1970, Biometrika, 57, 97

Haswell C. A., 2010, Transiting exoplanets. Cambridge University Press

Hatzes A. P., Cochran W. D., Endl M., McArthur B., Paulson D. B., Walker G. A. H., Campbell B., Yang S., 2003, ApJ, 599, 1383

Hayek W., Sing D., Pont F., Asplund M., 2012, A&A, 539, A102

Haynes K., Mandell A. M., Madhusudhan N., Deming D., Knutson H., 2015, ApJ, 806, 146

Haywood R. D., et al., 2014, MNRAS, 443, 2517

Haywood R. D., et al., 2016, MNRAS, 457, 3637

Hébrard G., et al., 2013, A&A, 549, A134

Henderson C. B., Gaudi B. S., Han C., Skowron J., Penny M. T., Nataf D., Gould A. P., 2014, ApJ, 794, 52

Heng K., 2016, ApJ, 826, L16

Henry G. W., Marcy G. W., Butler R. P., Vogt S. S., 2000, ApJ, 529, L41

Hoeijmakers H. J., Schwarz H., Snellen I. A. G., de Kok R. J., Bonnefoy M., Chauvin G., Lagrange A. M., Girard J. H., 2018, preprint, (arXiv:1802.09721)

Hollands M. A., Gänsicke B. T., Koester D., 2018, MNRAS,

Holman M. J., et al., 2010, Science, 330, 51

Horne K., 1986, PASP, 98, 609

Howard A. W., et al., 2012, ApJS, 201, 15

Howell S. B., 2006, Handbook of CCD astronomy. Vol. 5, Cambridge University Press

Howell S. B., et al., 2014, PASP, 126, 398

Hubbard W. B., Fortney J. J., Lunine J. I., Burrows A., Sudarsky D., Pinto P., 2001, ApJ, 560, 413

Huber K. F., Czesla S., Wolter U., Schmitt J. H. M. M., 2009, A&A, 508, 901

Huitson C. M., 2013, PhD thesis, University of Exeter

Huitson C. M., Sing D. K., Vidal-Madjar A., Ballester G. E., Lecavelier des Etangs A., Désert J.-M., Pont F., 2012, MNRAS, 422, 2477

Huitson C. M., et al., 2013, MNRAS, 434, 3252

Huitson C. M., Désert J.-M., Bean J. L., Fortney J. J., Stevenson K. B., Bergmann M., 2017, AJ, 154, 95

Husser T.-O., Wende-von Berg S., Dreizler S., Homeier D., Reiners A., Barman T., Hauschildt P. H., 2013, A&A, 553, A6

Ibgui L., Burrows A., Spiegel D. S., 2010, ApJ, 713, 751

Irwin M. J., Webster R. L., Hewett P. C., Corrigan R. T., Jedrzejewski R. I., 1989, AJ, 98, 1989

Irwin P., et al., 2008, Journal of Quantitative Spectroscopy and Radiative Transfer, 109, 1136

Jones E., Oliphant T., Peterson P., et al., 2001, SciPy: Open source scientific tools for Python, http://www.scipy.org/

Jordán A., et al., 2013, ApJ, 778, 184

Juvan I. G., Lendl M., Cubillos P. E., Fossati L., Tregloan-Reed J., Lammer H., Guenther E. W., Hanslmeier A., 2018, A&A, 610, A15

Kant I., 1755, Allgemeine Naturgeschichte und Theorie des Himmels

Kass R. E., Raftery A. E., 1995, Journal of the American Statistical Association, 90, 773

Kempton E. M.-R., Lupu R., Owusu-Asare A., Slough P., Cale B., 2017, PASP, 129, 044402

Khalafinejad S., et al., 2017, A&A, 598, A131

Kim S.-L., et al., 2016, Journal of Korean Astronomical Society, 49, 37

Kipping D. M., 2013, MNRAS, 435, 2152

Kirk J., Wheatley P. J., Louden T., Littlefair S. P., Copperwheat C. M., Armstrong D. J., Marsh T. R., Dhillon V. S., 2016, MNRAS, 463, 2922

Kirk J., Wheatley P. J., Louden T., Doyle A. P., Skillen I., McCormac J., Irwin P. G. J., Karjalainen R., 2017, MNRAS, 468, 3907

Kirk J., Wheatley P. J., Louden T., Skillen I., King G. W., McCormac J., Irwin P. G. J., 2018, MNRAS, 474, 876

Knutson H. A., et al., 2007, Nature, 447, 183

Kochanek C. S., et al., 2017, PASP, 129, 104502

Kozai Y., 1962, AJ, 67, 591

Kreidberg L., et al., 2014a, Nature, 505, 69

Kreidberg L., et al., 2014b, ApJ, 793, L27

Kreidberg L., et al., 2015, ApJ, 814, 66

Kurucz R., 1993, ATLAS9 Stellar Atmosphere Programs and 2 km/s grid. Kurucz CD-ROM No. 13. Cambridge, Mass.: Smithsonian Astrophysical Observatory, 1993., 13

Lam K. W. F., et al., 2017, A&A, 599, A3

Lammer H., Selsis F., Ribas I., Guinan E. F., Bauer S. J., Weiss W. W., 2003, ApJ, 598, L121

Langland-Shula L. E., Vogt S. S., Charbonneau D., Butler P., Marcy G., 2009, ApJ, 696, 1355

Latham D. W., Stefanik R. P., Mazeh T., Mayor M., Burki G., 1989, Nature, 339, 38

Laughlin G., 2018, Mass-Radius Relations of Giant Planets: The Radius Anomaly and Interior Models. Springer International Publishing, Cham, pp 1–17, doi:10.1007/978-3-319-30648-3˙1-1, `https://doi.org/10.1007/978-3-319-30648-3_1-1`

Laughlin G., Crismani M., Adams F. C., 2011, ApJ, 729, L7

Laureijs R., et al., 2011, preprint, p. arXiv:1110.3193 (`arXiv:1110.3193`)

Lavvas P., Koskinen T., 2017, ApJ, 847, 32

Lecavelier Des Etangs A., Pont F., Vidal-Madjar A., Sing D., 2008a, A&A, 481, L83

Lecavelier Des Etangs A., Vidal-Madjar A., Désert J.-M., Sing D., 2008b, A&A, 485, 865

Lecavelier Des Etangs A., et al., 2010, A&A, 514, A72

Leconte J., Chabrier G., Baraffe I., Levrard B., 2010, A&A, 516, A64

Leigh C., Collier Cameron A., Horne K., Penny A., James D., 2003, MNRAS, 344, 1271

Liang M.-C., Seager S., Parkinson C. D., Lee A. Y.-T., Yung Y. L., 2004, ApJ, 605, L61

Lin D. N. C., Bodenheimer P., Richardson D. C., 1996, Nature, 380, 606

Line M. R., Knutson H., Deming D., Wilkins A., Desert J.-M., 2013, ApJ, 778, 183

Lissauer J. J., 1993, ARA&A, 31, 129

Lodders K., 1999, ApJ, 519, 793

Lodders K., 2004, Science, 303, 323

Lomb N. R., 1976, Astrophysics and space science, 39, 447

Lopez E. D., Fortney J. J., 2013, ApJ, 776, 2

López-Morales M., Coughlin J. L., Sing D. K., Burrows A., Apai D., Rogers J. C., Spiegel D. S., Adams E. R., 2010, ApJ, 716, L36

Louden T., Wheatley P. J., 2015, ApJ, 814, L24

Louden T., Wheatley P. J., Irwin P. G. J., Kirk J., Skillen I., 2017, MNRAS, 470, 742

Lund M. N., et al., 2014, A&A, 570, A54

MacDonald R. J., Madhusudhan N., 2017, MNRAS, 469, 1979

195

Macintosh B. A., et al., 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. p. 18, doi:10.1117/12.788083

Macintosh B., et al., 2015, Science, 350, 64

Madhusudhan N., Crouzet N., McCullough P. R., Deming D., Hedges C., 2014, ApJ, 791, L9

Madhusudhan N., Agúndez M., Moses J. I., Hu Y., 2016, Space Sci. Rev., 205, 285

Magic Z., Asplund M., 2014, preprint, (`arXiv:1405.7628`)

Magic Z., Chiavassa A., Collet R., Asplund M., 2015, A&A, 573, A90

Mallonn M., et al., 2015a, A&A, 580, A60

Mallonn M., et al., 2015b, A&A, 583, A138

Mallonn M., et al., 2016, MNRAS, 463, 604

Mancini L., et al., 2013, MNRAS, 436, 2

Mancini L., et al., 2014, A&A, 562, A126

Mancini L., et al., 2017, MNRAS, 465, 843

Mandel K., Agol E., 2002, ApJ, 580, L171

Marcy G. W., Butler R. P., 2000, PASP, 112, 137

Marley M. S., Ackerman A. S., Cuzzi J. N., Kitzmann D., 2013, Clouds and Hazes in Exoplanet Atmospheres. University of Arizona Press, pp 367–391, doi:10.2458/azu˙uapress˙9780816530595-ch15

Marois C., Macintosh B., Barman T., Zuckerman B., Song I., Patience J., Lafrenière D., Doyon R., 2008, Science, 322, 1348

Marois C., Zuckerman B., Konopacky Q. M., Macintosh B., Barman T., 2010, Nature, 468, 1080

Martínez-Arnáiz R., Maldonado J., Montes D., Eiroa C., Montesinos B., 2010, A&A, 520, A79

Martioli E., et al., 2018, MNRAS, 474, 4264

Mayor M., Queloz D., 1995, Nature, 378, 355

Mayor M., et al., 2003, The Messenger, 114, 20

Mayor M., et al., 2011, preprint, (`arXiv:1109.2497`)

Mazeh T., Holczer T., Faigler S., 2016, A&A, 589, A75

McCullough P., Crouzet N., Deming D., Madhusudhan N., 2014, ApJ, 791, 55

McDonald I., et al., 2014, MNRAS, 445, 4137

McLaughlin D. B., 1924, ApJ, 60, 22

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, J. Chem. Phys., 21, 1087

Mills S. M., Fabrycky D. C., 2017, AJ, 153, 45

Mohler-Fischer M., et al., 2013, A&A, 558, A55

Montmerle T., Augereau J.-C., Chaussidon M., Gounelle M., Marty B., Morbidelli A., 2006, Earth, Moon, and Planets, 98, 39

Morley C. V., Fortney J. J., Kempton E. M.-R., Marley M. S., Visscher C., Zahnle K., 2013, ApJ, 775, 33

Morley C. V., Fortney J. J., Marley M. S., Zahnle K., Line M., Kempton E., Lewis N., Cahoy K., 2015, ApJ, 815, 110

Morton T. D., 2012, ApJ, 761, 6

Morton T. D., 2015, VESPA: False positive probabilities calculator, Astrophysics Source Code Library (ascl:1503.011)

Močnik T., Southworth J., Hellier C., 2017, MNRAS, 471, 394

Nelder J. A., Mead R., 1965, The computer journal, 7, 308

Nelson B. E., Ford E. B., Rasio F. A., 2017, AJ, 154, 106

Niemann H. B., et al., 1996, Science, 272, 846

Nikolov N., et al., 2014, MNRAS, 437, 46

Nikolov N., et al., 2015, MNRAS, 447, 463

Nikolov N., Sing D. K., Gibson N. P., Fortney J. J., Evans T. M., Barstow J. K., Kataria T., Wilson P. A., 2016, ApJ, 832, 191

Nikolov N., et al., 2018, MNRAS, 474, 1705

Norris C. M., Beeck B., Unruh Y. C., Solanki S. K., Krivova N. A., Yeo K. L., 2017, A&A, 605, A45

Nugroho S. K., Kawahara H., Masuda K., Hirano T., Kotani T., Tajitsu A., 2017, AJ, 154, 221

Öberg K. I., Murray-Clay R., Bergin E. A., 2011, ApJ, 743, L16

Oshagh M., Boisse I., Boué G., Montalto M., Santos N. C., Bonfils X., Haghighipour N., 2013, A&A, 549, A35

Oshagh M., Santos N., Ehrenreich D., Haghighipour N., Figueira P., Santerne A., Montalto M., 2014, A&A, 568, A99

Owen J. E., Jackson A. P., 2012, MNRAS, 425, 2931

Owen J. E., Wu Y., 2013, ApJ, 775, 105

Pál A., 2009, MNRAS, 396, 1737

Palle E., et al., 2017, A&A, 602, L15

Palmer C., Loewen E. G., 2005, Diffraction grating handbook. Newport Corporation New York

Parviainen H., 2017, preprint, (`arXiv:1711.03329`)

Parviainen H., Aigrain S., 2015, MNRAS, 453, 3821

Parviainen H., Pallé E., Nortmann L., Nowak G., Iro N., Murgas F., Aigrain S., 2016, A&A, 585, A114

Parviainen H., et al., 2017, preprint, (`arXiv:1709.01875`)

Penny M. T., et al., 2013, MNRAS, 434, 2

Pepper J., et al., 2007, PASP, 119, 923

Pepper J., et al., 2017, AJ, 153, 215

Perryman M., 2011, The Exoplanet Handbook. Cambridge University Press

Perryman M. A. C., et al., 2001, A&A, 369, 339

Perryman M., Hartman J., Bakos G. Á., Lindegren L., 2014, ApJ, 797, 14

Pinhas A., Madhusudhan N., 2017, preprint, (arXiv:1705.08893)

Pollacco D. L., et al., 2006, PASP, 118, 1407

Pollack J. B., Hubickyj O., Bodenheimer P., Lissauer J. J., Podolak M., Greenzweig Y., 1996, Icarus, 124, 62

Pont F., et al., 2007, A&A, 476, 1347

Pont F., Knutson H., Gilliland R. L., Moutou C., Charbonneau D., 2008, MNRAS, 385, 109

Pont F., Sing D. K., Gibson N. P., Aigrain S., Henry G., Husnoo N., 2013, MNRAS, 432, 2917

Press W. H., 2007, Numerical recipes 3rd edition: The art of scientific computing. Cambridge university press

Queloz D., Eggenberger A., Mayor M., Perrier C., Beuzit J. L., Naef D., Sivan J. P., Udry S., 2000, A&A, 359, L13

Rabus M., et al., 2009, A&A, 494, 391

Rackham B., et al., 2017, ApJ, 834, 151

Rafikov R. R., 2005, ApJ, 621, L69

Rajpaul V., Aigrain S., Osborne M. A., Reece S., Roberts S., 2015, MNRAS, 452, 2269

Rasio F. A., Ford E. B., 1996, Science, 274, 954

Rasmussen C. E., Williams C. K., 2006, Gaussian processes for machine learning. Vol. 1, MIT press Cambridge

Rauer H., et al., 2014, Experimental Astronomy, 38, 249

Rayleigh L., 1899, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 47, 375

Redfield S., Endl M., Cochran W. D., Koesterke L., 2008, ApJ, 673, L87

Ricker G. R., et al., 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Roberts S., Osborne M., Ebden M., Reece S., Gibson N., Aigrain S., 2013, Phil. Trans. R. Soc. A, 371, 20110550

Rossiter R. A., 1924, ApJ, 60, 15

Rostron J., 2015, PhD thesis, University of Warwick

Rostron J. W., Wheatley P. J., Anderson D. R., Collier Cameron A., Fortney J. J., Harrington J., Knutson H. A., Pollacco D. L., 2014, MNRAS, 441, 3666

Rowan D., et al., 2016, ApJ, 817, 104

Sahlmann J., Lazorenko P. F., Ségransan D., Martín E. L., Queloz D., Mayor M., Udry S., 2013, A&A, 556, A133

Santerne A., et al., 2016, A&A, 587, A64

Santos N. C., Israelian G., Mayor M., Rebolo R., Udry S., 2003, A&A, 398, 363

Scargle J. D., 1982, The Astrophysical Journal, 263, 835

Schwarz G., et al., 1978, The annals of statistics, 6, 461

Seager S., 2010, Exoplanet atmospheres: physical processes. Princeton University Press

Seager S., Mallén-Ornelas G., 2003, ApJ, 585, 1038

Seager S., Sasselov D. D., 2000, ApJ, 537, 916

Seager S., Kuchner M., Hier-Majumder C. A., Militzer B., 2007, ApJ, 669, 1279

Sedaghati E., et al., 2016, A&A, 596, A47

Sedaghati E., Boffin H. M. J., Delrez L., Gillon M., Csizmadia S., Smith A. M. S., Rauer H., 2017a, MNRAS, 468, 3123

Sedaghati E., et al., 2017b, Nature, 549, 238

Seeliger M., et al., 2015, MNRAS, 451, 4060

Shapiro A. I., Solanki S. K., Krivova N. A., Schmutz W. K., Ball W. T., Knaack R., Rozanov E. V., Unruh Y. C., 2014, A&A, 569, A38

Shappee B. J., et al., 2014, ApJ, 788, 48

Sheppard K. B., Mandell A. M., Tamburo P., Gandhi S., Pinhas A., Madhusudhan N., Deming D., 2017, ApJ, 850, L32

Showman A. P., Guillot T., 2002, A&A, 385, 166

Silva-Valio A., Lanza A. F., Alonso R., Barge P., 2010, A&A, 510, A25

Sing D. K., Vidal-Madjar A., Désert J.-M., Lecavelier des Etangs A., Ballester G., 2008a, ApJ, 686, 658

Sing D. K., Vidal-Madjar A., Lecavelier des Etangs A., Désert J.-M., Ballester G., Ehrenreich D., 2008b, ApJ, 686, 667

Sing D. K., et al., 2011a, MNRAS, 416, 1443

Sing D. K., et al., 2011b, A&A, 527, A73

Sing D. K., et al., 2012, MNRAS, 426, 1663

Sing D. K., et al., 2013, MNRAS, 436, 2956

Sing D. K., et al., 2015, MNRAS, 446, 2428

Sing D. K., et al., 2016, Nature, 529, 59

Smalley B., Smith K. C., Dworetsky M. M., 2001, UCLSYN User Guide (v 4.1), http://www.astro.keele.ac.uk/~bs/publs/uclsyn.pdf

Smith K. C., 1992, PhD thesis, University of London, University College London (United Kingdom

Smith K. C., Dworetsky M. M., 1988, in Adelman S. J., Lanz T., eds, Elemental Abundance Analyses. p. 32

Snellen I. A. G., Albrecht S., de Mooij E. J. W., Le Poole R. S., 2008, A&A, 487, 357

Snellen I. A. G., de Kok R. J., de Mooij E. J. W., Albrecht S., 2010, Nature, 465, 1049

Southworth J., 2008, MNRAS, 386, 1644

Southworth J., 2011, MNRAS, 417, 2166

Southworth J., Evans D. F., 2016, MNRAS, 463, 37

Southworth J., Wheatley P. J., Sams G., 2007, MNRAS, 379, L11

Southworth J., et al., 2009, MNRAS, 396, 1023

Southworth J., Mancini L., Maxted P. F. L., Bruni I., Tregloan-Reed J., Barbieri M., Ruocco N., Wheatley P. J., 2012, MNRAS, 422, 3099

Southworth J., et al., 2015, MNRAS, 447, 711

Southworth J., Mancini L., Madhusudhan N., Mollière P., Ciceri S., Henning T., 2017, AJ, 153, 191

Spergel D., et al., 2015, preprint, p. arXiv:1503.03757 (`arXiv:1503.03757`)

Stevenson K. B., 2016, ApJ, 817, L16

Stevenson K. B., et al., 2014, Science, 346, 838

Strassmeier K. G., 2009, A&A Rev., 17, 251

Sudarsky D., Burrows A., Pinto P., 2000, ApJ, 538, 885

Terquem C., Ajmia A., 2010, MNRAS, 404, 409

Thorngren D. P., Fortney J. J., 2017, preprint, (`arXiv:1709.04539`)

Tregloan-Reed J., Southworth J., Tappert C., 2013, MNRAS, 428, 3671

Triaud A. H. M. J., 2017, preprint, (`arXiv:1709.06376`)

Triaud A. H. M. J., et al., 2013, A&A, 551, A80

Triaud A. H. M. J., et al., 2015, MNRAS, 450, 2279

Tsiaras A., et al., 2017, preprint, (`arXiv:1704.05413`)

Turner J. D., et al., 2017, MNRAS, 472, 3871

Udalski A., 2003, ActaAstron, 53, 291

Unruh Y. C., Solanki S. K., Fligge M., 1999, A&A, 345, 635

Vanderburg A., et al., 2015, Nature, 526, 546

Vidal-Madjar A., Lecavelier des Etangs A., Désert J.-M., Ballester G. E., Ferlet R., Hébrard G., Mayor M., 2003, Nature, 422, 143

Vidal-Madjar A., et al., 2011, A&A, 527, A110

Vorobyov E. I., Elbakyan V., 2018, preprint, p. arXiv:1806.07675 (`arXiv:1806.07675`)

Wakeford H. R., Sing D. K., 2015, A&A, 573, A122

Wakeford H. R., et al., 2017a, Science, 356, 628

Wakeford H. R., Visscher C., Lewis N. K., Kataria T., Marley M. S., Fortney J. J., Mandell A. M., 2017b, MNRAS, 464, 4247

Wakeford H. R., et al., 2018, AJ, 155, 29

Walker G. A. H., Bohlender D. A., Walker A. R., Irwin A. W., Yang S. L. S., Larson A., 1992, ApJ, 396, L91

Weiss L. M., Marcy G. W., 2014, ApJ, 783, L6

Wheatley P. J., et al., 2017a, preprint, (`arXiv:1710.11100`)

Wheatley P. J., Louden T., Bourrier V., Ehrenreich D., Gillon M., 2017b, MNRAS, 465, L74

Wilson P. A., et al., 2015, MNRAS, 450, 192

Wilson D. J., Gänsicke B. T., Farihi J., Koester D., 2016, MNRAS, 459, 3282

Winn J. N., 2010, preprint, (`arXiv:1001.2010`)

Winn J. N., 2018, preprint, (`arXiv:1801.08543`)

Winn J. N., et al., 2008, ApJ, 683, 1076

Winn J. N., Fabrycky D., Albrecht S., Johnson J. A., 2010, ApJ, 718, L145

Wolszczan A., 1994, Science, 264, 538

Wolszczan A., Frail D. A., 1992, Nature, 355, 145

Wolter U., Schmitt J. H. M. M., Huber K. F., Czesla S., Müller H. M., Guenther E. W., Hatzes A. P., 2009, A&A, 504, 561

Wood P. L., Maxted P. F. L., Smalley B., Iro N., 2011, MNRAS, 412, 2376

Wu Y., Murray N., 2003, ApJ, 589, 605

Wyttenbach A., Ehrenreich D., Lovis C., Udry S., Pepe F., 2015, A&A, 577, A62

Wyttenbach A., et al., 2017, A&A, 602, A36

Yan F., Pallé E., Fosbury R. A. E., Petr-Gotzens M. G., Henning T., 2017, A&A, 603, A73

Zacharias N., Finch C. T., Girard T. M., Henden A., Bartlett J. L., Monet D. G., Zacharias M. I., 2012, VizieR Online Data Catalog, 1322

Zahnle K., Marley M. S., Fortney J. J., 2009a, preprint, (arXiv:0911.0728)

Zahnle K., Marley M. S., Freedman R. S., Lodders K., Fortney J. J., 2009b, ApJ, 701, L20

Zeng L., Sasselov D., 2013, PASP, 125, 227

Zhou G., Bayliss D. D. R., 2012, MNRAS, 426, 2483

Zuckerman B., Young E. D., 2017, Characterizing the Chemistry of Planetary Materials Around White Dwarf Stars. Springer International Publishing, Cham, pp 1–22, doi:10.1007/978-3-319-30648-3˙14-1, https://doi.org/10.1007/978-3-319-30648-3_14-1

de Wit J., Seager S., 2013, Science, 342, 1473

van Dokkum P. G., 2001, PASP, 113, 1420