Characterising Exoplanet Atmospheres: Bayesian Techniques for Transit Lightcurves

Thomas M. Evans

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2014

Characterising Exoplanet Atmospheres:
Bayesian Techniques for Transit Lightcurves

Thomas M. Evans
Balliol College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Trinity 2014

Precise measurements of transit lightcurves can be used to constrain the composition and structure of exoplanet atmospheres. Unfortunately, efforts to extract this information are usually hampered by the presence of correlated noise that is degenerate with the astrophysical signal of interest. A major theme of this thesis is the application of robust analysis methods to properly account for such degeneracies. In particular, I advocate the use of Bayesian inference for lightcurve fitting. For this study, the Bayesian framework is exploited by modelling lightcurves as Gaussian processes (GPs), which offer numerous advantages over traditional decorrelation methods. The main advantage is that GPs do not require a functional form to be specified for the poorly understood lightcurve systematics. Instead, the high-level properties of the signal covariance are parameterised, allowing complex correlations to be marginalised over relatively low-dimension parameter spaces.

I use GP models to analyse transit and eclipse lightcurves for the hot Jupiters HD 189733b, HD 209458b, and HAT-P-32b. The work is spread over three separate projects.

Firstly, I re-analyse the majority of the transits and eclipses that have been observed using the *Spitzer Space Telescope* Infrared Array Camera (IRAC) for HD 189733b and HD 20945b. The GP analyses generally produce uncertainties for inferred planet parameters that are factors of $\sim$ 1–5 larger than those quoted in the literature. In a number of cases, I obtain results that are fundamentally different to those published previously, with significant implications for the understanding of the atmospheres.

Secondly, I report an eclipse observation for HD 189733b that was made using the *Hubble Space Telescope* (*HST*) Space Telescope Imaging Spectrograph (STIS) over the 290–570 nm wavelength range. Geometric albedos of $A_g = 0.37^{+12}_{-13}$ and $A_g = 0.13^{+13}_{-12}$ are measured in the wavelength ranges 290–450 nm and 450–570 nm, respectively. This represents the first ever multi-wavelength eclipse measurement made for an exoplanet at visible wavelengths. The nonzero albedo in the 290–450 nm wavelength channel provides evidence for scattering in the atmosphere, possibly by haze/clouds or H_2 molecules. The relatively low albedo in the 450–570 nm wavelength channel is interpreted as being due to absorption by the wings of the Na 589 nm doublet.

Thirdly, I present two transit observations for HAT-P-32b made using the *Nordic Optical Telescope* (*NOT*) Andalucia Faint Object Spectrograph (ALFOSC) in multi-object spectroscopy mode over the 414–702 nm wavelength range. A joint GP model is fit to the two white lightcurves produced by integrating the spectra over the full wavelength range. Spectroscopic lightcurves are also generated by binning into 32 nm-wide wavelength channels, and preliminary lightcurve analyses are performed. The extracted transmission spectrum shows some evidence for absorption features, but this interpretation is currently very tentative. Further refinements to the data reduction and lightcurve analysis are suggested, which will allow the transmission spectrum to be evaluated more definitively.

Declaration

Research presented in this thesis was carried out under the supervision of Dr Suzanne Aigrain and Dr Neale Gibson between October 2010 and July 2014 at the University of Oxford. Except where indicated below or in the text of the thesis, the work presented is my own.

- Chapter 3 presents an analysis of observations made with the *Spitzer Space Telescope*, obtained from the NASA/IPAC Infrared Science Archive, both of which are operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration. I carried out all data reduction and lightcurve analyses. Joanna Barstow produced model atmosphere transmission and emission spectra.
- Chapter 4 presents the paper Evans et al., 2013, *The Deep Blue Color of HD 189733b: Albedo Measurements with Hubble Space Telescope/Space Telescope Imaging Spectrograph at Visible Wavelengths*, ApJL, 772, L16. The observations were made as part of *Hubble Space Telescope* Program GO-13006, with Frédéric Pont as Principal Investigator. The observations were designed and scheduled by Frédéric Pont. Data reduction and production of the calibrated time series spectra were done by David Sing. I performed the lightcurve analyses and wrote the paper.
- Chapter 5 presents observations made with the *Nordic Optical Telescope*, operated by the Nordic Optical Telescope Scientific Association at the Observatorio de los Muchachos, La Palma, Spain, of the Instituto de Astrofisica de Canarias. I was Principal Investigator for the observing program (Proposal 47-108), and led all aspects of the observations, including the design of the multi-object spectroscopy mask. I performed all of the data reduction and lightcurve analyses.

Thomas Evans
29 July 2014

Acknowledgements

I would like to start by thanking Suzanne Aigrain for being a wonderful, supportive supervisor. I've learned a lot from her over these past few years, and am extremely grateful for the warm and friendly environment she created within our research group, which helped me feel right at home in Oxford. Neale Gibson has played an enormous role during my PhD, and I thank him for the many stimulating discussions we've had, and for involving me in his various projects. Frédéric Pont has been a major influence in my work, and I thank him in particular for allowing me to analyse the STIS eclipse data for HD 189733b that is reported in Chapter 4. I am grateful to Jo Barstow, for producing the model spectra in Chapter 3 and assisting with the interpretation of observational data generally. Thanks also to: Stephen Roberts and Michael Osborne for opening my eyes to the delights of Bayesian inference; David Sing for reducing the STIS data that is presented in Chapter 4; Kevin Heng for discussions on exoplanet albedos; Hannu Parviainen for accompanying me on the observing run at the *NOT*; Ed and Ruth for being such good friends, both within our research group and outside; Thibaut, Jean, and Richard for being great housemates and helping to make my final year the best one; and all of the astrophysics grad students for the companionship and fun times. Finally, thank you to Mum, Dad, and Sylvie. Your love and support mean everything to me.

Contents

List of Figures

List of Tables

Thesis outline

The subject of this thesis is the observational characterisation of exoplanet atmospheres. In particular, it focuses on transiting hot Jupiter systems, and robust analysis methods to account for systematics that affect measured lightcurves.

The first two chapters provide a general background to the field and technical details for the analysis methods used in later chapters:

- Chapter 1 gives an overview of the observational study of exoplanets, with emphasis on the characterisation of hot Jupiter atmospheres. A brief historical outline of exoplanet detection is given, and important results from transit surveys are summarised, including some significant trends that have been uncovered for the hot Jupiter population. The use of transit lightcurves to characterise exoplanet atmospheres is explained and the specific aims of the thesis are stated.
- Chapter 2 elaborates on the Bayesian approach to model inference, and in particular how it applies to transit lightcurve fitting. Methods for marginalising the posterior distribution are described: Markov chain Monte Carlo, nested sampling, and variational methods. The theory of Gaussian processes (GPs) is introduced, laying an important foundation for later chapters, as the use of GP models to handle poorly understood systematics in transit lightcurves is a central tenet of the analyses presented in this work.

The following three chapters present the scientific contributions of this study:

- Chapter 3 presents an extensive analysis of archival lightcurves acquired using the *Spitzer Space Telescope* Infrared Array Camera (IRAC) for the canonical hot Jupiters HD 189733b and HD 209458b. The *Spitzer*/IRAC instrument is briefly described, and the ensemble of published transmission and emission data for both planets is reviewed to place the IRAC datasets in context. Details of the latter are given, and the production of lightcurves from the raw data frames is described. The lightcurves are modelled as GPs, allowing degeneracies between the astrophysical signal and instrumental systematics to be marginalised over. Transmission and emission results

obtained from these analyses are presented, a number of which provide a significant revision to previously published values. Implications for the atmospheres of each planet are discussed.

- Chapter 4 presents a secondary eclipse measurement for HD 189733b made using the *Hubble Space Telescope* Space Telescope Imaging Spectrograph (STIS) over the 290–570 nm wavelength range. At these wavelengths the planetary emission is comprised entirely of stellar radiation reflected by the atmosphere, with a negligible thermal contribution. Previous measurements of reflected light from exoplanet atmospheres are reviewed and the rationale for attempting the reflection measurement for HD 189733b using *HST*/STIS is explained. Details of the observation and data reduction are given. The lightcurves, which are binned into multiple wavelength channels, exhibit strong instrumental systematics and are once again modelled as GPs. Results of the lightcurve analyses are presented, including multi-wavelength constraints on the albedo. Implications for the planetary atmosphere are discussed.

- Chapter 5 presents two primary transit observations for the hot Jupiter HAT-P-32b that were made using the *Nordic Optical Telescope* (*NOT*) Andalucia Faint Object Spectrograph (ALFOSC) in multi-object spectroscopy (MOS) mode over the 414–702 nm wavelength range. The use of MOS to characterise exoplanet atmospheres is reviewed. The *NOT*/ALFOSC observations and data reduction are described. A joint GP analysis of the two "white" lightcurves (i.e. integrated over the 414–702 nm wavelength range) is presented. Analyses for spectroscopic lightcurves formed by binning the time series spectra into 32 nm-wide wavelength channels are also presented and the resulting transmission spectrum is discussed. However, only very tentative statements regarding the implications for the atmosphere can be made, due to the preliminary nature of the analysis.

The final conclusions are drawn together in Chapter 6.

Chapter 1

Characterising exoplanet atmospheres

A major theme of observational exoplanet science is the detailed characterisation of individual atmospheres. To date, this has mostly been done for transiting gas giant planets on short period orbits. Observations of three such systems are analysed for this thesis: HD 189733b, HD 209458b, and HAT-P-32b.

In this chapter, a general introduction to the field is given, including the methodology for extracting information about exoplanet atmospheres from transit lightcurves. Section 1.1 gives a historical overview of exoplanet detection, and Section 1.2 describes modern transit surveys. Significant trends that have been uncovered for the short-period gas giant population are summarised in Section 1.3. The important features of transiting exoplanet lightcurves are presented in Section 1.4, with an emphasis on how they can be used to constrain atmosphere properties. The specific aims of this thesis are given in Section 1.5.

1.1 Historical overview of exoplanet detection

Early steps towards the discovery of exoplanets were taken during the 19th century with the first indirect detections of unseen stellar companions (e.g. Mädler, 1842; Bessel, 1844). A notable claim was made by Jacob (1855), who deduced the existence of a planetary mass companion to 70 Ophiuchi with an

orbital period of 26 years. Subsequent astrometry measurements lent support to the planetary companion hypothesis (e.g. See, 1896; Reuyl & Holmberg, 1943), while others did not (e.g. Strand, 1937, 1952). Modern measurements and the accumulation of decades' worth of data have now shown that anomalies in the earlier datasets must have been due to systematic errors, rather than bona fide planetary signals (e.g. Heintz, 1988).

The 70 Ophiuchi case was not isolated. Throughout the 20th century, numerous claims of planetary-mass companions detected by astrometry were made (e.g. Strand, 1943; van de Kamp, 1944, 1969; Deich & Orlova, 1977a,b; Harrington et al., 1983). However, as with 70 Ophiuchi, none withstood further scrutiny (e.g. Gatewood & Eichhorn, 1973; Gatewood, 1974; Skrutskie et al., 1987; Choi et al., 2013).

While astrometry dominated the early exoplanet searches, Struve (1952) recommended two alternative methods for detecting a hypothetical population of close-in gas giant planets with orbital periods of order ∼1 day. His first suggestion was to monitor stellar radial velocities (RVs) by measuring spectral line shifts as a function of time. For a Jupiter-mass planet on a 1 day orbit, the expected RV signal would have an amplitude of $\sim 200\,\mathrm{m\,s^{-1}}$. The second approach was to use photoelectric instruments to detect the ∼ 1% drop in brightness that would occur when a Jupiter-sized planet transited across the face of its host star.

By the 1980s, spectrographs that could achieve long-term RV precisions of $\sim$10–100 $\mathrm{m\,s^{-1}}$ had been developed. Using one such instrument, Campbell et al. (1988) conducted a survey of 16 stars over six years. Although no definitive planet detections were made, a variation that repeated every 2.7 years was tentatively identified for the star γ Cephei A. Assuming the signal was caused by an unseen body orbiting the star, the amplitude implied a minimum mass[1] of just

[1] RV measurements allow us to determine $M_p \sin i$, where M_p is the planet mass and i is the orbital inclination. For more details, see e.g. Lovis & Fischer (2011).

$1.7 M_{\rm J}$ for the companion. Doubts were raised when the group noticed variations in the stellar activity cycle with a period similar to that of the RV variations (Walker et al., 1992). Since then, however, the original signal identified by Campbell et al. has been verified and shown to be independent of stellar activity, confirming that it is due to a companion (Hatzes et al., 2003).

Another early RV discovery of note was made by Latham et al. (1989), who detected a companion orbiting the F9V star HD 114762 with a period of 87 days and a minimum mass of 11 $M_{\rm J}$. Although this potentially puts the companion at the high end of the planetary mass regime, the unknown orbital inclination allows for a range of possible masses. Most likely, HD 114762b is a brown dwarf, drawn from the tail end of the stellar mass function.

The first unambiguous exoplanet detections were made by Wolszczan & Frail (1992), who serendipitously detected irregularities in the radio emission signals from the millesecond pulsar PSR 1257+12. The aperiodicity of the received pulses indicated the presence of two planets, both with masses $\sim 4\,M_{\oplus}$, and orbital periods of 98 and 67 days. Wolszczan (1994) confirmed these discoveries and also identified a third planet in the system with a period of 25 days and a mass of only $0.02\,M_{\oplus}$. Since then, only four other planets have been detected around pulsars, with a review of the field provided by Wolszczan & Kuchner (2011).

By the mid-1990s there still had not been a secure exoplanet detection around a main sequence star, despite the best efforts of dedicated surveys (e.g. Marcy & Butler, 1992; Cochran & Hatzes, 1994). The picture changed, however, when Mayor & Queloz (1995) measured a clear RV signal for the G5V star 51 Pegasus with a period of 4.2 days and an amplitude of $60\,{\rm m\,s^{-1}}$, corresponding to a minimum companion mass of $0.5\,M_{\rm J}$. The discovery was somewhat fortuitous, as Mayor & Queloz had been intending to refine the statistics of close-in, low-mass

stellar companions. To do this, they had selected a sample of 142 stars, and had adopted a strategy of observing each one intensely for short periods spanning a few days at a time. In contrast, surveys that were focused on detecting exoplanets tended to monitor smaller samples of stars over multi-year timescales in the hope of detecting gas giants with orbital periods of > 500 days, where planet formation theories predicted them to form (e.g. Lissauer, 1993; Boss, 1995). However, the discovery of 51 Peg b confirmed that, even if gas giants formed at or beyond the ice line,[2] they could then migrate inwards to orbits much closer to the star (e.g. Lin et al., 1996). With this realisation, gas giants on short-period orbits became prime targets for the RV searches, and additional discoveries soon followed (Marcy & Butler, 1996; Butler & Marcy, 1996; Butler et al., 1997, 1998; Fischer et al., 1999; Queloz et al., 2000).

The revelation that gas giants can exist on short-period orbits also spurred on efforts to measure exoplanet transits (e.g. Henry et al., 1997; Brown & Charbonneau, 1999). Success came when RV surveys uncovered a companion with a minimum mass of $0.6\,M_{\rm J}$ in a 3.5 day orbit around the G0V star HD 209458, and two groups subsequently observed it to transit (Charbonneau et al., 2000; Henry et al., 2000). This proved that HD 209458b was on an edge-on orbit, verifying that the true mass of the companion was very close to the minimum mass derived from RV, and thus unambiguously confirming its planetary nature. Furthermore, the drop in flux measured during transit allowed the relative size of the stellar and planetary discs to be determined. Using a model estimate for the stellar radius, the absolute planetary radius was then calculated to be $1.3\,R_{\rm J}$, implying a bulk density significantly lower than that of Saturn.

[2]Also known as the "snow line" or "frost line", the "ice line" refers to the region ~4–6 AU from the star where volatile species such as water and ammonia exist as solid ices in the protoplanetary disc (e.g. Boss, 1995). The high density of planet-building material is thought to allow rapid planetesimal core growth and runaway gas accretion to occur on timescales of a few million years, before the gas is lost from the disc.

In the following years, dozens of new transiting planet candidates were identified (e.g. Udalski et al., 2002a,b). Follow-up RV observations revealed most of these to be stellar binaries in configurations that produced planet-like transit signals (e.g. Brown, 2003). The first candidate confirmed to be a bona fide planet was OGLE-TR-56b, with an orbital period of 1.2 days, a mass of $0.9\,M_{\rm J}$, and a radius of $1.3\,R_{\rm J}$ (Konacki et al., 2003). Since then, many more planets have been uncovered by transit surveys (Section 1.2).

Indeed, to date, the vast majority of exoplanets have been discovered using either RV or transits. Other methods that have produced detections over the last decade, and which are likely to prove increasingly important going into the future, include: microlensing (e.g. Bond et al., 2004; Gaudi, 2012), direct imaging (e.g. Marois et al., 2008; Oppenheimer & Hinkley, 2009), and astrometry (e.g. Muterspaugh et al., 2010; Sozzetti et al., 2014).

1.2 Modern transit surveys

In the years after the first observed transit for HD 209458b, numerous surveys were established to search for more transiting exoplanets. A review of these early efforts was made by Horne (2003), who identified at least 23 such surveys that could be roughly divided into two categories: shallow and deep. The shallow surveys employed wide-angle CCDs mounted on small-aperture telescopes (5–50 cm) to observe relatively bright ($V < 13$ mag) stars over broad sky regions (tens of deg^2). The deep surveys used larger telescopes ($\sim$ 1–3 m) to target fainter stars ($V \sim 14$–18 mag) in narrower sky regions (< 1 deg^2).

The most successful deep ground-based survey has been the Optical Gravitational Lensing Experiment (OGLE), conducted using a dedicated 1.3 m telescope. Despite the primary intention being to detect gravitational microlensing signals of dark matter, the OGLE time series photometry has proven to be suitable for

the detection of transiting exoplanets, including the first planet discovered by the transit method, OGLE-TR-56b (Section 1.1). Since then, OGLE has uncovered an additional seven transiting exoplanets[3] (e.g. Pont et al., 2008b). In general, however, a disadvantage of the ground-based deep surveys has been the difficulty in following up planet candidates with RV observations, due to the faintness of the host stars.

Notable shallow ground-based surveys include TrES (Alonso et al., 2004), HATNet (Bakos et al., 2004), XO (McCullough et al., 2005), WASP (Pollacco et al., 2006), and MEarth (Nutzman & Charbonneau, 2008). Of these, the two that have discovered the most planets by far are HATNet and WASP, with a combined haul of well over 100 planets. Almost all of these are gas giants with orbital periods < 10 day and host star magnitudes of $V \sim 10$–12 mag. These properties lend the HATNet and WASP planets to detailed atmospheric characterisation from the ground: the short orbital periods result in frequent transits; the predominantly gaseous compositions and high equilibrium temperatures correspond to extended atmospheres that are relatively easy to detect; and the moderate host star brightnesses allow enough photons to be collected by 2–8 m telescopes without saturating the detectors. Importantly, it is also quite common for target stars in the $V = 10$–12 mag range to have one or more nearby (< 5 arcmin) comparison stars of similar brightness, which are necessary for performing differential spectrophotometry from the ground. One such system, HAT-P-32, forms the subject of Chapter 5.

Two space missions have also been launched with the specific goal of searching for transiting exoplanets: *CoRoT* (Auvergne et al., 2009) and *Kepler* (Borucki et al., 2010). The *CoRoT* satellite, equipped with a 27 cm telescope, was launched into an Earth-polar orbit in 2006 and was operational until 2012. The larger

[3]OGLE has also discovered exoplanets by microlensing (e.g. Gaudi et al., 2008).

Kepler satellite has a 95 cm telescope and was launched into an Earth-trailing orbit in 2009. It was fully operational until 2013, when the second of four pointing wheels failed; the mission continues operate with a revised set of science goals and observing strategy (Howell et al., 2014).

Over the course of its lifetime, the CoRoT satellite observed approximately 1.6×10^5 stars, typically spread across 3–4 fields per year (Moutou et al., 2013). To date, 27 transiting substellar companions have been published, of which 19 are gas giants with masses in the range 0.2–5 M_J and periods < 15 day. Of the remainder, four have masses $< 0.2\,M_J$ and three have masses $> 5\,M_J$, while one planet (CoRoT-9b; Deeg et al., 2010) has a mass of $0.8\,M_J$ and a longer orbital period of ~ 95 day.

The *Kepler* satellite adopted a strategy of staring at a fixed ~ 100 deg^2 field during the entire primary mission, acquiring continuous photometry for ~ 1.5×10^5 main sequence stars. From an analysis of the first 22 months of data, the tally of planet candidates discovered by *Kepler* stood at over 4000, with at least 880 of these validated as genuine planets at the time of writing (e.g. Burke et al., 2014; Marcy et al., 2014; Rowe et al., 2014). The frequency of planets is found to increase dramatically with decreasing planet size, from ~ $1\,R_J$ down to ~ $2\,R_\oplus$, and with increasing orbital period, out to 50 days (Howard et al., 2012). Thus, although short period gas giants – such as the three considered in this thesis (i.e. HD 189733b, HD 209458b, and HAT-P-32b) – provide the most favourable targets for atmospheric characterisation using transit lightcurves (Section 1.4.1), they are atypical of the overall planet population. Furthermore, the trend of increasing planet frequency with decreasing planet size and increasing orbital period explains the large disparity between the yields from the *Kepler* and *CoRoT* programs: firstly, with a smaller telescope, the sensitivity of *CoRoT* to smaller planets is considerably lower than that of *Kepler*; and secondly, by continuously

switching between fields, the sensitivity of *CoRoT* to longer period planets is reduced significantly compared to *Kepler*.

1.3 The hot Jupiters

Since the discovery of 51 Peg b (Section 1.1), an entire population of gas giant planets with orbital periods $P \lesssim 10\,\mathrm{day}$ and semimajor axes $a \lesssim 0.1\,\mathrm{AU}$ has been uncovered by RV and transit surveys (e.g. Butler et al., 2006; Udry & Santos, 2007; Howard et al., 2012; Wright et al., 2012). Orbiting at such short distances from their parent stars, the upper atmospheres of these planets are heated to temperatures of 1000–3000 K. As such, those with masses in the range ~ 0.5–$5\,M_\mathrm{J}$ are commonly referred to as "hot Jupiters", three of which are studied in subsequent chapters: HD 209458b (Chapter 3), HD 189733b (Chapters 3 and 4), and HAT-P-32b (Chapter 5).

Hot Jupiters are not believed to have formed in situ, due to the limited amount of material available for building planets within the ice line (e.g. Papaloizou & Terquem, 2006). Rather, they are likely to have formed further out from the star, in the vicinity of the ice line, before migrating inwards. The primary mechanism for this migration is still a topic of debate, with possibilities including gravitational interactions with the protoplanetary disc (e.g. Lin et al., 1996), planet-planet scattering (e.g. Chatterjee et al., 2008), and Kozai interactions with a third body in the system (e.g. Fabrycky & Tremaine, 2007).

The detection of the hot Jupiters has been helped by the fact that they are the easiest to detect using the RV and transit methods. This is due to a variety of reasons, including: the transit depth increases with planet radius; the RV amplitude increases with planet mass; the geometric transit probability increases with decreasing semimajor axis; and small orbits correspond to short orbital periods, allowing repeat signals to be accumulated rapidly.

Analysing the planet candidates from the first 16 months of *Kepler* data, Fressin et al. (2013) deduced that $0.43 \pm 0.05\%$ of FGK stars have a gas giant planet with an orbital period $P < 10$ day. This is consistent with the frequency of $0.31^{+0.43}_{-0.18}\%$ for $P < 5$ day hot Jupiters found by Gould et al. (2006) based on the findings of the OGLE transit survey. However, somewhat higher estimates have been obtained by analyses of RV survey results. For instance, Wright et al. (2012) obtained a frequency of $1.2 \pm 0.4\%$ by analysing data from the California Planet Survey. The reason for these differences remains unclear, but it is likely related to the distinct properties of the stellar populations probed by *Kepler* and the RV surveys.

The observational characterisation of hot Jupiter systems has revealed a number of significant trends. These include:

- Most hot Jupiter orbits are consistent with being circular (e.g. Pont et al., 2011). This is in line with expectations that strong tidal forces from the nearby star will act to circularise the orbit of a typical hot Jupiter on timescales of ~ 1 Gyr or less (e.g. Rasio et al., 1996). Tidal forces are also expected to synchronise the rotation period of hot Jupiters with their orbital periods over $\sim$ 1–10 Myr timescales, so that the same hemisphere always faces the star (Correia & Laskar, 2011).

- There is a peak in the distribution of hot Jupiters with semimajor axes of ~ 0.04 AU, corresponding to orbital periods of ~ 3 day (e.g. Butler et al., 2006; Udry & Santos, 2007; Wright et al., 2009). The frequency of gas giants drops off again between $\sim$ 0.08–0.8 AU, before starting to rise at larger semimajor axes. This suggests that the inward migration of hot Jupiters is halted quite sharply at ~ 0.04 AU, probably due to tidal interaction with the star (Plavchan & Bilinski, 2013).

- Approximately one third of the ~ 50 hot Jupiters with measured obliquities[4] are on orbits that are misaligned with the stellar rotation axis (Winn et al., 2010; Schlaufman, 2010; Albrecht et al., 2012). One possible explanation is that most orbits are misaligned initially, but become aligned over time through the action of tidal forces. This hypothesis is based on an apparent correlation between the orbit misalignment and the tidal timescale of the system. If borne out, it could provide evidence that dynamical scattering, rather than inward migration through the protoplanetary disc, is primarily responsible for transporting hot Jupiters close to their host stars.

- The occurrence rate of hot Jupiters, and gas giants in general, is positively correlated with host star metallicity (Fischer & Valenti, 2005). This supports core accretion theories of planet formation, which predict that the efficiency of massive planet formation increases as the metallicity of the protoplanetary disc increases (e.g. Mordasini et al., 2009).

- A significant fraction of hot Jupiters have radii that are ~ 20–50% larger than predicted by evolution models (Fortney & Nettelmann, 2010), with the amount of radius inflation correlating with the incident stellar flux (Weiss et al., 2013). Various mechanisms have been suggested for injecting stellar energy into the planetary interior including downward flux of kinetic energy due to strong winds (Showman & Guillot, 2002) and Ohmic dissipation caused by interactions between ionised winds and the planetary magnetic field (Batygin & Stevenson, 2010).

The atmospheres of dozens of hot Jupiters have also been characterised in one form or another, primarily through transit, eclipse, and phase curve observations (see Seager & Deming, 2010, for a review). The basic theory behind these techniques is described in Sections 1.4.1 and 1.4.2 below.

[4]The obliquity is the angle between the stellar rotation axis and the orbit axis.

1.4 Transits, eclipses, and phase curves

Deriving an expression for time-dependent flux of a transiting exoplanet system can be divided into two parts. Firstly, the position of the planet with respect to the star must be determined as a function of time, which requires knowledge of the orbital motion. Secondly, the total system flux must be calculated, given the configuration of the star and planet at a given instant.

In the simplest case, a Keplerian orbit is assumed for the planetary motion.[5] Such an orbit describes an ellipse, with shape determined by the semimajor axis a and eccentricity e. To define the orientation of the orbit in 3D space, an additional three parameters are required: the inclination i, which gives the angle between the sky plane and orbital plane; the longitude of the ascending node Ω, which gives the orientation between an arbitrary reference vector on the sky plane and the ascending node of the orbit; and the argument of periapse ω, which gives the angle on the orbital plane traced from the ascending node to the periapse. Two additional parameters constrain the location of the planet on the orbit as a function of time: the orbital period P and the time of periapse passage T_0.

With a set of values for the seven parameters $\{a, e, i, \Omega, \omega, P, T_0\}$, it is fairly straightforward to calculate the projected distance on the sky plane between the planet and star as a function of time, as well as the planet's coordinate along the line-of-sight relative to the star (e.g. Murray & Correia, 2010). Typically, the longitude of the ascending node is fixed to $\Omega = \pi$. If the orbit is circular (i.e. $e = 0$), the argument of periapse ω can also be set arbitrarily. For example, if $\Omega = \pi$ and $\omega = 3\pi/2$, then the periapse epoch T_0 will coincide with the transit

[5] The assumption of Keplerian motion will be invalid if gravitational perturbations by additional companions in the system are non-negligible (e.g. Holman & Murray, 2005). However, for the three hot Jupiters considered in this thesis (HD 189733b, HD 209458b, HAT-P-32b), the orbits are consistent with being periodic and circular.

mid-time $T_{\rm mid}$.

The sky-projected distance between the star and planet can be used to calculate the amount of overlap between the two circular discs as seen by a distant observer.[6] If the flux from the stellar disc is approximated as being uniform, and the planetary disc is assumed to be uniformly opaque,[7] then the drop in the relative flux of the system δ_p during primary transit evolves as:

$$\delta_p(\alpha,\lambda) = \mathcal{A}(\alpha,\lambda)\left[\frac{R_p(\lambda)}{R_\star}\right]^2\left[1-\frac{F_p(\alpha,\lambda)}{F_\star(\lambda)}\right], \quad (1.1)$$

where α is the orbital phase angle of the planet, such that $\alpha = 0$ corresponds to superior conjunction (i.e. "full phase"); λ is the observation wavelength; $\mathcal{A}$ is the fraction of the planet disc that overlaps the stellar disc; F_p is the disc-averaged flux of the planet per unit area; and $F_\star$ is the stellar flux per unit area. If the nightside flux from the planet can be ignored, then $F_p/F_\star = 0$ during transit (i.e. $\alpha \approx \pi$). Note that according to Equation 1.1, the planetary radius can vary with wavelength. This is to allow for wavelength-dependent variations in the atmospheric opacity, and will be elaborated on in Section 1.4.1.

In practice, the assumption of the stellar flux being uniform is inadequate and it is necessary to account for the effects of limb darkening. For the lightcurves considered in this work, the stellar limb darkening profile is approximated using a quadratic law of the form:

$$\frac{I(\mu,\lambda)}{I(0)} = 1 - c_1(\lambda)\left(1-\mu\right) - c_2(\lambda)\left(1-\mu\right)^2, \quad (1.2)$$

where I is the stellar brightness as a function of $\mu = \cos\theta$, with θ being the angle between the normal to the stellar surface and the line-of-sight, and c_1 and c_2 are the wavelength-dependent limb darkening coefficients. Equation 1.2 describes a

[6] The assumption that the star and planet have spherical shapes implies that higher-order effects such as tidal bulges are negligible.

[7] In reality, the opacity of the planetary disc will decrease towards the edge, as the column density of the atmosphere decreases. These effects are addressed in Section 1.4.1.

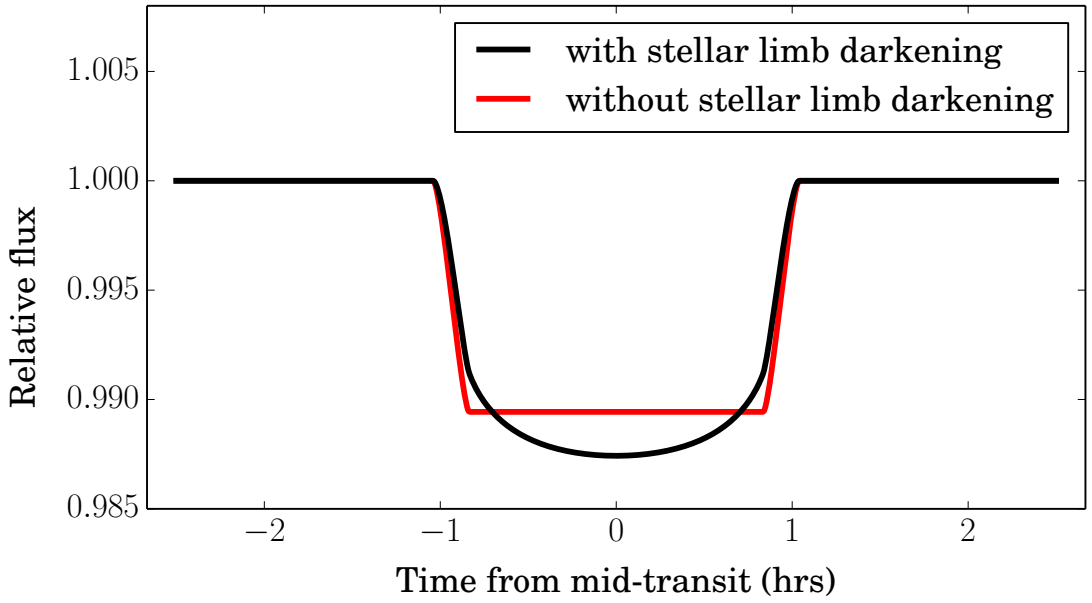

Figure 1.1: Example transit signals generated using the analytic equations of Mandel & Agol (2002) for a hot Jupiter with $R_p/R_\star = 0.1$, $a/R_\star = 8.8$, $e = 0$, $i = 88.7$ deg, and $P = 2.2$ day. The black line shows the signal obtained when a quadratic limb darkening law is used for the stellar brightness profile, with coefficients $c_1 = 0.35$ and $c_2 = 0.30$. The red line shows the signal that would be obtained for the exact same system parameters if the effects of limb darkening are ignored (i.e. $c_1 = 0$, $c_2 = 0$).

normalised radial profile that decreases towards the stellar limb, with contours of constant μ corresponding to annuli on the disc.

Calculation of the flux change during transit therefore requires integrating Equation 1.2 over the area eclipsed by the planet as a function of phase angle $\mathcal{A}(\alpha, \lambda)$. Without reproducing the mathematics here, Mandel & Agol (2002) provide analytic formulae for doing this, which are adopted in subsequent chapters. The equations can be parameterised in terms of $R_p/R_\star$ and $a/R_\star$, such that the model is fully constrained by the set of parameters $\{R_p/R_\star, c_1, c_2, a/R_\star, e, i, \omega, P, T_0\}$, with $\Omega = \pi$ as per convention. Figure 1.1 shows an example of a transit signal computed for a typical hot Jupiter system, including an illustration of the effect of stellar limb darkening.

At the other side of the orbit, when the planet disappears behind the star

during secondary eclipse, the relative change in the system flux is:

$$\delta_s(\alpha,\lambda) \;=\; \mathcal{A}(\alpha)\left(\frac{R_p}{R_\star}\right)^2 \frac{F_p(\alpha,\lambda)}{F_\star(\lambda)}\,. \tag{1.3}$$

Again, note that Equation 1.3 is only strictly valid if the hemisphere of the planet that is visible just before and after eclipse (i.e. the dayside hemisphere) has a uniform flux. If there are spatial inhomogeneities, then these should be integrated over the intersection area $\mathcal{A}$, producing subtle effects on the shape of the ingress and egress curves (e.g. Rauscher et al., 2007). Given the precision of current lightcurves, however, treating the planetary dayside hemisphere as a disc with uniform flux is a very good approximation. Thus, the eclipse component of the signal can be well modelled using the Mandel & Agol (2002) formulae with limb darkening coefficients set to zero.

Meanwhile, the phase-dependent planetary flux on the righthand sides of Equations 1.1 and 1.3 can be expressed as:

$$\frac{F_p(\alpha,\lambda)}{F_\star(\lambda)} \;=\; A_g(\lambda)\left(\frac{R_p}{a}\right)^2 \Phi(\alpha,\lambda)\,, \tag{1.4}$$

where A_g is the geometric albedo and Φ is a phase function normalised such that $\Phi(0,\lambda) = 1$. Note that it can be shown that Equation 1.4 follows from the definition of the geometric albedo: namely, the visible flux from a planet at full phase ($\alpha = 0$) relative to the total stellar flux incident upon it (e.g. Russell, 1916).

Defined in this way, the geometric albedo provides a means of expressing the reflectivity of a planet using practically observable quantities: namely, the normalised semimajor axis $a/R_\star$, which can be deduced from the transit, and the relative planet brightness at full phase $F_p/F_\star$, which can be deduced from the eclipse. Before proceeding, however, it is also worth mentioning two other types of albedo that arise frequently in the literature:

- The spherical albedo A_s, which gives the fraction of incident stellar radiation scattered back to space over all solid angles at a given wavelength. For

the case of a uniform, isotropically-scattering atmosphere, it can be shown that $A_g(\lambda) = \frac{2}{3} A_s(\lambda)$ (e.g. Seager, 2010).

- The Bond albedo A_B, which gives the fraction of incident stellar radiation scattered back to space over all solid angles and all wavelengths, such that $A_B = \int_0^\infty A_s(\lambda)\, d\lambda$. The Bond albedo determines the energy budgets of intensely irradiated planets such as hot Jupiters, as it quantifies the total amount of stellar energy that is absorbed by the planet (i.e. $1 - A_B$), which in turn drives the physics and chemistry of the atmosphere.

Returning to Equation 1.4, the flux from the planet $F_p(\alpha, \lambda)$ comprises two components: intrinsic thermal emission and scattered stellar radiation. The phase function can therefore be decomposed as:

$$\Phi(\alpha, \lambda) \;=\; \frac{1}{2} \left[\, \Phi_{\text{therm}}(\alpha, \lambda) \;+\; \Phi_{\text{scatt}}(\alpha, \lambda) \,\right] , \tag{1.5}$$

where $\Phi_{\text{therm}}(\alpha, \lambda)$ is the thermal phase function and $\Phi_{\text{scatt}}(\alpha, \lambda)$ is the scattering phase function. Both phase functions are normalised such that $\Phi_{\text{therm}} = \Phi_{\text{scatt}} = 1$ at superior conjunction.

For tidally locked hot Jupiters, variations in the measured thermal emission can be well approximated using a low-order expansion of sinusoidal basis functions, such that:

$$\Phi_{\text{thermal}}(\alpha, \lambda) \;=\; 1 + \sum_n a_n(\lambda) \cos(n\alpha) + \sum_n b_n(\lambda) \sin(n\alpha) , \tag{1.6}$$

with $n \leq 2$ proving adequate for currently available lightcurves (Cowan & Agol, 2008). Note that Equation 1.6 allows for the possibility of spatial inhomogeneities in the temperature map of the planet, such as multiple hot spots distributed in latitude and longitude.

Similarly, for the simple case of a uniform atmosphere that scatters isotropically at a given wavelength, it can be shown that the scattering phase function is

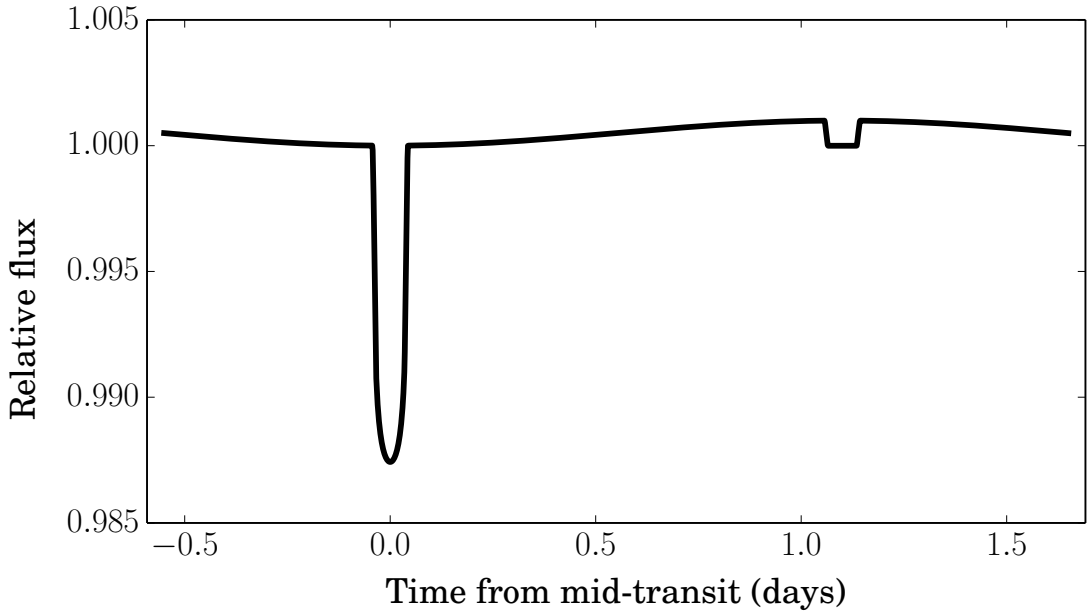

Figure 1.2: Example lightcurve for a transiting hot Jupiter covering a full orbit. The signal was generated using the analytic equations of Mandel & Agol (2002) and the same system parameters as Figure 1.1, with an eclipse depth of $\delta_s = 0.1\%$ and a sinusoidal phase function. Note the rounded shape of the primary transit due to stellar limb darkening, compared to the flat-bottomed secondary eclipse for which planetary limb darkening is ignored.

given by:

$$\Phi_{\text{scatt}}(\alpha, \lambda) \quad = \quad \frac{1}{\pi} \left[\sin\alpha + (\pi - \alpha) \cos\alpha \right] \ . \tag{1.7}$$

However, cloud particles in solar system planets tend to scatter relatively strongly in the backwards direction, causing the phase function to be more strongly peaked at superior conjunction (i.e. $\alpha = 0$) than Equation 1.7 would suggest (e.g. Sudarsky et al., 2005, and references therein). It is to be expected that hot Jupiters will also exhibit a range of different scattering phase functions, due to variations in the composition and structure of their atmospheres.

1.4.1 Atmosphere signals

Figure 1.2 draws together the elements of the previous section and illustrates the signal for a transiting hot Jupiter over a full orbit. It shows the primary transit, secondary eclipse, and a simple sinusoidal phase function. By accurately

measuring these lightcurve features, information can be extracted about the nature of the planetary atmosphere. Early works that laid out the fundamental theory for the characterisation of transiting exoplanet atmospheres include: Seager & Sasselov (1998, 2000); Seager et al. (2000); Marley et al. (1999); Sudarsky et al. (2000); Barman et al. (2001); Brown (2001). A brief overview is provided in the following sections.

1.4.1.1 Transmission

When the planet passes across the face of the star during primary transit, a fraction of the stellar radiation gets transmitted through the upper layers of the atmosphere. Variations in the transit depth as a function of wavelength therefore provide information about the wavelength-dependent opacity of the atmosphere, averaged over the limb of the disc. Sources of opacity can include the destruction of photons by absorption (e.g. transitions between electronic energy levels of atoms or rotation-vibration states of molecules) and scattering of photons out of the grazing beam (e.g. by cloud/haze particles).

Models for hot Jupiter atmospheres, such as those of Fortney et al. (2010), predict that the dominant opacity sources at visible wavelengths are the electronic transitions of alkali metals, most notably Na and K, as well as collision-induced absorption and Rayleigh scattering by H_2 molecules. Other possible absorbers include TiO and VO, although to date there have been no definitive detections of these species in a hot Jupiter atmosphere. Transitioning to the infrared, molecular rotation-vibration transitions become responsible for the bulk of absorption, with the most common species expected to be H_2O, CH_4, CO_2, and CO under conditions of chemical equilibrium. At all wavelengths, clouds/hazes of condensate species can potentially provide additional sources of opacity.

To estimate the expected signal magnitude, it is helpful to introduce the gaseous atmospheric scale height H, which corresponds to the vertical distance

over which the gas pressure decreases by a factor e^{-1}. For an atmosphere in hydrostatic equilibrium, it is straightforward to show that this is equal to:

$$H = \frac{kT}{\mu g}, \tag{1.8}$$

where k is the Boltzmann constant, T is the gas temperature, μ is the mean molecular mass, and g is the local gravitational acceleration (e.g. Seager, 2010).

A crude estimate for the atmospheric scale height can be obtained using the planet's equilibrium temperature $T_{\rm eq}$, i.e. the temperature that the planet would have if it radiated as a blackbody and was in thermal equilibrium. If the star is approximated as a blackbody with effective temperature $T_{\rm eff,\star}$, then according to energy balance:

$$T_{\rm eq} = T_{\rm eff,\star} \left(\frac{R_\star}{a}\right)^{1/2} \left[f_R\,(1-A_B)\right]^{1/4}, \tag{1.9}$$

where f_R is a parameter that quantifies the efficiency of heat redistribution from the dayside to nightside. For uniform heat redistribtion $f_R = 1/4$. Alternatively, if all radiation is re-emitted from the dayside hemisphere $f_R = 2/3$. Equilibrium temperatures obtained by setting $A_B = 0$ and $f_R = 1/4$ are given in Table 1.1 for HD 189733b (Chapters 3 and 4), HD 209458b (Chapter 3), and HAT-P-32b (Chapter 5). Substituting these temperatures into Equation 1.8, along with the gravitational accelerations g for each planet, we obtain the gas scale heights H that are also listed in Table 1.1.

To estimate the expected wavelength dependence of the transmission signal, we now consider an atmosphere with opacities κ_1 and κ_2 at two different wavelengths. It follows that the effective planetary radii[8] measured for those wavelengths will differ by $N = \ln(\kappa_1/\kappa_2)$ scale heights (e.g. Brown, 2001). Given that prominent absorption features in hot Jupiter transmission spectra are expected to have opacities that are enhanced by a factor of $\sim$ 10–100 relative to the

[8]The effective planetary radius is taken to be the radius at which a grazing beam through the atmosphere has an optical depth of unity.

Table 1.1: Basic system parameters for HD 189733, HD 209458, and HAT-P-32. Values for $T_{\mathrm{eff},\star}$, $R_\star$, R_p, and M_p have been taken from Torres et al. (2008) for HD 189733 and HD 209458, and from Hartman et al. (2011) for HAT-P-32. These parameters have in turn been used to derive the values listed for $\log g$, H (Equation 1.8), and T_{eq} (Equation 1.9). Note that the equilibrium temperatures T_{eq} were calculated using $A_B = 0$ and $f_R = 1/4$.

Parameter	Unit	HD 189733	HD 209458	HAT-P-32
V	mag	7.6	7.6	11.4
K	mag	5.5	6.3	10.0
$T_{\mathrm{eff},\star}$	K	5040	6070	6210
$R_\star$	$R_\odot$	0.8	1.1	1.2
R_p	R_J	1.1	1.4	1.8
M_p	M_J	1.1	0.7	0.9
$a/R_\star$	—	8.8	8.8	6.1
$\log g$	cgs	3.3	3.0	2.8
H	km	230	540	1060
T_{eq}	K	1200	1450	1780

continuum (e.g. see Figure 3 in Fortney et al., 2008), this translates to $N \sim 2$–5 scale heights. For the narrow cores of strong absorption features, such as the Na doublet at 589 nm, the opactiy ratio can reach more like $\kappa_1/\kappa_2 \sim 10^4$, such that $N \sim 10$. From Equation 1.1 it follows that an increase in the effective planetary radius by N atmospheric scale heights approximately corresponds to an increase in the measured transit depth of:

$$\Delta\delta_p = \left(\frac{R_p + NH}{R_\star}\right)^2 - \left(\frac{R_p}{R_\star}\right)^2 \approx \frac{2NH}{R_p}\left(\frac{R_p}{R_\star}\right)^2 \quad , \tag{1.10}$$

where it has been assumed that $(H/R_\star)^2 \approx 0$. Substituting the values from Table 1.1 into Equation 1.10 and adopting $N = 3$, we obtain expected transmission signals of $\Delta\delta_p = 340$ ppm for HD 189733b, $\Delta\delta_p = 560$ ppm for HD 209458b, and $\Delta\delta_p = 1170$ ppm for HAT-P-32b. To rigorously extract information about the chemical composition and temperature structure of the upper atmosphere from a set of transmission measurements at different wavelengths, it is necessary to

perform a so-called “atmosphere retrieval”. Methods for doing this are described further in Section 1.4.2.

The first successful measurement of opacity variations in the atmosphere of an exoplanet was made by Charbonneau et al. (2002), who used the *Hubble Space Telescope* (*HST*) Space Telescope Imaging Spectrograph (STIS) to measure an excess in transmission absorption due to the Na 589 nm doublet for HD 209458b. Following this pioneering result, STIS has proven to be a productive instrument for performing transmission spectroscopy (e.g. Sing et al., 2011b, 2013; Huitson et al., 2013; Nikolov et al., 2014). Numerous observations have also been made across the ultraviolet-to-near-infrared wavelength range using the other spectroscopic facilities of *HST*: the Advanced Camera for Surveys (ACS) (e.g. Pont et al., 2008a), the Near Infrared Camera and Multi-Object Spectrometer (NICMOS) (e.g. Swain et al., 2008b; Tinetti et al., 2010), and the Wide Field Camera 3 (WFC3) (e.g. Berta et al., 2012; Deming et al., 2013). Further into the infrared, multi-wavelength transit observations have been made using two *Spitzer Space Telescope* instruments: the Infrared Array Camera (IRAC) (e.g. Désert et al., 2009, 2011b; Beaulieu et al., 2010) and the Multiband Imaging Photometer for Spitzer (MIPS) (Richardson et al., 2006; Crossfield et al., 2012b). A large fraction of these *HST* and *Spitzer* observations have targeted the hot Jupiters HD 189733b and HD 209548b, and are reviewed in more detail in Section 3.2.

Ground-based observations have also been used to detect exoplanet atmospheres, starting with Redfield et al. (2008) who used high-resolution spectroscopy to measure Na absorption in transmission at 589 nm for HD 189733b. Similar observations have since revealed Na absorption in the atmosphere of HD 209458b (Snellen et al., 2008; Jensen et al., 2011). The first detection of K absorption was made by Sing et al. (2011a) by observing three transits of XO-2b simultaneously with two narrow-bandpass tunable filters centered around the

766 nm doublet. High-resolution spectroscopy has been used by Snellen et al. (2010a) to detect CO absorption in the atmosphere of HD 189733b by measuring the cross-correlation between time series spectra acquired during transit with model spectra containing forests of Doppler-shifted absorption lines.

Another technique that has risen to prominence in recent years is differential spectrophotometry, with time series spectra acquired for the target and one or more comparison stars over the course of a transit. The comparison stars provide a reference spectrum that can be used to calibrate out the systematics from the target fluxes. This technique was first applied successfully by Bean et al. (2010), who measured a featureless transmission spectrum for the super-Earth[9] GJ 1214b over the 0.8–1.0 μm wavelength range. The lack of detected absorption features was interpreted as evidence that they are either obscured by an unknown opacity source, such as a high-altitude cloud layer, or that the mean molecular weight of the atmosphere (i.e. μ in Equation 1.8) is high, resulting in a small atmospheric scale height H.

Using the same differential spectrophotometry method, flat transmission spectra have been measured for the transiting gas giants WASP-29b (Gibson et al., 2013a) and HAT-P-32b (Gibson et al., 2013b). The latter measurements have been interpreted as evidence for clouds, because the atmospheres are predominantly composed of H_2 and should thus have large scale heights with correspondingly large absorption features. Further details of the differential spectrophotometry technique are given in Chapter 5, along with new observations for HAT-P-32b made with the *Nordic Optical Telescope*.

1.4.1.2 Emission

The secondary eclipse depth δ_s gives the dayside hemisphere flux of the planet (Equation 1.3). Due to the zenith viewing angle, the emission probed by the

[9]The term super-Earth refers to planets with masses in the range 2–10 Earth masses.

secondary eclipse originates from layers that are considerably deeper in the atmosphere than those probed by the grazing geometry of primary transit. Emission measurements thus provide a greater insight into the vertical structure and bulk composition of the upper atmosphere.

For thermal emission, the brightness of the dayside hemisphere is determined by the temperature of the planetary photosphere. Given that the altitude of the photosphere is a function of the atmosphere opacity, which varies as a function of wavelength, measurements of the dayside thermal emission at multiple wavelengths provide constraints for the vertical temperature profile of the atmosphere. Such measurements also constrain the wavelength-dependence of the column-integrated opacity, and thus the abundances for the absorbing species in the atmosphere. Extracting this information, however, is a degenerate problem, discussed further in Section 1.4.2.

The first detection of thermal emission from an exoplanet was made by Charbonneau et al. (2005) by detecting the secondary eclipse of TrES-1b in the 4.5μm and 8.0μm wavelength channels of *Spitzer*/IRAC. Since then, secondary eclipses have been measured for over 40 exoplanets using *Spitzer*'s infrared instruments: IRAC (e.g. Knutson et al., 2008; Todorov et al., 2010; Blecic et al., 2014); MIPS (e.g. Charbonneau et al., 2008; Deming et al., 2005); and the Infrared Spectrograph (IRS) (e.g. Richardson et al., 2007; Grillmair et al., 2008). The majority of these observations have been made for hot Jupiters, which tend to produce relatively deep eclipses of $\delta_s \sim 1000$ ppm in the IRAC bandpasses, and even deeper eclipses at longer wavelengths where the planet-star flux contrast improves. Spectroscopic secondary eclipse observations have also been made using the near-infrared instruments of *HST*: namely, NICMOS (e.g. Swain et al., 2009a,b) and WFC3 (e.g. Wilkins et al., 2014; Ranjan et al., 2014). As with the space-based transmission spectroscopy, many of the *Spitzer* and *HST* ther-

mal emission measurements have targeted the hot Jupiters HD 189733b and HD 209458b, and are described further in Section 3.2.

Thermal emission has been detected from the ground by observing secondary eclipses in broad bandpasses such as H and K (e.g. Snellen & Covino, 2007; de Mooij & Snellen, 2009; Anderson et al., 2010; Gibson et al., 2010a; Zhao et al., 2012; de Mooij et al., 2011, 2013; Chen et al., 2014). Ground-based eclipse measurements using spectroscopy have only had limited success (e.g. Crossfield et al., 2012a; Bean et al., 2013) or else produced controversial results (e.g. Swain et al., 2010; Mandell et al., 2011; Waldmann et al., 2012). Meanwhile, detections of CO and H_2O absorption in a number of planetary dayside hemispheres have been made using the same high-resolution spectroscopy cross-correlation technique that was used to detect CO in transmission for HD 189733b (Section 1.4.1.1) (e.g. Brogi et al., 2012; Rodler et al., 2013; de Kok et al., 2013; Birkby et al., 2013).

The dayside emission may also include a contribution from stellar radiation that has been scattered back to space by the planetary atmosphere. The strength of this signal will depend upon the relative efficiency of scattering and absorption in the atmosphere, which is encoded in the geometric albedo (i.e. A_g in Equation 1.4). The wavelength dependence of the geometric albedo is controlled by the interplay between the scattering particles on the one hand, and the absorbing species located at higher levels in the atmosphere on the other. For instance, a wavelength-independent geometric albedo would be expected if there is an optically-thick layer of "white" cloud located at the very top of the atmosphere, as incident stellar radiation would be scattered back to space before it can be absorbed by other species. However, if the same layer of cloud was located at a lower altitude, the emergent spectrum would be sculpted by absorbing species located in the atmosphere above. The strength of the superimposed absorption

features will depend on factors such as the cloud altitude, which determines the column-density of atmosphere overhead.

In general, scattered light signals for hot Jupiters are much more difficult to measure than the thermal emission. Typical eclipse depths are $\lesssim 100$ ppm, which follows by substituting $R_p = 1\,R_J$, $a = 0.04$ AU, and $A_g = 0.5$ into Equation 1.4. Indeed, models predict the albedos of hot Jupiters will be low ($A_g < 0.2$) in the absence of high-altitude reflective cloud layers, due to alkali absorption at visible wavelengths and absorption by molecules such as H_2O and CH_4 at infrared wavelengths (e.g. Sudarsky et al., 2000). To date, most of the useful constraints on reflection from exoplanet atmospheres have been provided by space-based photometers, such as *Kepler* (e.g. Heng & Demory, 2013). These are reviewed in Chapter 4 as background to one of the main contributions of the present study: a spectroscopic eclipse observation for HD 189733b using *HST*/STIS (Evans et al., 2013), which represents the first multi-wavelength detection of reflected light for an exoplanet.

Meanwhile, in addition to the eclipse signal, the planetary phase curve provides further constraints on the atmospheric emission, with the amplitude of the variation giving the difference between the disc-integrated fluxes of the dayside and nightside hemispheres. Combined with the eclipse depth, which provides the dayside flux, this allows the nightside flux to be determined.

For a tidally-locked planet with a static brightness distribution, the phase curve period will equal the orbital period of the planet (Equation 1.6). In the case of a simple brightness distribution, peaking at the substellar point and decreasing smoothly towards the antistellar point, the phase curve minimum and maximum will coincide with the transit and eclipse mid-times, respectively (Figure 1.2). However, if the maxima and minima of the brightness distribution are offset with respect to the substellar and antistellar points, the measured phase curve will

also be shifted.

For thermal emission, the phase curve offset provides an insight into the atmospheric circulation of the planet (e.g. Showman & Guillot, 2002; Showman et al., 2009; Perna et al., 2012). The magnitude of the shift is governed by the ratio between the radiative and advective timescales in the atmosphere: a significant shift can be expected if the advective timescale is small relative to the radiative timescale.

Knutson et al. (2007a) were the first to measure the thermal phase curve of an exoplanet, using *Spitzer*/IRAC to monitor the system flux for HD 189733 over half an orbit in the 8.0 μm bandpass. The authors found that the hot spot was offset from the substellar point by 16 ± 6 deg, consistent with advection by strong winds as predicted by circulation models (e.g. Showman et al., 2009). Since then, thermal phase curves have also been published for HD 149026b (Knutson et al., 2009a), WASP-12b (Cowan et al., 2012), HAT-P-2b (Lewis et al., 2013), WASP-18b (Maxted et al., 2013), and HD 209458b (Zellem et al., 2014). In most cases, the measured phase curves exhibit significant hot spot offsets, indicating advection by winds. The only exception is WASP-18b, for which the lack of hot spot offset is accompanied by a high day-night temperature contrast, implying a low efficiency for heat redistribution between the two hemispheres.

Broadly speaking, circulation models predict that both the radiative and advective timescales increase with depth in the atmosphere, but the radiative timescale increases more rapidly (Showman et al., 2011). Consequently, larger thermal phase curve offsets can be expected as the altitude of the photosphere decreases. This means that, in principle, a comparison of phase curve offsets at different wavelengths can provide constraints for the wavelength-dependent opacity of the atmosphere, which is primarily responsible for determining the photosphere altitude. However, the picture is complicated by other factors,

such as possible variations in the atmospheric drag at different altitudes, which would affect advective timescales (e.g. Rauscher & Menou, 2012). So far, only tentative steps have been taken towards constraining atmospheric circulation and chemistry using multi-wavelength thermal phase curve measurements (e.g. Knutson et al., 2012).

For scattered radiation, offsets in the phase curve relative to the substellar point will be caused by spatial variations in the reflectivity of the dayside atmosphere. For instance, Demory et al. (2013) measured an offset of 41 ± 12 deg for the phase curve maximum of Kepler-7b in the *Kepler* bandpass, spanning the wavelength range 400–900 nm. The authors interpret this as the signature of a large-scale asymmetry in the distribution of reflective clouds in the dayside atmosphere. Measurements such as these thus have the potential to provide insight into the weather patterns and atmospheric chemistry of exoplanets. Also, as was noted above, the shape of the phase curve is determined by the properties of the dominant scattering species in the atmosphere, e.g. a non-sinusoidal phase curve that peaks sharply at or near superior conjunction is indicative of strong back scattering. However, no reflection phase curve measurements have yet been made at the precision necessary to resolve a non-sinusoidal shape.

1.4.2 Atmosphere retrievals

Atmosphere retrievals provide a means to reverse-engineer information about the atmosphere from measurements of its transmission (Section 1.4.1.1) and emitted flux (Section 1.4.1.2). Retrievals work by parameterising the atmosphere in terms of its basic properties, such as the pressure-temperature (PT) profile and the abundances of major chemical species. In principle, these properties could vary within the atmosphere in 3D space and time; in practice, however, current retrieval methods assume they are static and only vary along the vertical spatial

axis. These assumptions are justified by the limited precisions, resolutions, and wavelength coverages of available datasets, as well as computational tractability.

There are two common ways to specify the PT profile. The first approach is to divide the atmosphere into a series of horizontal layers stacked vertically, and assume that the pressure and temperature are constant within each layer (e.g. Irwin et al., 2008; Lee et al., 2012; Barstow et al., 2014). For example, Barstow et al. (2014) use a layered atmosphere with 50 pressure levels. Under the assumption of radiative equilibrium, flux is conserved across layer boundaries, giving a continuous PT profile. The second approach is to design a simple parameterisation for the PT profile (e.g. Madhusudhan & Seager, 2009; Benneke & Seager, 2012; Line & Yung, 2013). Using a small number ($\sim$ 5–10) of free parameters, a functional expression can be designed with sufficient flexibility to capture the main features expected for an atmospheric PT profile, including a deep, adiabatic convective zone and a stratosphere that may or may not have a temperature inversion.

Gas abundances are generally assumed to be vertically uniform, and treated as free parameters in the retrieval. As was mentioned above, important sources of gas opacity for hot Jupiter atmospheres include absorption by atoms such as Na and K, and by molecules such as H_2O, CO_2, CO, CH_4, and NH_3. Collision-induced absorption and Rayleigh scattering by molecules, including H_2, must also be accounted for. Clouds/hazes can provide additional sources of opacity, and make the retrieval task significantly more challenging (e.g. Barstow et al., 2014).

Once the model has been defined, an output spectrum is produced by solving the radiative transfer equation along the appropriate line-of-sight, i.e. a grazing ray travelling through the uppermost levels of the atmosphere in transmission or a vertical ray travelling up through the atmosphere in emission. To account for geometrical effects, the output spectrum is computed by taking a weighted

sum of such beams around the limb of the planet in transmission or across the visible hemisphere for emission. The model spectrum is provided as input to a likelihood function that evaluates the quality of the fit to the data, such as the χ^2 (Equation 2.23). Model parameters can then be inferred by either optimising the likelihood function (e.g. Lee et al., 2012; Line & Yung, 2013; Barstow et al., 2014) or marginalising over the parameter space with Markov chain Monte Carlo sampling (Section 2.2.1) (e.g. Madhusudhan & Seager, 2009; Benneke & Seager, 2012).

1.5 Aims of this thesis

In this thesis, multi-wavelength transit and eclipse observations are analysed for the three hot Jupiters: HD 189733b, HD 209458b, and HAT-P-32b. The fundamental aim is to use the principles outlined in Section 1.4.1 to characterise the planetary atmospheres.

In Chapter 3, archival transit and eclipse datasets are analysed for HD 189733b and HD 209458b in each of the *Spitzer*/IRAC wavelength channels (3.6μm, 4.5μm, 5.8μm, 8.0μm). These analyses aim to re-evaluate a number of claims that have been made in the literature regarding the atmospheres. In particular, these include:

- a claimed detection of H_2O absorption in the transmission spectrum of HD 189733b (Tinetti et al., 2007);
- an alternative claim of CO absorption in the transmission spectrum of HD 189733b (Désert et al., 2009);
- a claimed detection of H_2O absorption in the transmission spectrum of HD 209458b (Beaulieu et al., 2010);

- the inference of a thermal inversion in the atmosphere of HD 209458b based on the eclipse data (Knutson et al., 2008).

In Chapter 4, a secondary eclipse observation made for HD 189733b using *HST*/STIS over the 290–570 nm wavelength range is presented. The aim of this observation is to measure the wavelength-dependent geometric albedo of the atmosphere and test the hypothesis of Pont et al. (2013) that the dayside hemisphere is dominated by a thick layer of reflective cloud/haze.

In Chapter 5, two primary transits are presented for HAT-P-32b, acquired with the 2.5m *Nordic Optical Telescope* Andalucia Faint Object Spectrograph in multi-object spectroscopy mode over the 414–702 nm wavelength range. The aim of these observations is to investigate the potential of small ground-based telescopes for characterising exoplanet atmospheres. The specific scientific aims are to constrain transmission features such as Rayleigh scattering by haze or H_2 molecules, and absorption by the Na 589 nm doublet.

A common aim for all lightcurve analyses presented in this work is to demonstrate the application of Gaussian process models for the robust treatment of instrumental systematics. The details of Gaussian processes, and the formulation of lightcurve modelling as a Bayesian inference problem more generally, comprise the subject material for the next chapter.

Chapter 2

Bayesian methods for lightcurve modelling

The main goal of this work is to characterise exoplanet atmospheres through the analysis of transit lightcurves. Unfortunately, the vast majority, if not all, of the lightcurves that have been measured for this purpose are affected by instrumental systematics with larger amplitudes than the planetary signals of interest. This is the case throughout the published literature, and is true of observations made from space with telescopes such as *Spitzer* (Chapter 3) and *Hubble* (Chapter 4), and observations made from the ground (Chapter 5). This is not surprising, given that none of the instruments used to date for characterising transiting exoplanet atmospheres were originally designed with this purpose in mind. A careful treatment of systematics is therefore of central importance when interpreting transit lightcurves and drawing conclusions about planetary atmospheres. In this chapter, I lay out some basic tools for performing this task within a Bayesian framework.

I start in Section 2.1 by outlining the basic principles behind the Bayesian approach to probabilistic modelling, and how it differs from the traditional frequentist view. In Section 2.2, three practical methods for marginalising the posterior distribution of a Bayesian model are described: Markov chain Monte Carlo (MCMC), nested sampling, and variational integration. As a demonstration,

I include here the transit lightcurve MCMC analysis that I performed for the hot Jupiter CoRoT-18b, which was published in Hébrard et al. (2011). In Section 2.3, Gaussian process models are introduced, and I explain how they can be used to simultaneously model planetary signals and poorly understood instrumental systematics in a robust manner. This methodology is adopted extensively in subsequent chapters. I close in Section 2.4 by briefly mentioning the `pyhm` and `gps` software packages, which have been custom-written in the Python programming language for this work.

2.1 Frequentist and Bayesian approaches to inference

Fundamental to the scientific endeavour is reasoning in the face of uncertainty. In our present context, we make a measurement of a transit lightcurve and use this to infer information about the properties of the star-planet system (Section 1.4). However, the measurement is necessarily imperfect – not only is there intrinsically random measurement error, but correlations in the data due to effects such as instrumental systematics can be degenerate with features of the astrophysical signal, reducing our ability to accurately determine certain planet properties. These sources of uncertainty must be propagated to our model parameter inferences in a principled manner.

The "frequentist" way of approaching this problem is to treat the measured data, which I will denote by the vector $\boldsymbol{d}$, as just one of many possible datasets that could have been measured and assume that the underlying parameters of our model, which I will denote by the vector $\boldsymbol{\alpha}$, have unique, fixed values. Estimates for the latter are obtained by maximising the likelihood function with respect to $\boldsymbol{\alpha}$, where the likelihood function is defined as the conditional probability distribution $p(\boldsymbol{d}|\boldsymbol{\alpha})$. The uncertainty in these estimates can then be evaluated by

considering the distribution of possible data sets. A practical means of achieving this is to bootstrap the residuals of the best-fit model multiple times and repeat the optimisation process for each resulting dataset. The spread in estimated parameter values then gives an indication of the appropriate errorbars. This type of approach is common in the exoplanet literature for transit lightcurve fitting (e.g. Gillon et al., 2007; Désert et al., 2009; Gibson et al., 2010b).

Under the "Bayesian" worldview, however, we consider only one dataset; namely, the one that has been observed, $\boldsymbol{d}$. We acknowledge that we cannot infer the values of the model parameters $\boldsymbol{\alpha}$ with infinite precision, and instead place probability distributions on them. We thus introduce the intuitive concept of probability as representing our degree of belief in something, rather than as merely being associated with the outcomes of random events. In doing so, the standard product and sum rules of probabilities can still be applied, leading to a self-consistent system for performing deductive reasoning (Cox, 1946). This may be concisely expressed using Bayes theorem, which states that:

$$p(\boldsymbol{\alpha}|\boldsymbol{d}) \quad = \quad \frac{p(\boldsymbol{d}|\boldsymbol{\alpha})\, p(\boldsymbol{\alpha})}{p(\boldsymbol{d})} \,, \tag{2.1}$$

where $p(\boldsymbol{\alpha})$ is the prior distribution over the parameters, $p(\boldsymbol{\alpha}|\boldsymbol{d})$ is the posterior distribution having seen the data, and $p(\boldsymbol{d})$ is the model evidence.

Simply evaluating the righthand side of Equation 2.1 for different parameter values $\boldsymbol{\alpha}$, however, would not be very illuminating. What we would really like to do is marginalise the posterior distribution, such that:

$$G \quad = \quad \int_{\Omega} p(\boldsymbol{\alpha}|\boldsymbol{d})\, d\boldsymbol{\alpha} \,, \tag{2.2}$$

where Ω is some volume of the parameter space that encapsulates a fraction G of the total probability mass. For instance, it is common to quote so-called 1σ uncertainties on the unknown parameters by setting $G = 0.68$, as this is (approximately) the probability mass contained within a standard deviation of

the mean for a normal distribution.[1] Whatever we choose G to be, marginalising the posterior allows us to quantitatively express plausible ranges of parameter values, given the observed data and our prior assumptions.

Marginalisation also allows us to integrate over nuisance parameters that we are not interested in, using:

$$p(\boldsymbol{\alpha}|\boldsymbol{d}) \quad = \quad \int p(\boldsymbol{\alpha}, \boldsymbol{\alpha}'|\boldsymbol{d})\, d\boldsymbol{\alpha}' \,, \tag{2.3}$$

where $\boldsymbol{\alpha}'$ denotes the nuisance parameters. This is a big advantage of the Bayesian approach – by marginalising over distributions of parameters, we can avoid over-fitting to the data. This is in contrast to the frequentist paradigm, where point estimates are obtained for all parameters of the model, and we must resort to ad hoc methods to determine those that are actually justified by the data and those that are not.

Not only does a Bayesian formulation allow us to refine our knowledge of the unknown parameters for a given model, it also allows us to rank the relative probabilities of different models. If we denote two models $\mathcal{M}_1$ and $\mathcal{M}_2$, we can show this by recasting Bayes theorem from Equation 2.1 to:

$$\frac{p(\mathcal{M}_1|\boldsymbol{d})}{p(\mathcal{M}_2|\boldsymbol{d})} \quad = \quad \frac{p(\boldsymbol{d}|\mathcal{M}_1)\, p(\mathcal{M}_1)}{p(\boldsymbol{d}|\mathcal{M}_2)\, p(\mathcal{M}_2)} \,, \tag{2.4}$$

where the ratio on the lefthand side is known as the Bayes factor and gives the relative probability between the two models; on the righthand side, $p(\mathcal{M}_1)$ and $p(\mathcal{M}_2)$ are our subjective priors for each model, while $p(\boldsymbol{d}|\mathcal{M}_1)$ and $p(\boldsymbol{d}|\mathcal{M}_2)$ are the model evidences. Note that the latter are the same as the model evidence $p(\boldsymbol{d})$ in the denominator of Equation 2.4, but here I have made explicit the fact that the model evidence is conditional on the form of the assumed model.

[1] Of course, there is no guarantee that $p(\boldsymbol{\alpha}|\boldsymbol{d})$ is normally distributed.

2.2 Methods for marginalising the posterior

Assuming we are in possession of a model that we believe can explain the observed data, our next task is to marginalise the posterior distribution $p(\boldsymbol{\alpha}|\boldsymbol{d})$ to identify plausible ranges for the parameter values. Naively, we might hope to do this by directly integrating the righthand side of Bayes theorem (Equation 2.1). However, such an approach is complicated by the presence of the model evidence $p(\boldsymbol{d})$ in the denominator, which can be re-written as:

$$p(\boldsymbol{d}) \quad = \quad \int p(\boldsymbol{d}|\boldsymbol{\alpha})\, p(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} \;, \tag{2.5}$$

i.e. the model evidence is the probability of observing a given dataset marginalised over every possible parameter value. Equation 2.5 is analytically intractable for almost all real world problems, including transit lightcurve modelling, while standard numerical integration techniques can rarely be applied in practice given their poor scaling with the dimensionality of the parameter space. In this section, I introduce three different methods that circumvent these problems in one way or another: Markov chain Monte Carlo, nested sampling, and variational approximation.

2.2.1 Markov chain Monte Carlo

Markov chain Monte Carlo (MCMC) methods are widely used for performing model inference in all areas of exoplanet science (e.g. Ford, 2005; Bouchy et al., 2008; Madhusudhan et al., 2011), and indeed, in astrophysics more generally (e.g. Tegmark et al., 2004; Walker et al., 2009; Courteau et al., 2011). They are based on the property that Markov chains generate samples from the posterior distribution when the sufficient condition of detailed balance is met; namely:

$$p(\boldsymbol{\alpha}_i|\boldsymbol{d})\, T(\boldsymbol{\alpha}_i \rightarrow \boldsymbol{\alpha}_j) \quad = \quad p(\boldsymbol{\alpha}_j|\boldsymbol{d})\, T(\boldsymbol{\alpha}_j \rightarrow \boldsymbol{\alpha}_i) \tag{2.6}$$

where $T(\boldsymbol{\alpha}_i \rightarrow \boldsymbol{\alpha}_j)$ is the transition probability between states $\boldsymbol{\alpha}_i$ and $\boldsymbol{\alpha}_j$.

A widely-used algorithm that satisfies detailed balance is the Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970). It proceeds by first taking a random draw from the prior $p(\boldsymbol{\alpha})$, denoted as $\boldsymbol{\alpha}_0$. A second random sample $\boldsymbol{\alpha}_1$ is then drawn from a proposal distribution conditional on $\boldsymbol{\alpha}_0$, denoted as $q(\boldsymbol{\alpha}|\boldsymbol{\alpha}_0)$, and accepted with probability:

$$p(\boldsymbol{\alpha}_0 \rightarrow \boldsymbol{\alpha}_1) \quad = \quad \min\left[1, \frac{q(\boldsymbol{\alpha}_1|\boldsymbol{\alpha}_0)\, p(\boldsymbol{\alpha}_1|\boldsymbol{d})}{q(\boldsymbol{\alpha}_0|\boldsymbol{\alpha}_1)\, p(\boldsymbol{\alpha}_0|\boldsymbol{d})}\right] . \tag{2.7}$$

If the step is accepted, the chain progresses to state $\boldsymbol{\alpha}_1$; otherwise, we set $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_0$. A new state $\boldsymbol{\alpha}_2$ is then randomly generated from the proposal distribution $q(\boldsymbol{\alpha}|\boldsymbol{\alpha}_1)$, and the process is repeated until some stopping criterion is met. The transition probability of the resulting Markov chain is thus given by:

$$T(\boldsymbol{\alpha}_i \rightarrow \boldsymbol{\alpha}_j) \quad = \quad q(\boldsymbol{\alpha}_j|\boldsymbol{\alpha}_i)\, p(\boldsymbol{\alpha}_i \rightarrow \boldsymbol{\alpha}_j) , \tag{2.8}$$

which satisfies the condition of detailed balance given by Equation 2.6. Thus, the resulting chain of states $\{\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots\}$ will eventually converge to the posterior distribution $p(\boldsymbol{\alpha}|\boldsymbol{d})$. Note in particular that at no point are we required to evaluate the problematic model evidence $p(\boldsymbol{d})$ term.

Ideally, we would like to have some way of determining whether or not the chain has converged to the true distribution at some level of tolerance. Of course, this is not possible, as the true distribution is unknown (otherwise we would not need to estimate it). It is possible, however, to prove that a chain has *not* converged. For instance, if multiple chains are run in parallel and have converged, they should have identical statistical properties. This is the principle behind diagnostics such as the Gelman-Rubin statistic (Gelman & Rubin, 1992), which compares the inter-chain variances and intra-chain variances for each parameter. If these quantities are found to differ by more than a few percent, it is a strong indication that at least one of the chains has not yet converged.

In terms of choosing a proposal distribution, a multivariate normal centered on the current state is often adopted, such that:

$$q(\boldsymbol{\alpha}|\boldsymbol{\alpha_i}) \quad = \quad \mathcal{N}(\boldsymbol{\alpha_i}, \mathbf{S}) \; , \tag{2.9}$$

where the righthand side denotes a multivariate normal distribution with mean vector $\boldsymbol{\alpha_i}$ and covariance matrix $\mathbf{S}$. It is common for a diagonal covariance matrix to be used, with $\mathbf{S} = \mathrm{diag}\left(s_1^2, \ldots, s_D^2\right)$ where the s_i terms are the step sizes for each of the D parameters. The latter are usually tuned so that the overall step acceptance rate is ~20–40%, which allows the chain to achieve optimal mixing throughout the parameter space (e.g. Gelman et al., 1996; Roberts et al., 1997). This can be achieved in practice by running short chains and iteratively adjusting the step sizes during the burn-in phase.

A problem with using a diagonal covariance matrix for the proposal distribution, however, is that it results in relatively inefficient sampling when significant correlations exist between the parameters $\boldsymbol{\alpha}$. MCMC algorithms that employ more sophisticated sampling routines to mitigate against such correlations, and offer other enhancements such as being parallelisable, are gaining increasing popularity in the exoplanet community (e.g. Ford, 2006; Eastman et al., 2013; Foreman-Mackey et al., 2013).

2.2.1.1 An example: CoRoT-18b

CoRoT-18b is a hot Jupiter planet in orbit around a G9V star that was discovered as part of the *CoRoT* mission (Section 1.2). It has mass $3.47 \pm 0.38\,M_\mathrm{J}$, radius $1.31 \pm 0.18\,R_\mathrm{J}$, and orbital period 1.9000693 ± 0.0000028 days. I performed the original MCMC lightcurve analysis and was second author on the discovery paper (Hébrard et al., 2011). It is included here as a simple demonstration of using MCMC to marginalise the posterior of a transit lightcurve model.

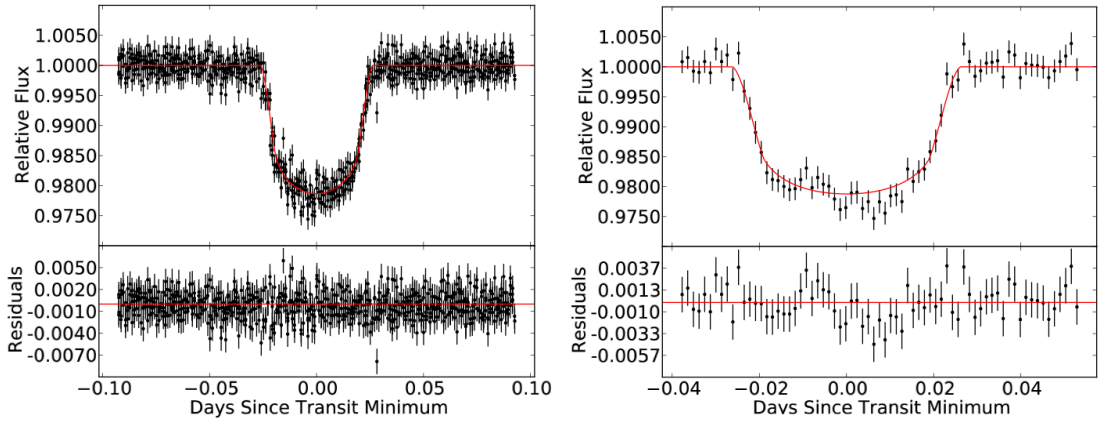

Figure 2.1: Primary transit lightcurves and best-fit residuals for CoRoT-18b obtained with the *CoRoT* (left) and *Euler* (right) telescopes. The *CoRoT* lightcurve consists of 13 consecutive transits, and has been phase-folded at the best-fit orbital period. Red lines show the transit models with parameters set to the maximum likelihood values.

Figure 2.1 shows the phase-folded lightcurve acquired with the CoRoT telescope, comprised of 13 consecutive transits. Also shown is an additional follow-up transit that was observed using the 1.2 m *Euler Telescope* at La Silla Observatory, Chile. Both lightcurves were modelled simultaneously, assuming a circular orbit and allowing five parameters to vary: the orbital period P, the time of central transit T_c, the normalised system scale $a/R_\star$, the impact parameter $b = a\cos i/R_\star$, and the radius ratio $R_p/R_\star$. A quadratic limb darkening law was used, with coefficients fixed to those given by Sing (2010) for the CoRoT lightcurve and to those given by Claret (2004) for the Euler lightcurve.

Maximum likelihood estimates (MLEs) for the unknown parameter values were obtained using the Levenberg-Marquardt nonlinear least squares algorithm (Levenberg, 1944). This allowed a suitable region of the parameter space in which to initialise the MCMC chains to be identified. The Metropolis-Hastings algorithm was then used with a proposal distribution of the form given by Equation 2.9. Step sizes were adjusted using a short burn-in chain to give acceptance rates of 20–40%. Five separate chains were initiated at different starting locations randomly perturbed away from the MLE values, and run in parallel for 2×10^5

Table 2.1: Results of the MCMC analysis for CoRoT-18b. Quoted values give the chain medians with uncertainties corresponding to the symmetric ranges about the median that encompass ±34% of the chain samples.

Free Parameters	Value
Planet orbital period, P (days)	$1.9000693^{+0.0000028}_{-0.0000029}$
Periastron passage, T_0 (HJD−2400000)	$55321.72412^{+0.00018}_{-0.00018}$
Scaled semimajor axis, $a/R_\star$	$6.49^{+0.32}_{-0.29}$
Radius ratio, $R_p/R_\star$	$0.1342^{+0.0018}_{-0.0018}$
Impact parameter, $b = a\cos i/R_\star$	$0.40^{+0.08}_{-0.13}$

Deduced Parameters	Value
Orbit inclination, i (degrees)	$86.5^{+1.2}_{-0.9}$
Transit duration, T_{14} (hours)	$2.389^{+0.036}_{-0.036}$
Ingress/Egress duration, $T_{12} = T_{34}$ (hours)	$0.332^{+0.038}_{-0.035}$
Stellar density, $\rho_\star$ (g cm^{-3})	$1.43^{+0.22}_{-0.18}$

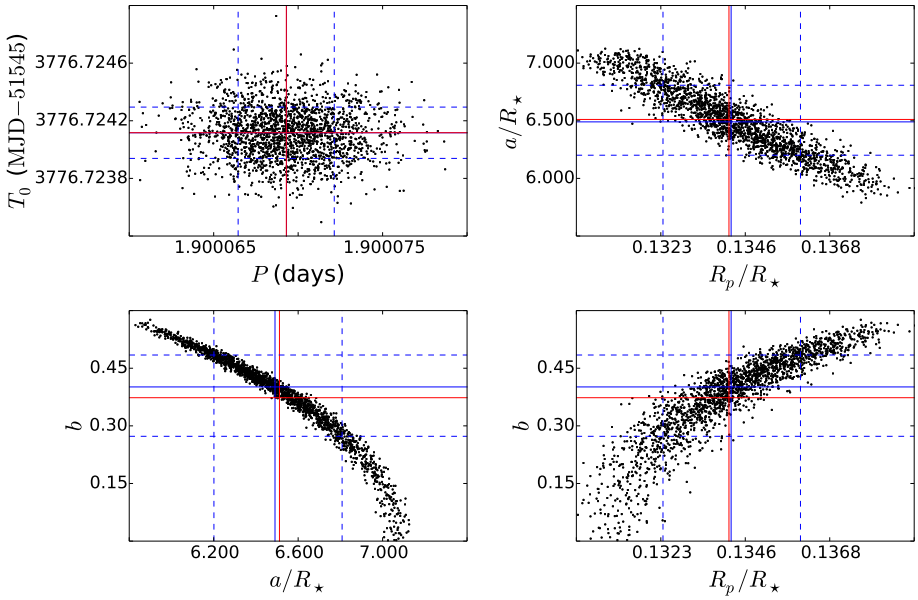

Figure 2.2: MCMC results for the CoRoT-18 lightcurve analysis. Black points show chain samples for different parameter pairs, representing random draws from the model posterior distribution. Solid blue lines indicate median chain values for each parameter, and blue dashed lines indicate the ranges about the median that encompass ±34% of the chain samples. Solid red lines show maximum likelihood estimates.

steps. By this point, the Gelman-Rubin statistic was within ~1% of unity for each of the free parameters, indicative of a stable state having been reached. The chains were then combined to obtain marginalised posterior distributions for each of the free parameters, with medians and uncertainties reported in Table 2.1 and examples plotted in Figure 2.2.

2.2.2 Nested sampling

A relatively new method for both marginalising the posterior $p(\boldsymbol{\alpha}|\boldsymbol{d})$ and evaluating the model evidence $p(\boldsymbol{d})$ at modest computational expense is nested sampling (Skilling, 2004). Although nested sampling has not been used in the present study, I briefly highlight it here in the interest of completeness and because it appears to be worthy of consideration for future lightcurve modelling work, given the advantages it offers. For instance, by computing model evidences, it becomes possible to perform model selection using Bayes factors (Equation 2.4), e.g. for choosing between two covariance kernels in a Gaussian process fit to a transit lightcurve (Section 2.3). A more technical treatment of nested sampling can be found in the original paper by Skilling, or alternatively, another readable account such as that by Allison & Dunkley (2014), in which nested sampling is applied to a simple multivariate normal dataset and a real-world cosmological dataset.

Nested sampling starts by defining the ordered prior mass according to:

$$X(\lambda) \quad = \quad \int_{p(\boldsymbol{\alpha}|\mathbf{d})>\lambda} p(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} \;, \tag{2.10}$$

where λ is some likelihood threshold value. If we use $\mathcal{L}(X)$ to denote the likelihood value corresponding to prior mass X, then it follows that the evidence is given by:

$$p(\boldsymbol{d}) \quad = \quad \int_0^1 \mathcal{L}(X)\, dX \;, \tag{2.11}$$

where the multi-dimensional integral of Equation 2.5 has been reduced to a one-dimensional integral. To evaluate this integral, we start by drawing N samples from the prior, and record the point $\boldsymbol{\alpha_0}$ that has the lowest likelihood $\mathcal{L}_0$. We then discard $\boldsymbol{\alpha_0}$ from the active set, and continue drawing from the prior until a new point is obtained with a likelihood greater than $\mathcal{L}_0$. At this point, we again record and remove the sample $\boldsymbol{\alpha_1}$ from the active set that has the lowest likelihood $\mathcal{L}_1$, and continue repeating the process.

Each of the likelihood values $\{\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2, \ldots\}$ generated in this manner will have associated prior mass values $\{X_0, X_1, X_2, \ldots\}$. Given how the former were generated, the latter will follow a sequence of the form:

$$X_0 = 1\,, \quad X_1 = t_1 X_0\,, \quad X_2 = t_2 X_1\,, \quad \ldots \tag{2.12}$$

where the t_i values are distributed as $p(t_i) = N t_i^{N-1}$. This in turn implies that the expectation values for the prior mass values are given by:

$$\mathbb{E}[X_i] \;=\; \exp\left(-\frac{i}{N}\right)\,. \tag{2.13}$$

Therefore, the algorithm proceeds by sampling from the prior within surfaces of increasing likelihood, or equivalently, decreasing prior mass – hence the name nested sampling.

If we then take $X_i \approx \mathbb{E}[X_i]$, Equation 2.11 can be approximated using trapezoidal integration as:

$$p(\boldsymbol{d}) \;=\; \sum_{i=1}^{M} w_i \mathcal{L}_i\,, \tag{2.14}$$

where $w_i = \frac{1}{2}(X_{i-1} - X_{i+1})$ and M is the number of iterations. For justifications of Equations 2.12 and 2.13, as well as discussions on choosing M according to a stopping criterion and quantifying the uncertainty associated with Equation 2.14, see Skilling (2004).

In addition to evaluating the model evidence, nested sampling also allows the normalised posterior distribution to be recovered. Specifically, the posterior value associated with state α_i is given by:

$$p(\boldsymbol{\alpha_i}|\boldsymbol{d})\, d\boldsymbol{\alpha} \;=\; \frac{\mathcal{L}_i\, w_i}{\sum_{i=1}^{M} \mathcal{L}_i\, w_i}\,, \tag{2.15}$$

which follows from Equation 2.14 and the fact that:

$$\mathcal{L}_i\, w_i \;=\; \mathcal{L}(X_i)\, dX \;=\; p(\boldsymbol{d}|\boldsymbol{\alpha_i})\, p(\boldsymbol{\alpha_i})\, d\boldsymbol{\alpha} \quad , \tag{2.16}$$

by construction. In practice, the posterior is marginalised by assigning each of the recorded samples $\{\boldsymbol{\alpha_0}, \boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \ldots\}$ with individual weights according to Equation 2.15 and then binning them into histograms, similar to what is done for the samples generated using MCMC.

2.2.3 Variational methods

Variational Bayes (VB) methods make use of variational calculus[2] to approximate the posterior distribution. To do this, the model evidence is first decomposed into:

$$p(\boldsymbol{d}) \;=\; \mathrm{KL}(q\|p) + L(q) \tag{2.17}$$

where q is some unspecified, but analytically tractable, function that depends on the model parameters $\boldsymbol{\alpha}$, and:

$$\mathrm{KL}(q\|p) \;=\; -\int q(\boldsymbol{\alpha}) \log\left[\frac{p(\boldsymbol{\alpha}|\boldsymbol{d})}{q(\boldsymbol{\alpha})}\right] d\boldsymbol{\alpha}\,, \tag{2.18}$$

$$L(q) \;=\; \int q(\boldsymbol{\alpha}) \log\left[\frac{p(\boldsymbol{d},\boldsymbol{\alpha})}{q(\boldsymbol{\alpha})}\right] d\boldsymbol{\alpha}\,. \tag{2.19}$$

The first term $\mathrm{KL}(q\|p)$ is a quantity known as the Kullback-Leibler divergence, and commonly arises in the field of information theory. It can be thought of as

[2]The calculus of functionals, rather than functions as in standard calculus.

a measure of the distance between the function $q(\boldsymbol{\alpha})$ and the posterior $p(\boldsymbol{\alpha}|\boldsymbol{d})$. Therefore, a function $q(\boldsymbol{\alpha})$ that minimises the Kullback-Leibler divergence will provide the best possible approximation to the true posterior. Given that the model evidence on the lefthand side of Equation 2.17 does not depend on $q(\boldsymbol{\alpha})$, this is equivalent to finding a function $q(\boldsymbol{\alpha})$ that maximises $L(q)$.

For certain cases, it is possible to analytically derive an iterative algorithm that is guaranteed to converge to a solution that maximises $L(q)$ and hence approaches the true posterior distribution. One such example is a linear basis model; namely, a model μ of the form:

$$\mu(\boldsymbol{v}) \quad = \quad w_0 + \sum_{i=1}^{M} \phi_i(\boldsymbol{v})\, w_i \ , \tag{2.20}$$

where $\boldsymbol{v}$ is the input vector, the ϕ_i terms are the basis functions, and the w_i terms are the linear weights. I do not derive the details of the algorithm here, but they can be found in a reference such as Bishop (2006).

I made considerable effort to investigate the possibility of using VB with linear basis models of the form given by Equation 2.20 to account for poorly understood systematics in transit lightcurves. I used tightly overlapping Gaussians for the basis functions ϕ_i with amplitudes controlled by the linear weights w_i, resulting in models that were capable of modelling an extremely broad range of smooth signals. However, I encountered a number of difficulties with this approach. Firstly, the transit signal is nonlinear in the parameters of interest, such as $R_p/R_\star$, $a/R_\star$, etc., and hence cannot be incorporated into the VB framework. This meant that I had to iterate between modelling the transit signal with an approach like MCMC and treating the systematics with VB. Secondly, a large number of overlapping Gaussian basis functions were often required to adequately reproduce the structure of the systematics component, causing the computation times to become impractically large.

2.3 Gaussian processes for transit lightcurves

Formally, a Gaussian process (GP) is defined as a collection of data points, any subset of which has a multivariate normal distribution (e.g. Rasmussen & Williams, 2006). One of the greatest advantages of GP models is that they allow us to naturally handle correlations in the data that may be poorly understood from a first principles standpoint by specifying only high-level properties of the covariance. As Gibson et al. (2012b) have pointed out, this makes GPs particularly suitable for modelling many of the transit lightcurves that have been measured to date.

Let us start by considering N flux measurements denoted by the vector $\boldsymbol{d} = [\, d_1, d_2, \ldots, d_N \,]^{\mathrm{T}}$. For each data point, we have $d_i = f_i + \varepsilon_i$, where f_i is the underlying non-random component of the signal (i.e. the latent signal) and ε_i is some unknown random measurement error, which obeys Poisson counting statistics. In the limit of large photon numbers, we can approximate the Poisson errors as being normally distributed (i.e. Gaussian errors), such that $p(\varepsilon_i) = \mathcal{N}(0, \sigma_i^2)$, where σ_i is the uncertainty associated with the ith measurement. If we assume the measurement uncertainties are independent, the likelihood function is given by:

$$p(\boldsymbol{d}|\boldsymbol{\alpha}) \;=\; \prod_{i=1}^{N} \mathcal{N}\left(d_i - \mu_i, \sigma_i^2\right) \;=\; \mathcal{N}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \,, \tag{2.21}$$

where μ_i is the model value associated with the ith data point, $\boldsymbol{\Sigma} = \mathrm{diag}\left(\sigma_1^2, \ldots, \sigma_N^2\right)$. The likelihood of the dataset is therefore given by a multivariate normal distribution, thus satisfying the definition of a GP given above.

Writing out the log likelihood in full, we have:

$$\ln p(\boldsymbol{d}|\boldsymbol{\alpha}) \;=\; -\frac{1}{2}\boldsymbol{r}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{r} - \frac{1}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}N\ln 2\pi \,, \tag{2.22}$$

where $\boldsymbol{r} = \boldsymbol{d} - \boldsymbol{\mu}$ are the model residuals. In particular, we note that the first term on the righthand side of Equation 2.22 is proportional to the negative χ^2 statistic,

which is widely used for model fitting in the exoplanet literature. Explicitly:

$$\chi^2 \;=\; \sum_{i=1}^{N} \left(\frac{r_i}{\sigma_i} \right)^2 \;=\; \boldsymbol{r}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \boldsymbol{r} \;. \tag{2.23}$$

If the covariance matrix $\boldsymbol{\Sigma}$ is fixed, it follows that maximising $p(\boldsymbol{d}|\boldsymbol{\alpha})$ in Equation 2.22 with respect to the model parameters $\boldsymbol{\alpha}$ is exactly equivalent to the familiar process of minimising the χ^2 statistic.

In the real world, the latent signal $\boldsymbol{f} = \left[f_1, f_2, \ldots, f_N \right]^{\mathrm{T}}$ will be a combination of the planet signal and systematics (i.e. correlated "noise"). If we formulate our model using Equation 2.21, it is necessary to combine both of these components into the mean function $\boldsymbol{\mu}$, such that:

$$\mu \;=\; (E \times H) + B \;, \tag{2.24}$$

where E is the astrophysical signal of interest, and H and B are multiplicative and additive systematics terms, respectively. However, this requires that we understand the systematics well enough to be in possession of a functional form that describes them, which is very rarely the case. A common compromise is to instead approximate the systematics as low-order polynomials of variables that are believed to correlate with the measured signal, such as xy pixel coordinates of the PSF on the detector, airmass, atmospheric humidity, space-telescope orbital phase, etc. The dangers of this approach have been highlighted by Gibson et al. (2011). In particular, a low-order polynomial decorrelation may be inadequate for modelling the systematics, while adding higher-order terms runs the risk of overfitting to the data. Gibson et al. showed that certain claims of molecular absorption features detected in exoplanet atmospheres (e.g. Swain et al., 2008b; Tinetti et al., 2010) could potentially be attributable to effects such as these, rather than genuine astrophysical signals.

An alternative option is to incorporate the systematics into our model by allowing for nonzero off-diagonal entries in the covariance matrix of the likelihood

function, such that:

$$p(\boldsymbol{d}|\boldsymbol{\alpha},\boldsymbol{\gamma}) \quad = \quad \mathcal{N}\left(\boldsymbol{\mu}, \mathbf{K}+\boldsymbol{\Sigma}\right) , \tag{2.25}$$

where K_{ij} gives the covariance between the ith and jth data points, and $\boldsymbol{\gamma}$ are parameters that control the behaviour of the covariance. The log likelihood function of Equation 2.22 then generalises to:

$$\ln p(\boldsymbol{d}|\boldsymbol{\alpha},\boldsymbol{\gamma}) \quad = \quad -\frac{1}{2}\boldsymbol{r}^{\mathrm{T}}(\mathbf{K}+\boldsymbol{\Sigma})^{-1}\boldsymbol{r} - \frac{1}{2}\ln|\mathbf{K}+\boldsymbol{\Sigma}| - \frac{1}{2}N\ln 2\pi \ . \tag{2.26}$$

Although we no longer have to provide an explicit functional form for the systematics contribution, we must now specify a kernel function to populate the off-diagonal terms of the covariance matrix **K** (see Section 2.3.1). However, by modelling the covariance rather than the systematics signal directly, it is possible to capture a broad range of behaviours with relatively few free parameters. Thus, GP models are simultaneously parsimonious and highly flexible.

2.3.1 Covariance kernels

Entries of the covariance matrix **K** are constructed using a kernel function, such that $K_{ij} = k(\boldsymbol{v_i}, \boldsymbol{v_j})$ where $\boldsymbol{v_i}$ and $\boldsymbol{v_j}$ are the inputs associated with the ith and jth data points, respectively. By "inputs", I refer to variables that correlate with the measured signal – these are often the same variables that would be used for a standard polynomial decorrelation, examples of which were mentioned above. The only requirement for a valid covariance kernel is that the resulting matrix **K** be positive semi-definite, and hence invertible, so that the likelihood function can be evaluated (Equation 2.26). In this section, I describe two of the most common covariance kernels used in the machine learning literature, which are employed for the analyses presented in later chapters. A much fuller discussion of covariance kernels can be found in Rasmussen & Williams (2006).

The first kernel we consider is the squared exponential, which is defined as:

$$k\left(\boldsymbol{v_i}, \boldsymbol{v_j}\right) \;=\; A^2 \exp\left[-\left(\frac{r_{ij}}{L}\right)^2\right], \tag{2.27}$$

where the modulus $r_{ij} = \| \boldsymbol{v_i} - \boldsymbol{v_j} \|$ gives the Euclidean distance between the two input vectors, A is the covariance amplitude, and L is the characteristic length scale. According to Equation 2.27, the covariance exhibits the sensible behaviour of decreasing exponentially as the distance between two data points increases. This results in functions that are very smooth, examples of which are illustrated in Figure 2.3.

The second kernel we consider is the Matérn function, with general form given by:

$$k\left(\boldsymbol{v_i}, \boldsymbol{v_j}\right) \;=\; A^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left[\frac{r_{ij}}{L}\sqrt{2\nu}\right]^{\nu} K_{\nu}\left(\frac{r_{ij}}{L}\sqrt{2\nu}\right), \tag{2.28}$$

where K_{ν} is the modified Bessel function of the second kind, Γ is the Gamma function, L is the covariance length scale, and ν is a non-negative parameter controlling the smoothness of the signal. As $\nu \to \infty$, the Matérn kernel converges to the squared exponential kernel (Equation 2.27), while as $\nu \to 0$ it describes functions that become increasingly "rough". We will restrict our attention to the specific case where $\nu = 3/2$, in which case Equation 2.28 simplifies to:

$$k\left(\boldsymbol{v_i}, \boldsymbol{v_j}\right) \;=\; A^2 \left[1 + \frac{r_{ij}}{L}\sqrt{3}\right] \exp\left(-\frac{r_{ij}}{L}\sqrt{3}\right). \tag{2.29}$$

This kernel is more appropriate than the squared exponential for handling high-frequency correlations in transit lightcurves, e.g. due to aperture losses from short timescale pointing jitter or variable atmospheric conditions. Examples of signals generated by the Matérn $\nu = 3/2$ kernel are shown in Figure 2.4.

For both the squared exponential and Matérn kernels, we can assign separate length scales for each of the inputs by replacing the r_{ij}/L terms in Equations

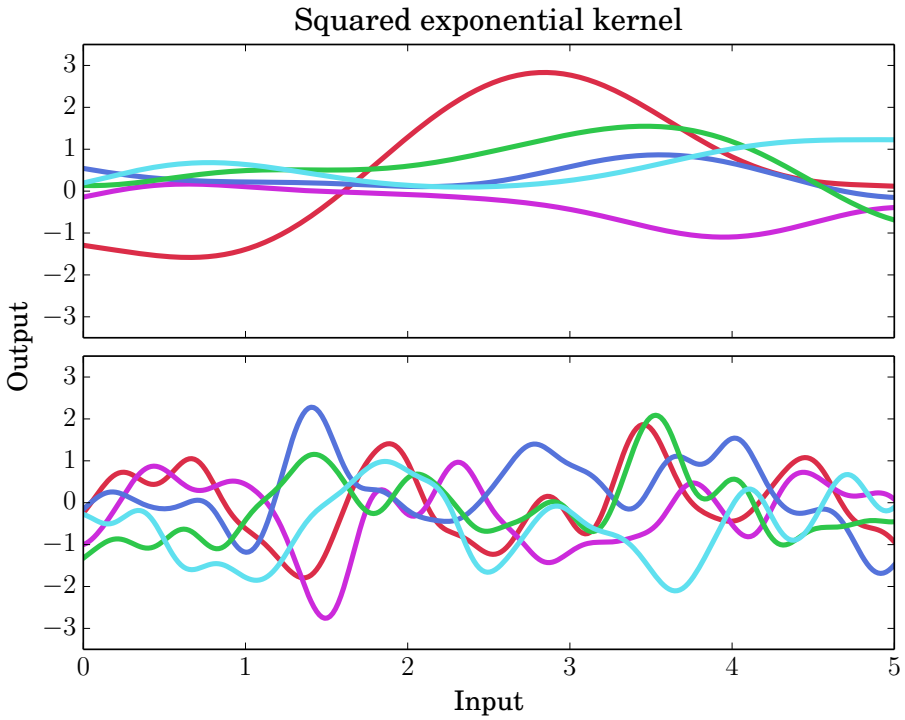

Figure 2.3: (Top) Random draws from a noiseless, zero-mean GP, i.e. likelihood function given by $\mathcal{N}(\mathbf{0}, \mathbf{K})$, with a squared exponential kernel, defined according to Equation 2.27. (Bottom) The same as the top panel, but with a covariance length scale L that is five times smaller.

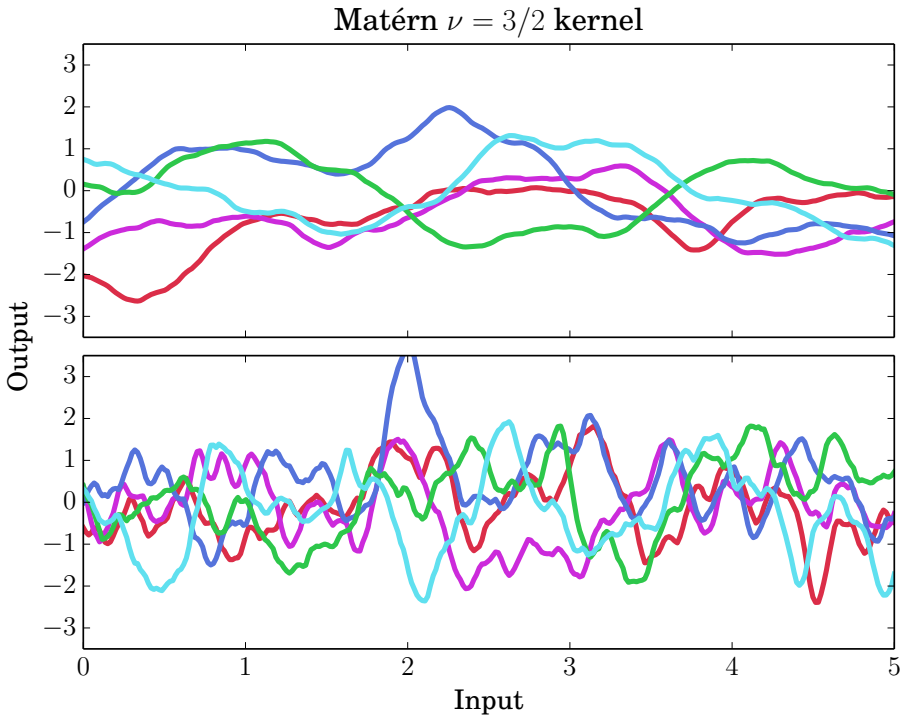

Figure 2.4: (Top) The same as the top panel of Figure 2.3, but for a Matérn $\nu = 3/2$ kernel, defined according to Equation 2.29. (Bottom) The same as the top panel, but with a covariance length scale L that is five times smaller.

2.27–2.29 with:

$$\sqrt{\sum_{k=1}^{D}\left(\frac{r_{ij,k}}{L_k}\right)^2} \;=\; \sqrt{\sum_{k=1}^{D}\left(\frac{v_{i,k}-v_{j,k}}{L_k}\right)^2}\,, \tag{2.30}$$

where $r_{ij,k}$ is the Euclidean distance between the ith and jth data points for the kth input variable, and L_k is the covariance length scale associated with the variable. This allows the model to cope better with heterogeneous systematics that are a combination of different processes. For example, in Chapter 4 we will see that the *HST*/STIS systematics are correlated with the orbital phase of the satellite and the orientation of the spectral trace on the detector. Furthermore, assigning separate length scales to each input can help us identify those that are not strongly correlated with the measured signal, and can thus be discarded. In practice, such inputs will be identified during the inference process (Section 2.2) as those for which $L_k \to \infty$.

2.3.2 Inbuilt Occam's razor

GP models have the desirable property of automatically implementing the principle of Occam's razor. We can see this by considering the role played by the first two terms on the righthand side of Equation 2.26 (the third term is a fixed constant).

The first term, $-\frac{1}{2}\boldsymbol{r}^{\mathrm{T}}(\mathbf{K}+\boldsymbol{\Sigma})^{-1}\boldsymbol{r}$, serves as a goodness-of-fit term. For a given covariance matrix, it increases as the residuals become smaller, rewarding mean functions $\boldsymbol{\mu}$ that match the data well.

The second term, $-\frac{1}{2}\ln|\mathbf{K}+\boldsymbol{\Sigma}|$, can be thought of as a complexity penalty. As it decreases, the likelihood function becomes less sharply peaked near the mean $\boldsymbol{\mu}$ and the probability mass of the model becomes more diffusely spread throughout function space. The model thus becomes more complex in the sense that similar probabilities are attributed to a greater diversity of functions.

An important point to make is that, unlike the covariance, complexity built into the mean function $\boldsymbol{\mu}$ is not penalised. Generally speaking, flexibility built into $\boldsymbol{\mu}$ will tend to allow a closer match to the data, and thus be rewarded by the goodness-of-fit term. It is good practice then to use the mean function for modelling only those elements of the signal that have a well-understood functional form.

2.3.3 Probability distribution over function space

For the lightcurve fitting in Chapters 3, 4, and 5, the transit signal is set as the mean function, and the covariance kernel is used to account for residual sources of correlated noise. Such a model is illustrated in Figure 2.5, which shows a synthetic lightcurve with correlated noise added in and overplotted with the best-fit GP model using a Matérn $\nu = 3/2$ covariance kernel. The best-fit model here actually refers to a distribution over function space, corresponding to the multivariate normal distribution of the GP model conditioned on the observed data. Thus, the probability of measuring some value $d_\star$ at input location $\boldsymbol{v}_\star$ having seen the data $\boldsymbol{d}$ is given by:

$$p(d_\star \mid \boldsymbol{d}) \;=\; \mathcal{N}\left(\eta_\star, c_\star^2 + \sigma_\star^2\right) , \tag{2.31}$$

with:

$$\eta_\star \;=\; \mu_\star + \boldsymbol{k}_\star \left(\mathbf{K} + \boldsymbol{\Sigma}\right)^{-1} \boldsymbol{r} , \tag{2.32}$$

$$c_\star^2 \;=\; k(\boldsymbol{v}_\star, \boldsymbol{v}_\star) - \boldsymbol{k}_\star^{\mathrm{T}} \left(\mathbf{K} + \boldsymbol{\Sigma}\right)^{-1} \boldsymbol{k}_\star , \tag{2.33}$$

where the vector $\boldsymbol{k}_\star$ has elements $k(\boldsymbol{v}_i, \boldsymbol{v}_\star)$ for $i = 1, \ldots, N$. In the context of lightcurve fitting, the predictive distribution described by Equations 2.31–2.33 is merely convenient for plotting our best-fit model, by evaluating it at the input locations of the measured data (Figure 2.5). However, in other applications where interpolation and prediction are necessary, the GP predictive distribution

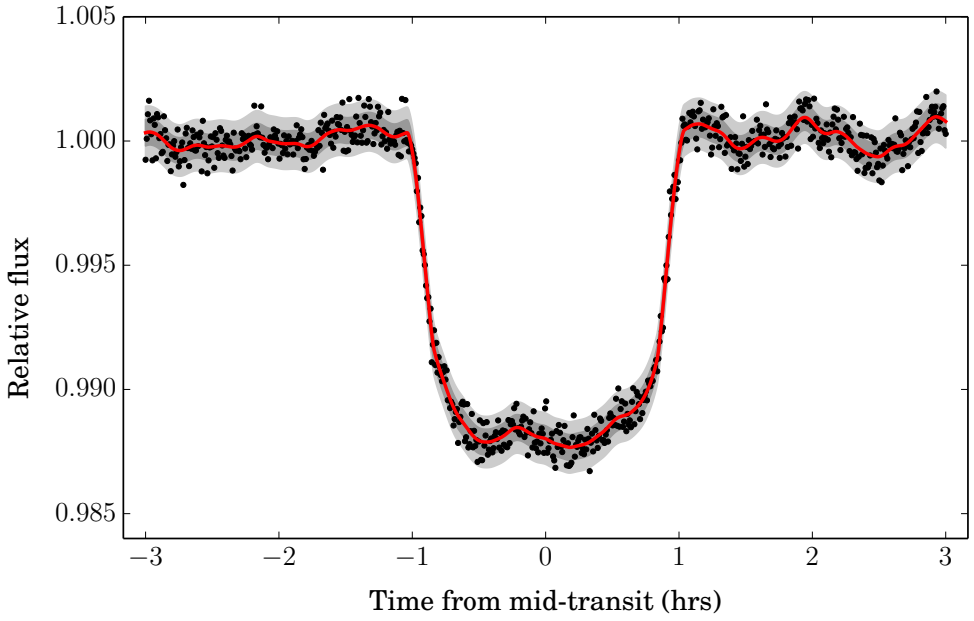

Figure 2.5: An example of a synthetic transit lightcurve with correlated noise and white noise added. The red line shows the predictive mean (Equation 2.32) of a GP model with a transit lightcurve mean function and Matérn $\nu = 3/2$ covariance kernel (Equation 2.29). Dark and light grey shading indicate the 1σ and 2σ ranges of the predictive distribution (Equation 2.33).

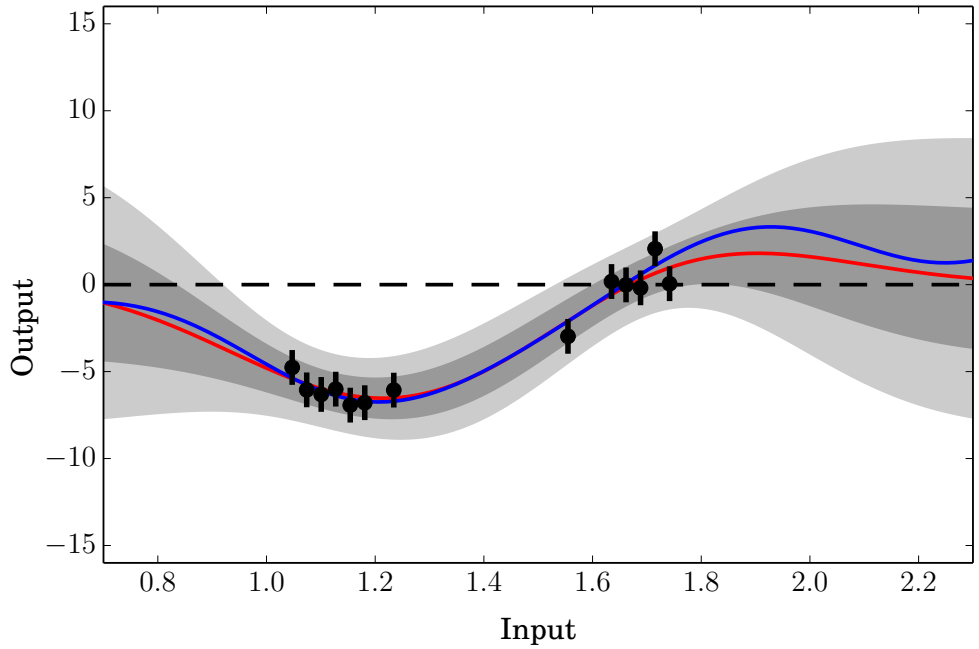

Figure 2.6: An example GP predictive distribution, with mean (solid red line), 1σ range (dark grey shading), and 2σ range (light grey shading) indicated. The GP has been conditioned on observed data (black points) that were generated from some underlying process (solid blue line). A squared exponential kernel was used for the covariance. Notice how the uncertainty of the predictive distribution increases as the distance from the observed data increases. Moving away from the observed data, the predictive distribution reverts to the mean distribution unconditioned on the data, which in this case is a zero-value function (dashed black line).

provides a principled means of propagating uncertainty. As illustrated in Figure 2.6, the variance of the predictive distribution exhibits the sensible behaviour of shrinking in regions of input space that are well-constrained by data, and reverting to the prior distribution (Equation 2.25) away from the measured data.

2.3.4 Computational complexity

Once we have defined the mean function and covariance kernel for our GP model, the next task is to perform inference, i.e. optimise the likelihood with respect to the unknown parameters or marginalise over the parameter space (Section 2.2). This process typically involves multiple evaluations of Equation 2.26, each of which requires inverting the $N \times N$ matrix $\mathbf{K} + \mathbf{\Sigma}$. The computational cost of this operation scales as $\mathcal{O}\left(N^3\right)$, which in practice makes GP models impractical for datasets where $N \gtrsim 1000$.

2.4 The `pyhm` and `gps` software packages

I wrote two custom software packages in the Python programming language based on the material presented in this chapter, and used these to perform the lightcurve analyses described in Chapters 3, 4, and 5. The first is the `pyhm` package, which allows users to define Bayesian models for arbitrary datasets. It includes routines for optimising the model likelihood with respect to the unknown parameters and for performing MCMC with the Metropolis-Hastings algorithm. The second is the `gps` package, which allows GP objects to be defined with user-specified mean functions and covariance kernels. Class methods include evaluating the GP log likelihood (Equation 2.26) and computing predictive distributions (Equation 2.31). Both packages have been made available online.[3]

[3] `http://www.github.com/tomevans`

Chapter 3

Gaussian processes applied to *Spitzer*/IRAC lightcurves

Over the past decade, the *Spitzer Space Telescope* has proven to be a highly productive facility for characterising the atmospheres of exoplanets. The ability of its instruments to probe the ~3–25μm wavelength range has provided constraints on the thermal emission from many exoplanets, as well as atmospheric transmission in a region dominated by absorption from molecular species such as H_2O, CH_4, CO, and CO_2. Given their large atmospheric scale heights and favourable contrast ratios at these wavelengths, hot Jupiters have offered especially favourable targets for such observations.

In this chapter, I consider transiting exoplanet lightcurves that have been acquired using one *Spitzer* instrument in particular: the Infrared Array Camera (IRAC). A selection of significant IRAC observations in the field of exoplanet atmosphere characterisation were mentioned in Section 1.4.1. The success of these observations has depended crucially upon instrumental systematic corrections, which are particularly strong in IRAC lightcurves. The contribution of the present work is to apply the machinery of Gaussian processes (GPs), which were introduced in Section 2.3, to the task of robust lightcurve modelling. To focus the analysis, I restrict attention to transits and eclipses that have been observed for the canonical hot Jupiters HD 189733b and HD 209458b.

The chapter is arranged as follows. Brief descriptions of the *Spitzer* observatory and IRAC instrument are given in Section 3.1. In Section 3.2, I place HD 189733b and HD 209458b in context by surveying the important transmission and emission data that has been published for these planets to date. In Section 3.3, I provide details of the IRAC datasets that have been analysed for the current analysis. My custom-written pipeline for producing lightcurves from raw IRAC data frames, `pyirac`, is detailed in Section 3.4. In Section 3.6, I describe the instrumental systematics that are characteristic of IRAC lightcurves, and the methods that have previously been used to treat them. In Section 3.7, I describe the GP framework that I employ for lightcurve modelling, and discuss the relative advantages and disadvantages that it has compared to other approaches. In Section 3.8, I present the results of my lightcurve analyses for HD 189733b and HD 209458b. I discuss the implications for the planetary atmospheres in Section 3.9, before concluding in Section 3.10.

3.1 *Spitzer Space Telescope*

Launched into a heliocentric Earth-trailing orbit in August 2003, the NASA *Spitzer Space Telescope* started taking scientific data in December 2003. With a mirror diameter of 85cm, *Spitzer* is equipped with three scientific instruments: the Infrared Array Camera (IRAC) for imaging in four wavelength channels centered at 3.6μm, 4.5μm, 5.8μm, and 8.0μm; the Infrared Spectrograph (IRS) for low-resolution single-slit spectroscopy between 5.2–38.0μm and high-resolution echelle spectroscopy between 9.9–37.2μm; and the Multiband Imaging Photometer for *Spitzer* (MIPS) for imaging in three wavelength channels centered at 24μm, 70μm, and 160μm.

In May 2009 *Spitzer* exhausted its helium coolant, and from July 2009 entered a *Warm Mission* phase during which only the 3.6μm and 4.5μm IRAC

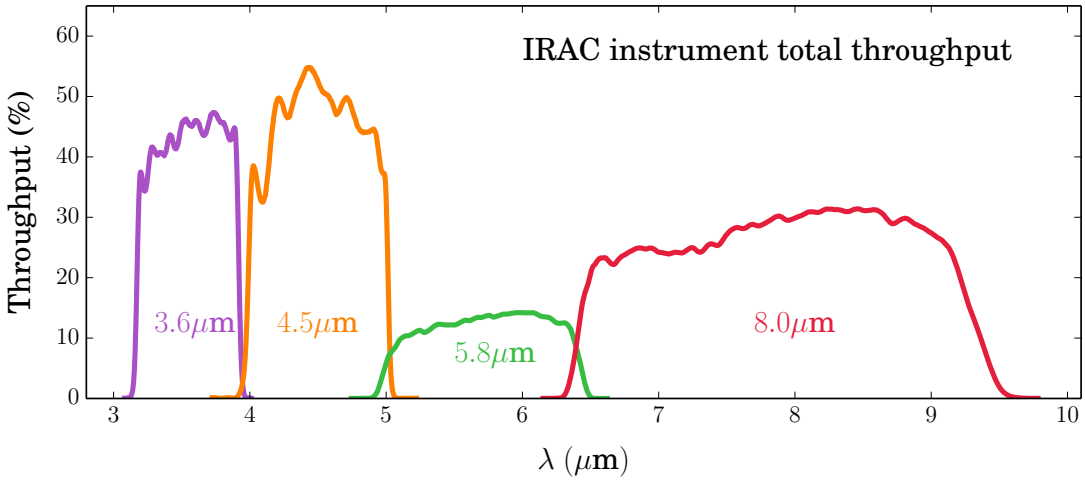

Figure 3.1: Throughputs of the IRAC 3.6μm, 4.5μm, 5.8μm, and 8.0μm channels.

channels were operational. Both of these channels continue to perform at a level comparable to the cryogenic phase of the mission. The *Warm Mission* is currently funded through to September 2014, with funding beyond this time subject to review.

3.1.1 Infrared Array Camera (IRAC)

Light that enters IRAC is first divided into two adjacent, non-overlapping 5.2′ × 5.2′ fields of view (FOVs) by a pair of pickoff mirrors. Both light beams are refocused by doublet lenses onto dichroic beam splitters, which reflect the shorter wavelengths and transmit the longer wavelengths. The four resulting light beams then pass through filters that isolate the different bandpasses shown in Figure 3.1, before arriving at the detectors. With this optical setup, the first FOV is imaged in the 3.6μm and 5.8μm channels and the second FOV is imaged in the 4.5μm and 8.0μm channels. There is, however, a small shift of ~10 pixels between the FOVs of the 3.6μm and 5.8μm wavelength channels, and also for the 4.5μm and 8.0μm wavelength channels.

All four detector arrays are 256 × 256 pixels. The 3.6μm and 4.5μm channel detectors are made of indium antimonide (InSb), and the 5.8μm and 8.0μm

channels are made of arsenic-doped silicon (Si:As). Fowler sampling is used to read data frames, with full array and subarray modes available. For the subarray mode, sets of 64 frames covering 32×32 pixels ($38'' \times 38''$) are taken in succession, with a readout time of 0.01 sec per frame compared with 0.2 sec for the full 256×256 pixel array. Frame times[1] of 0.02 sec, 0.1 sec, and 0.4 sec are available in subarray mode,[2] while full array mode allows frame times of 0.4 sec in the 3.6μm and 4.5μm channels and 2 sec in the 5.8μm and 8.0μm channels.[3] Although full array mode was used for some of the earliest exoplanet observations, subarray mode was soon adopted as the standard due to its higher duty cycle and suitability for bright targets. Further details of the IRAC instrument design and operating modes can be found in Fazio (2004), Reach et al. (2005), and the instrument handbook.[4]

3.2 Review of published observations

Before proceeding to describe the details of the archival IRAC datasets that have been analysed for the current study, I pause here to review the transmission and emission observations that have made the greatest contributions to our understanding of the atmospheres of HD 189733b and HD 209458b. The data are shown in Figures 3.2 and 3.3, respectively. Indeed, most of the IRAC lightcurves that I have re-analysed are already published, and I summarise the results of those studies below. In doing so, I aim to place the IRAC datasets in an overall context, and highlight those that have produced anomalous or conflicting results.

[1] Frame time is defined as the total time taken to reset the array, make the pre-reads, integrate, and make the post-reads. Hence, $t_f = t_e + n_F t_r + t_d$ if t_f is the frame time, t_e is the exposure time, n_F is the Fowler number, t_r is the readout time, and t_d is the detector reset time. Typical values for n_F are 1–4, and t_r is 0.2s for full array mode and 0.01s for subarray mode. The detector reset time t_d is negligible.

[2] Due to high data rates, only a single channel can be used at a time for the 0.02s frame time, whereas data for all four channels are read out for the 0.1s and 0.4s frame times.

[3] Frame times of 12s, 30s, and 100s can also be employed simultaneously by all four channels in full array mode, but these are never used for bright exoplanet host stars.

[4] Available at `http://irsa.ipac.caltech.edu/data/SPITZER/docs/irac`

This will set the scene for introducing the GP modelling approach in Section 3.7, which I argue avoids a number of fundamental problems that have been inherent to previous lightcurve analyses.

3.2.1 HD 189733b

The first claimed detection of molecular absorption in an exoplanet atmosphere was made for HD 189733b by Tinetti et al. (2007), who inferred the presence of H_2O based on transit measurements made in the IRAC 3.6μm, 5.8μm, and 8.0μm wavelength channels. Key to this interpretation was the measurement of a significantly higher opacities in the 5.8μm and 8.0μm channels relative to the 3.6μm channel, based on the lightcurve analyses that were published in Knutson et al. (2007a) and subsequently in Beaulieu et al. (2008). The 3.6μm and 5.8μm lightcurves were re-analysed by Désert et al. (2009), who were unable to replicate the results, finding no significant variation in the opacity between the two channels. In addition, Désert et al. analysed new lightcurves in the 4.5μm and 8.0μm channels, and found an enhanced opacity in the 4.5μm channel relative to the others. This feature was tentatively attributed to CO absorption, and has subsequently been confirmed with further IRAC observations (Knutson et al., 2012).

Swain et al. (2008b) used *HST*/NICMOS in spectroscopic mode to observe a transit of HD 189733b over the wavelength range 1.5–2.5μm, and reported the detection of absorption features due to CH_4 and H_2O. However, using NICMOS to observe a transit in photometric mode over two narrow bandpasses centered on 1.66μm and 1.87μm, Sing et al. (2009) were able to rule out significant absorption by H_2O. Furthermore, the robustness of observations made in NICMOS spectroscopic mode has been called into question by Gibson et al. (2011), who demonstrated that the signal proposed by Swain et al. could easily be caused by

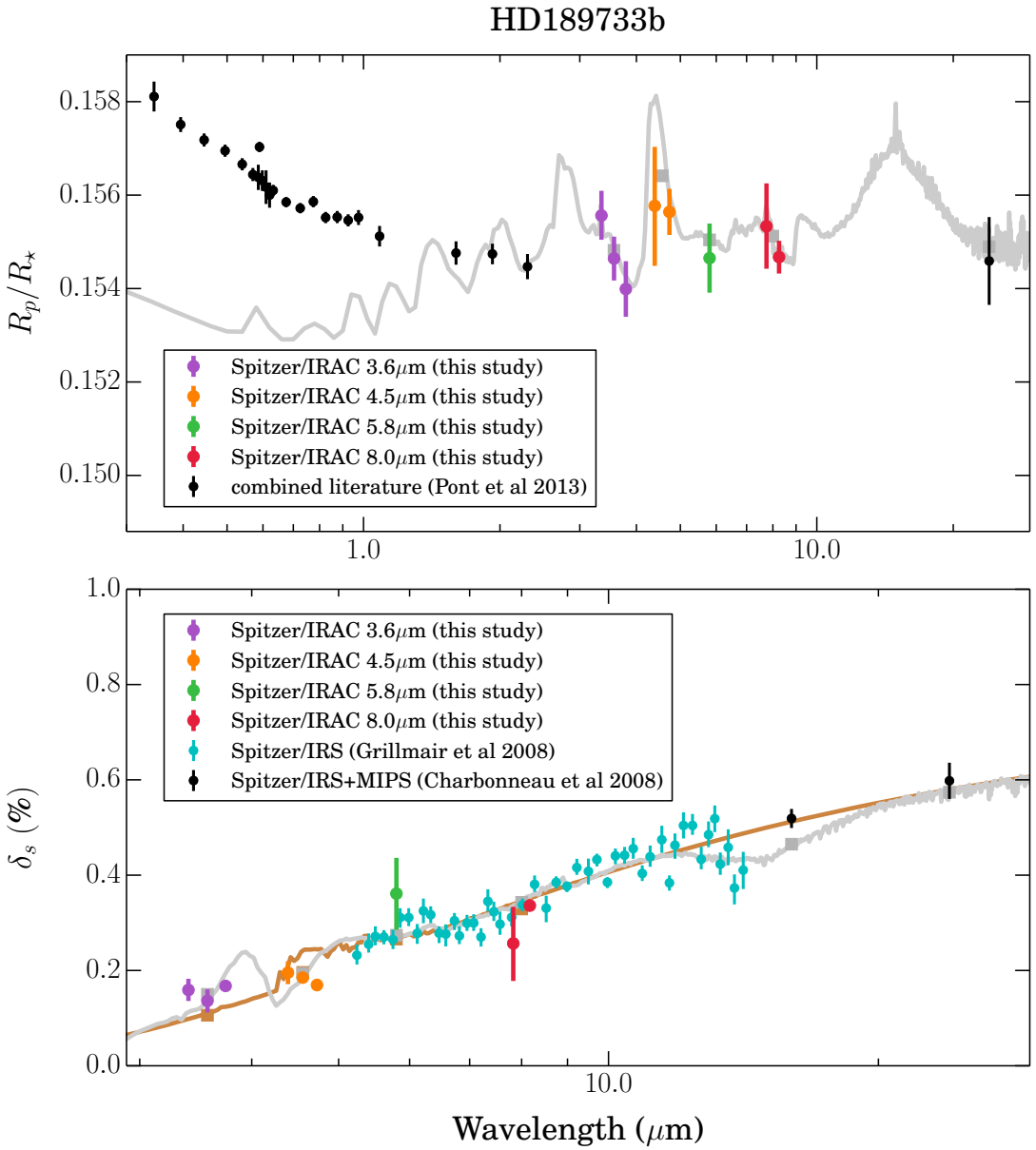

Figure 3.2: Transmission $R_p/R_\star$ (top) and emission δ_s (bottom) data for HD 189733b. The IRAC data shown are those values obtained in the current study, with small horizontal offsets applied to measurements made in the same channel at different epochs for clarity. Data at other wavelengths have been taken from the published literature. All transmission values have been corrected for the effect of variation in star spot coverage between observation epochs, as described in Section 3.8.1. The transmission data from Pont et al. (2013) were adapted from the following original publications: Gibson et al. (2012b,a), Huitson et al. (2012), Knutson et al. (2009b), Pont et al. (2008a), and Sing et al. (2009, 2011b). Solid grey lines show NEMESIS models for clear atmospheres, as described in Section 3.9.1. The enhanced transmission opacity measured at wavelengths $< 2\,\mu$m is thought to be the Rayleigh scattering signature of a high-altitude haze of sub-micron dust grains. For emission, the solid brown shows the signal expected if the planet radiated as an isothermal blackbody with a temperature of 1212 K, assuming an ATLAS spectrum for the star (see Section 3.8.1). Square symbols give model values integrated over photometric bandpasses. Both horizontal axes are on a log scale.

an imperfect correction of the NICMOS instrumental systematics. In a follow-up paper, Gibson et al. (2012b) presented a new transmission spectrum extracted from the original NICMOS spectroscopy using GPs to treat the instrumental systematics. Although the shape of the Gibson et al. spectrum is qualitatively similar to the one obtained by Swain et al., the more comprehensive treatment of measurement uncertainties causes the errorbars to be inflated by a factor of ~2–3, greatly reducing the detection significance of any absorption features that may be present in the data. Indeed, Gibson et al. are quite explicit in acknowledging that even their GP analysis is likely to have underestimated the true uncertainties, owing to the poorly understood nature of the NICMOS systematics. For instance, this could occur if the systematics are correlated with an auxiliary variable that is not measured and hence cannot be fed as an input to the GP model. The NICMOS spectroscopy results are therefore treated with skepticism by a significant fraction of the exopolanet community.

At shorter wavelengths, Pont et al. (2008a) used the *HST* Advanced Camera for Surveys (ACS) to measure the transmission spectrum over the 0.55–1.05μm wavelength range. Expecting to detect the strong absorption features due to sodium and potassium that are predicted by models, Pont et al. instead measured a near-featureless spectrum, with the atmosphere opacity observed to decrease smoothly with increasing wavelength. This has been interpreted as the signature of Rayleigh scattering by a high altitude haze of sub-μm haze particles, possibly composed of a refractory condensate such as $MgSiO_3$ (enstatite) (e.g. Lecavelier Des Etangs et al., 2008). Observations made by Sing et al. (2011b) using *HST*/STIS have since confirmed that the λ^{-4} scattering profile extends across the 290–570 nm wavelength range. Also observing with STIS, Huitson et al. (2012) detected the narrow core of the Na I doublet at 589 nm, and did not detect the absorption wings. This was interpreted as evidence either of an obscuring high

altitude haze, or a depletion of Na in the atmosphere. Meanwhile, evidence that haze remains a significant opacity source in the near-infrared has come from Gibson et al. (2012a), who used *HST*/WFC3 to measure transits in narrow photometric bands centered at 1.1μm and 1.5μm, and found that the effective planetary radius remained higher than predictions for clear atmosphere models at these wavelengths.

In emission, Deming et al. (2006) measured the first secondary eclipse for HD 189733b at 16μm with *Spitzer*/IRS. The same instrument was used in spectroscopic mode by Grillmair et al. (2007) to observe an eclipse over the 8–15μm wavelength range. This observation produced a flat spectrum, lacking the absorption features due to H_2O or CH_4 predicted by models. However, an additional measurement made by the same group, again with *Spitzer*/IRS, gave an emission spectrum consistent with strong H_2O absorption (Grillmair et al., 2008). A convincing explanation for the disagreement between these two observations has not been provided. Rather, the authors speculated that planet-wide weather variations between the two epochs may be responsible.

Charbonneau et al. (2008) observed a secondary eclipse while continuously cycling between the four IRAC channels (see Section 3.3), and observed another eclipse at 24μm with *Spitzer*/MIPS. Notably, the authors found the emission from the planet to be significantly higher in the 3.6μm channel relative to the other three channels. Charbonneau et al. suggested the suppressed emission in the longer wavelength channels was due to H_2O absorption, with the peak at 3.6μm corresponding to an opacity window.

An additional six transits and six eclipses have been measured by Agol et al. (2010) in the 8.0μm channel. Notably, Agol et al. found significant variations in the measured transit depths across epochs, suggesting that stellar activity is likely to bias the results even for wavelengths as long as 8.0μm. On the other

hand, they concluded that the measured eclipse depths were consistent with a constant value at the level of the uncertainties. This is unsurprising, as the measured eclipse depth should be relatively unaffected by variations in star spot coverage.

Swain et al. (2009b) have presented an emission spectrum for HD 189733b measured across the 1.5–2.5μm wavelength range with *HST*/NICMOS. According to the authors, absorption by H_2O, CO_2, and CO is required to explain the observed spectrum. However, this result comes with the same caveats as the other NICMOS observations described above. A re-analysis of the data using GPs has been presented by Barstow et al. (2014), who found uncertainties for the inferred eclipse depths that were a factor of ~ 3 larger than those originally reported by Swain et al.

Using *Spitzer*, thermal phase curves have also been measured in the 8.0μm IRAC channel (Knutson et al., 2007a), with MIPS at 24μm (Knutson et al., 2009b), and in the 3.6μm and 4.5μm IRAC channels (Knutson et al., 2012). All of these lightcurves have revealed that the hottest part of the dayside hemisphere is shifted approximately 15–30 deg eastward of the substellar point, indicating the advection of heat around the planet in super-rotating equatorial jets as predicted by circulation models (e.g. Showman & Guillot, 2002; Showman et al., 2009; Perna et al., 2012). Examining the multi-wavelength data together, Knutson et al. (2012) found a reasonable agreement with the 1D radiative transfer models of Burrows et al. (2007) and the 3D general circulation models of Showman et al. (2009). However, a compelling match between the data and models apparently requires the treatment of effects such as condensate species, non-equilibrium chemistry, non-synchronous planetary rotation, and non-solar metallicity. As the quality and wavelength coverage of exoplanet data continues to improve, the development of comprehensive models that address these issues will become

increasingly pressing.

3.2.2 HD 209458b

Observing four transits of HD 209458b with *HST*/STIS, Charbonneau et al. (2002) were the first to detect absorption by the Na I doublet in the atmosphere – indeed, this milestone observation was the first detection of any chemical species in an exoplanet atmosphere (Section 1.4.1). Charbonneau et al. noted that the measured absorption was significantly weaker than would be expected for a clear atmosphere with solar Na abundance. One obvious explanation for this would be that the atmosphere of HD 209458b is depleted in Na relative to solar. Alternatively, high-altitude haze could be masking the absorption feature, reducing the amplitude of the core and dampening the wings. Sing et al. (2008a,b), however, were subsequently able to resolve the narrow core of the Na doublet with further *HST*/STIS observations, which appeared to disfavour a hazy atmosphere scenario.

More recently, Deming et al. (2013) have used *HST*/WFC3 to measure the transmission spectrum across the 1.0–1.7μm wavelength range. This observation resulted in a clear detection of the H_2O absorption bandhead centered at 1.4μm, albeit with a lower amplitude than predicted for clear atmospheres. Deming et al. found a more satisfactory match to the data by including an additional grey absorber in their model, as a crude attempt to simulate the effect of an obscuring haze in the atmosphere.

A transmission spectrum has also been acquired over the 0.3–1μm wavelength range by Knutson et al. (2007b) using *HST*/STIS. It revealed variations in the effective planet radius as a function of wavelength, but Knutson et al. were reluctant to attribute these to specific absorption features, preferring to defer this work to a future analysis. Barman (2007), however, argued that the measured

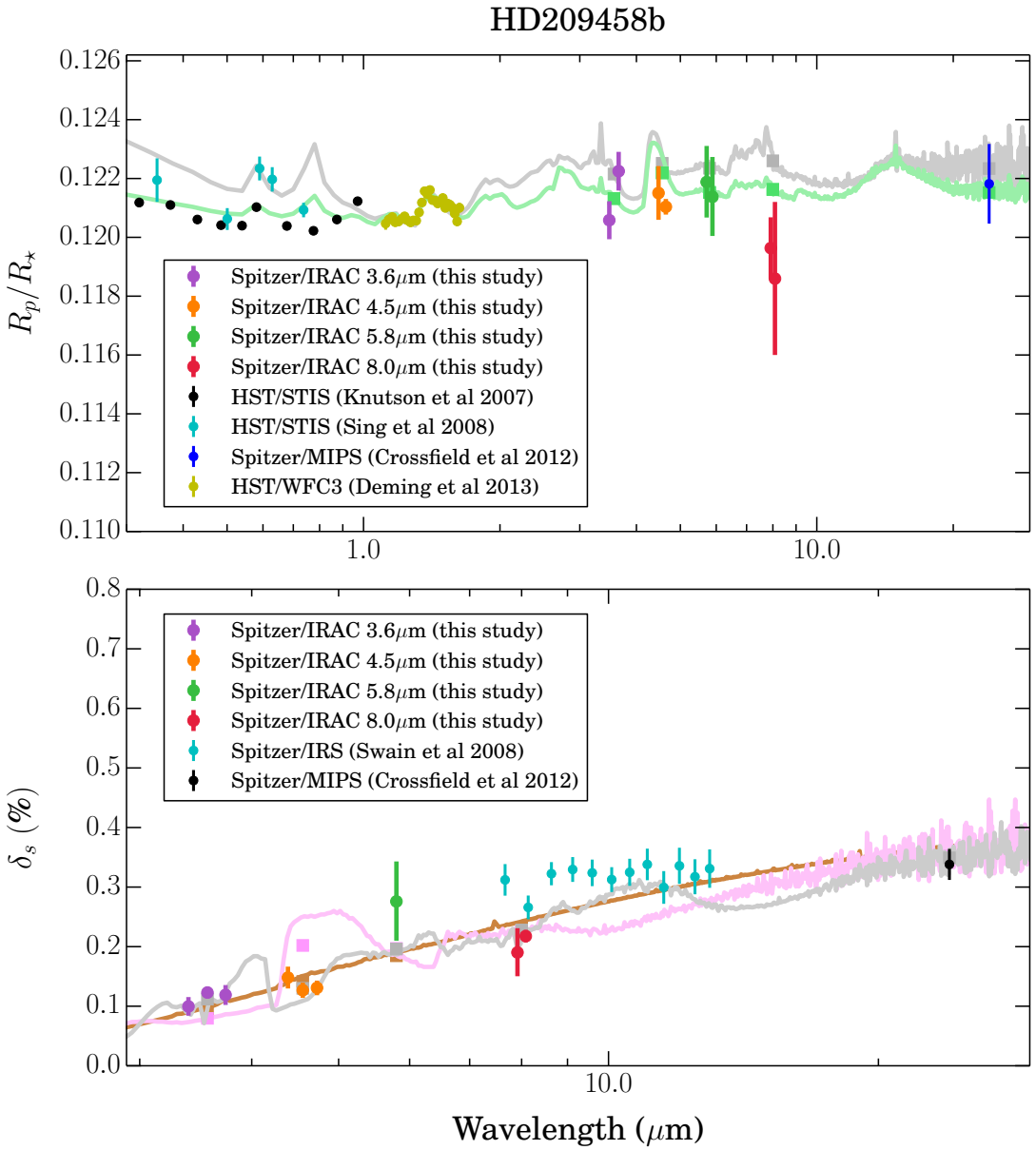

Figure 3.3: Similar to Figure 3.3, but for HD 209458b. In both panels, solid grey lines show NEMESIS models with non-inverted pressure-temperature profiles (see Figure 3.14) and molecular abundances that have been adjusted to give the best fit to the data, as described in Section 3.9.2. In the top panel, the green line shows a NEMESIS model with abundances fixed to those retrieved for the dayside hemisphere, but then rescaled to match the amplitude of the 1.4 μm H_2O feature measured by the WFC3 data. In the bottom panel, the pink line shows a NEMESIS model with an inverted pressure temperature profile (see Figure 3.14) and molecular abundances that have been adjusted to give the best fit to the IRAC data of Knutson et al. (2008), which are not shown. This emphasises how the revised 4.5 μm eclipse depths presented in this study affect the interpretation of the emission spectrum, which is best explained by an atmosphere without a temperature inversion (Section 3.9.2). The solid brown line shows the emission signal expected if the planet radiated as an isothermal blackbody with a temperature of 1460 K, assuming an ATLAS spectrum for the star (Section 3.8.2). Square symbols give model values integrated over photometric bandpasses. Both horizontal axes are on a log scale.

variations were consistent with being caused by H_2O absorption in a cloud-free atmosphere.

Transits of HD 209458b have been observed in all four IRAC channels, and published by Beaulieu et al. (2010). The effective planetary radius was measured by those authors to be significantly higher in the 5.8μm and 8.0μm channels relative to the 3.6μm and 4.5μm channels, which was interpreted as evidence for H_2O absorption. However, Deming et al. (2013) have noted that the degree of H_2O absorption required to explain these IRAC measurements is incompatible with their WFC3 measurement.

Deming et al. (2005) published the first detection of thermal emission from an exoplanet, measuring the secondary eclipse of HD 209458b with *Spitzer*/MIPS at 24μm. Knutson et al. (2008) reported secondary eclipses in each of the four IRAC channels. In particular, Knutson et al. found that the planetary emission in the 4.5μm and 5.8μm channels significantly exceeded that observed for the 3.6μm channel, indicative of H_2O in emission rather than absorption. This could naturally be explained if the atmosphere has a temperature inversion (e.g. Burrows et al., 2007).

An additional two primary transits and two secondary eclipses observed with *Spitzer*/MIPS at 24μm have been reported by Crossfield et al. (2012b). The authors also re-examined the original MIPS eclipse presented in Deming et al. (2005). They found the transmission at 24μm to be consistent with an atmosphere exhibiting absorption features with similar amplitude to the H_2O feature measured by Deming et al. (2013), rather than that claimed by Beaulieu et al. (2010). Meanwhile, the emission was revised upwards slightly relative to the original value proposed by Deming et al. (2005), bringing it more in line with the general circulation model predictions of Showman et al. (2009) for an atmosphere with a temperature inversion.

An emission spectrum has also been obtained by Richardson et al. (2007) across the 8–15μm wavelength range using *Spitzer*/IRS. The authors tentatively attributed an emission excess near 10μm as being the signature of H_2O absorption. However, an independent analysis of the same data by Swain et al. (2008a) failed to replicate this result, instead revealing an essentially featureless spectrum.

Zellem et al. (2014) have published a full phase observation taken in the IRAC 4.5μm channel. They inferred a lower planetary emission in this wavelength channel compared to that found by Knutson et al. (2008), giving better agreement with predictions of general circulation models (e.g. Showman et al., 2009). As with HD 189733b, the dayside hemisphere hot spot was observed to be shifted eastward of the substellar point by ~ 40 deg, indicating advection by winds.

3.3 Archival IRAC datasets analysed for this study

For this study, I have re-analysed almost all of the IRAC lightcurves that were mentioned in the previous section.[5] In total, this comprises eight primary transit and nine secondary eclipse lightcurves for HD 189733b, and nine primary transit and nine secondary eclipse lightcurves for HD 209458b. Details of the relevant *Spitzer* observing programs are given in Table 3.1, and more specific information for the individual lightcurves is given in Table 3.2.

For the eclipse observations presented in Charbonneau et al. (2008) for HD 189733 and Knutson et al. (2008) for HD 209458, which were taken as part of Programs 261 and 20523, respectively, the following observing strategy was employed. Using the subarray readout mode, four sets of 64 frames were acquired in a given channel before the telescope was repointed to be centered on the next channel, and another four sets of 64 frames were taken. This process was cycled through each of the four channels, and repeated for the duration of

[5]The only omissions are the six transits and six eclipses for HD 189733b in the 8.0μm channel published by Agol et al. (2010), which were observed as part of Program 40238 (PI: Agol).

Table 3.1: Datasets analysed for this study.

Target	Program	PI [a]	Type [b]	Channels (μm)				References [c]
				3.6	4.5	5.8	8.0	
HD 189733b	261	D. Charbonneau	Eclipse	Yes	Yes	Yes	Yes	Ch08
	30825	D. Charbonneau	Half-phase	-	-	-	Yes	Kn07 Ag10
	30590	A. Vidal-Madjar	Transit	Yes	-	Yes	-	Be08 Dé09 Dé11
	40732	A. Vidal-Madjar	Transit	Yes	Yes	-	Yes	Dé09 Dé11
	60021	H. Knutson	Full phase	Yes	Yes	-	-	Kn12
HD 209458b	20523	D. Charbonneau	Eclipse	Yes	Yes	Yes	Yes	Kn08
	40280	H. Knutson	Half-phase	-	-	-	Yes	-
	461	G. Tinetti	Transit	Yes	Yes	Yes	Yes	Be10
	60021	H. Knutson	Full phase	Yes	Yes	-	-	Ze14

Notes:

[a] Principal investigator of the observing program.

[b] The type of observation performed. The full-phase observations for HD 189733b and HD 209458b made as part of Program 60021 cover the planets' entire orbital periods, but in this study I only analyse the transit and two eclipses. Similarly, I only anlayse a transit and eclipse for the half-phase observations made as part of Programs 30825 and 40280.

[c] Ag10 (Agol et al., 2010), Ba10 (Ballard et al., 2010), Be08 (Beaulieu et al., 2008), Be10 (Beaulieu et al., 2010), Ch08 (Charbonneau et al., 2008), Dé09 (Désert et al., 2009), Dé11 (Désert et al., 2011a), Kn07 (Knutson et al., 2007a), Kn08 (Knutson et al., 2008), Kn12 (Knutson et al., 2012), Ze14 (Zellem et al., 2014).

the observations. As with Charbonneau et al. (2008) and Knutson et al. (2008), I discarded the first set of 64 frames from each cycle, as the stellar PSF was still drifting significantly during this time following the repointing. Like those authors, I also discarded the first 10 frames and the 58th frame from each set of 64 for the 5.8μm and 8.0μm channels, as these exhibited count levels consistently below the median.

Following Programs 261 and 20523, it was realised that constantly repointing the telescope was a bad approach for performing exoplanet observations as it exacerbated the effect of the instrumental systematics, which are described below in Section 3.6. Subsequent lightcurves listed in Table 3.2 were therefore acquired using stare mode. A number of these observations were made in full array mode, with data taken simultaneously in a pair of wavelength channels (Programs 30590, 40732, and 461). It has also been common, however, to use subarray mode to observe in a single channel, as this allows for shorter exposure times (Programs 30825, 40280, and 60021). The latter reduces the risk of saturation for bright targets such as HD 189733 and HD 209458, and improves the duty cycle. During the *Warm Mission*, only the 3.6 μm and 4.5 μm channels have been available, and for this reason all observations have been made in subarray mode. The numbers of data frames acquired for each lightcurve are listed in Table 3.2, and range from $\sim 10^3$ for single transits observed in full array mode, to $\sim 10^6$ for full-phase observations observed in subarray mode. Note, however, that for this study I do not model the complete half- and full-phase datasets acquired for Programs 30825, 40280, and 60021. Instead, for these lightcurves only the subsections containing the eclipses and transits are analysed. This translates to a single transit and eclipse for the half-phase observations, and two eclipses and a transit for the full-phase observations.

Table 3.2: Lightcurve details.

Target	Program	Type	Channel (μm)	Date (UT)	Δt [a] (hr)	Mode [b]	N [c]
HD 189733	261	Eclipse	3.6	2005 Nov 24	4.9	sub, 0.1	21760
		Eclipse	4.5	2005 Nov 24	4.9	sub, 0.1	21760
		Eclipse	5.8	2005 Nov 24	4.9	sub, 0.1	21760
		Eclipse	8.0	2005 Nov 24	4.9	sub, 0.1	21760
	30825	Half-phase	8.0	2006 Oct 29–30	10.0	sub, 0.4	84416
	30590	Transit	3.6	2006 Oct 31	4.5	full, 0.4	1936
		Transit	5.8	2006 Oct 31	4.5	full, 2.0	1936
	40732	Transit	3.6	2007 Nov 25	4.5	sub, 0.1	122880
		Transit	4.5	2007 Nov 23	4.5	full, 0.4	1920
		Transit	8.0	2007 Nov 23	4.5	full, 2.0	1920
	60021	Full-phase	3.6	2010 Dec 27–29	68.5	sub, 0.1	1722624
		Full-phase	4.5	2009 Dec 22–24	66.7	sub, 0.1	1711872
HD 209458	20523	Eclipse	3.6	2005 Nov 27–28	8.1	sub, 0.1	35840
		Eclipse	4.5	2005 Nov 27–28	8.1	sub, 0.1	35840
		Eclipse	5.8	2005 Nov 27–28	8.1	sub, 0.1	35840
		Eclipse	8.0	2005 Nov 27–28	8.1	sub, 0.1	35840
	40280	Half-phase	8.0	2007 Dec 23–26	52.3	sub, 0.4	439680
	461	Transit	3.6	2007 Dec 30–31	6.7	full, 0.4	2860
		Transit	3.6	2008 Jul 18–19	6.7	full, 0.4	2860
		Transit	4.5	2008 Jul 22	6.7	full, 0.4	2860
		Transit	5.8	2007 Dec 30–31	6.7	full, 2.0	2860
		Transit	5.8	2008 Jul 18–19	6.7	full, 2.0	2860
		Transit	8.0	2008 Jul 22	6.7	full, 2.0	2860
	60021	Full-phase	3.6	2011 Jan 12–16	102.6	sub, 0.1	2597440
		Full-phase	4.5	2010 Jan 17–21	96.6	sub, 0.4	795200

Notes:

[a] Time elapsed between the start and end of observations.

[b] Readout mode and frame time in seconds. See Section 3.1.1 for further details.

[c] Total number of frames in the dataset.

3.4 The `pyirac` software package

The raw data frames listed in Table 3.2 were reduced and aperture photometry performed using the `pyriac` software package, which was custom-written in the Python programming language for this work and has been made publicly available online.[6] The following sections describe the important steps of the `pyirac` pipeline.

3.4.1 Basic calibrated data

The `pyirac` pipeline takes IRAC Basic Calibrated Data (BCD) frames as input. The BCD frames have had basic pre-processing and calibrations applied, such as dark subtraction and flatfielding. They are provided in Flexible Image Transport System (FITS) format, with headers that contain information about the instrument configuration and observation parameters, and pixel values expressed as fluxes in 32-bit floating point arrays. For frames taken in full array mode the pixel value data is provided as a 256×256 array, while for frames taken in subarray mode it is in the form of a $32 \times 32 \times 64$ cube, where the third dimension corresponds to the frame number.

3.4.2 Centroiding

The first step of the `pyirac` pipeline is to locate the stellar centroid in each BCD frame. Accurate centroiding is especially important for performing IRAC photometry, as the IRAC PSF is undersampled – the FWHM of the PSF varies from ~1.3–1.6 pixels across the four wavelength channels. Consequently, small errors in the centroiding can translate to relatively large errors in the photometry, as fractional pixels make a non-negligible contribution to the total flux. Further information about the IRAC PSF is given in Appendix A.

[6] `http://www.github.com/tomevans`

To perform the centroiding, an approximate guess for the centroid coordinates in the first frame must be provided, accurate to $\sim$ 1–2 pixels. The background value is then estimated and subtracted from the pixel values. Details of the options available for background estimation are given below in Section 3.4.5. Note, however, that the background values computed during this stage are only preliminary estimates made for the purpose of determining the centroids. The final background values are computed after the centroiding and photometry has been completed, at the last step in the pipeline.

A subarray centered on the estimated centroid coordinates and typically measuring $\sim 10 \times 10$ pixels is then extracted from the frame. This subarray is provided as input for either Gaussian profile fitting or calculating the flux-weighted mean, depending on which method the user has specified (details of both these methods are given in Appendix A). By using a subarray for this step, the chance of nearby stars or cosmic ray hits biasing the result is reduced.

If the resulting centroid coordinates differ from those of the previous frame by > 2 pixels, the current frame is flagged as bad. Otherwise, the centroid coordinates of the current frame are used as the initial guess for the next frame, and the process is repeated until centroids have been determined for the entire dataset.

3.4.3 Pre-cleaning

Pre-cleaning involves flagging anomalous frames and excluding them from further analysis. Firstly, the median and standard deviation in the centroid coordinates are computed from the 30 frames that immediately precede and follow each frame. Any frames with centroid coordinates that diverge by $> 5\sigma$ are flagged as bad, and this process is iterated twice. Secondly, the median and standard deviation pixel counts are computed from the 30 frames that immediately precede

and follow each frame. This is done for each individual pixel in a subarray spanning the photometric aperture, the size of which will have been specified by the user. Frames that contain one or more $> 5\sigma$ outlying pixels determined in this manner are flagged as bad. Again, this process is iterated twice. I find that this typically results in $\lesssim 1\%$ of the frames being flagged as bad and discarded.

3.4.4 Aperture photometry

Photometry is performed on each frame by summing up the pixel values contained within a circular aperture of user-specified radius centered on the centroid coordinates. However, due to the undersampled nature of the IRAC PSF, a significant fraction of the pixels are only partially contained within the photometric aperture. There are two basic approaches for handling this situation without resorting to PSF fitting.[7] The first is to geometrically calculate the fractional area of each pixel that is contained within the aperture, and weight its contribution accordingly. The second option, which is the one adopted in the `pyirac` pipeline, is to bilinearly interpolate the pixel values onto a finer grid. This approach is arguably more suited to an undersampled PSF, as it allows for the flux being non-constant across pixels. By default, the `pyirac` pipeline interpolates on to a 10×10 grid per pixel, but this can be adjusted by the user. Grid points that are separated by less than the aperture radius from the centroid are then added towards the total aperture sum for each frame. At this step, flux values are converted from the units of $\mathrm{MJy\,sr^{-1}}$ provided in the BCD frames to photon counts using the `GAIN`, `EXPTIME`, and `FLUXCONV` keywords in the FITS headers.[8]

[7] For information about PSF fitting, see Appendix A.

[8] For further details on these keywords, refer to the *Spitzer* IRAC Instrument Handbook.

3.4.5 Background subtraction

The final step in the `pyirac` pipeline is to estimate the background contributions for each frame and subtract these from the aperture fluxes determined in the previous step. Background pixels are identified using one of three methods, which must be selected by the user. The first option is to use the pixels located within a circular annulus centered on the stellar centroid. The second option is to use pixels located in the corners of the array, in order to avoid the diffraction spikes. The third option is for the user to directly provide a map of the background pixels. Once the background pixels have been identified for a given frame, two iterations of 5σ culling are applied to remove obvious outliers. The per pixel background contribution is then estimated by taking either the mean or median of the remaining pixel values, depending on which has been specified by the user. Conversion from MJy sr^{-1} to photons is made as before using information stored in the FITS header. The per pixel background value is then multiplied by the fractional number of pixels contained within the photometric aperture and subtracted from the corresponding aperture flux. This process is repeated for each frame.

3.5 Data reduction

The `pyirac` pipeline was used to produce lightcurves from BCD frames for the datasets listed Table 3.2. Centroids were computed by taking the flux-weighted mean[9] in a box spanning 7×7 pixels centered on the approximate location of the stellar PSF center. Photometry was performed using aperture radii of 3 pixels for the 3.6μm and 4.5μm frames, and 4 pixels for the 5.8μm and 8.0μm frames. Background levels were calculated by taking the median of the 8×8 pixel

[9] I found the flux-weighted mean method to be more suitable overall for computing IRAC centroids than Gaussian profile fitting. See Appendix A for a detailed comparison of centroiding methods applied to IRAC data.

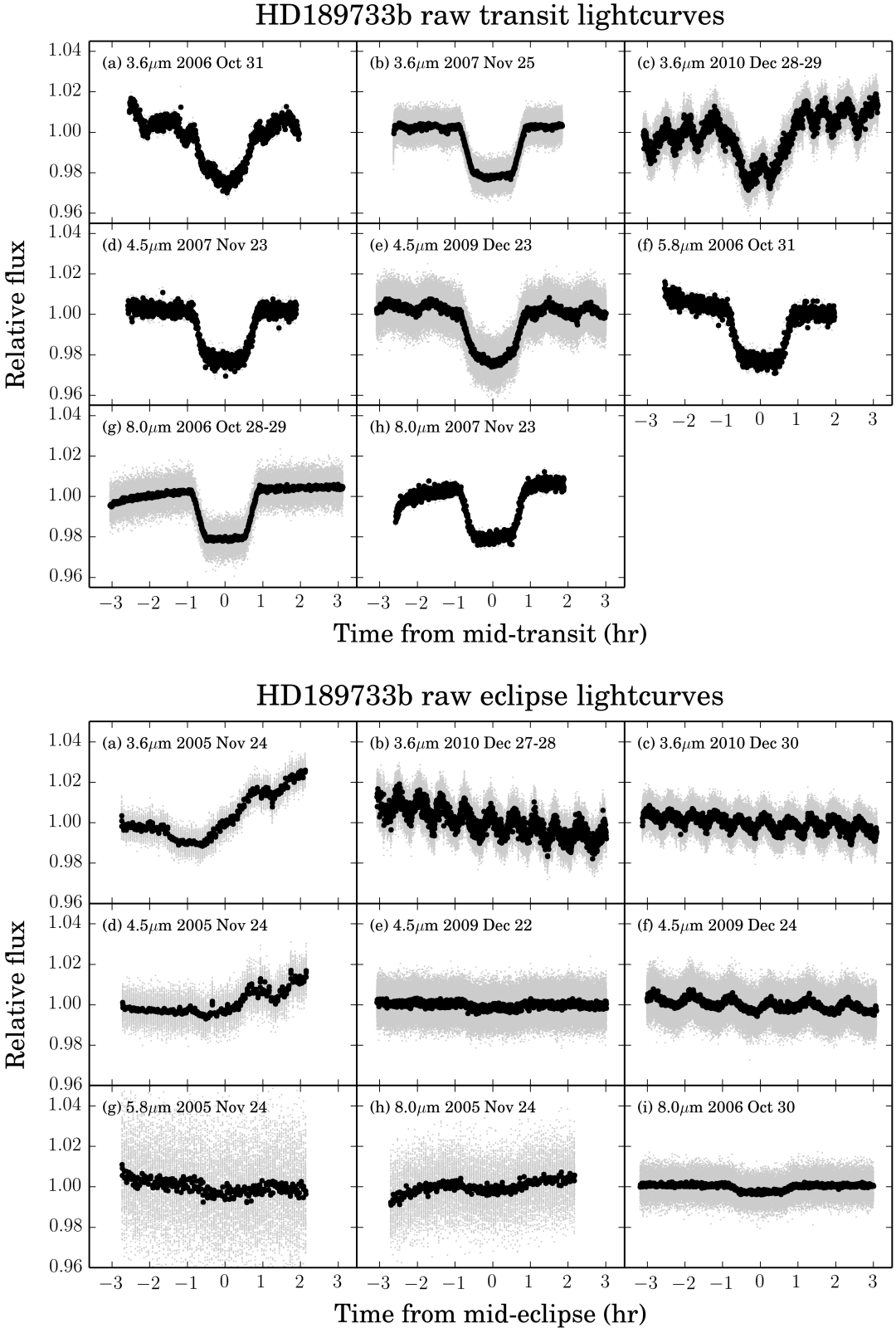

Figure 3.4: Raw transit (top) and eclipse (bottom) lightcurves for HD 189733b. Light grey points show unbinned data and black points show binned data used for the lightcurve analyses. Note that for some of the lower cadence observations (e.g. 3.6 μm 2006 Oct 31, 4.5 μm 2007 Nov 23), the unbinned data are barely visible.

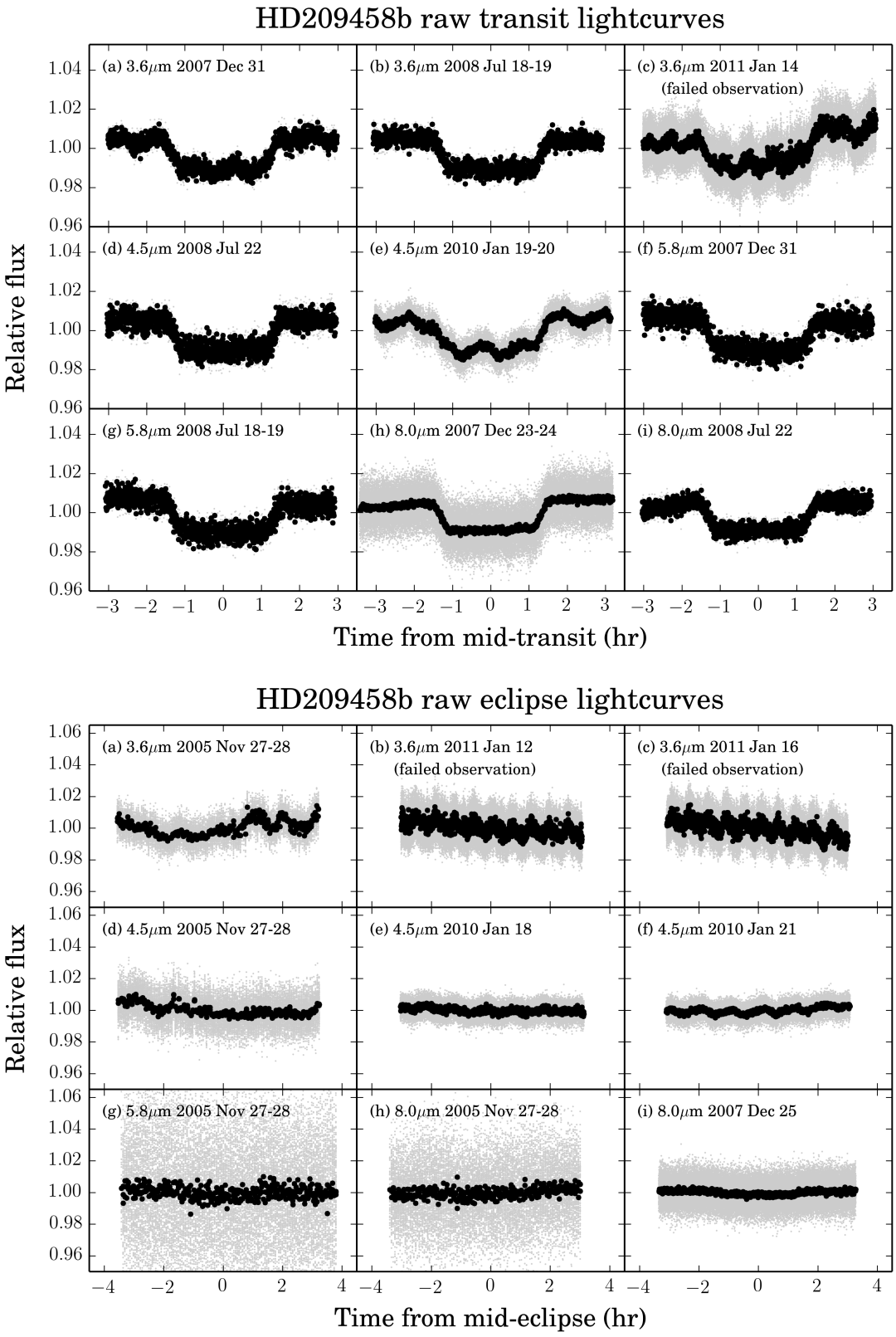

Figure 3.5: Same as Figure 3.4, showing raw IRAC transit (top) and eclipse (bottom) lightcurves for HD 209458b.

subarrays in the frame corners, and then subtracted from the aperture fluxes. The resulting lightcurves are shown in Figures 3.4 and 3.5.

3.6 IRAC lightcurve systematics

Lightcurves acquired with IRAC are affected by well-documented instrumental systematics, which are clearly evident in Figures 3.4 and 3.5. These can be divided into two main types: intrapixel sensitivity variations, which affect the InSb detectors (3.6μm and 4.5μm), and time-dependent sensitivity ramps, which affect the Si:As detectors (5.8μm and 8.0μm). Illustrative examples of the different types of systematics encountered in each channel are highlighted in Figure 3.6 and described further in the following sections, along with the common approaches that have been used to treat them in the literature.

3.6.1 Intrapixel sensitivity variations

In the 3.6μm and 4.5μm channels, the measured flux correlates with the position of the stellar PSF on the detector array (Figure 3.6). As noted in the IRAC instrument handbook, this effect is caused by variations in the quantum efficiency across individual pixels. Pointing drift during observations, combined with the undersampled nature of the PSF, therefore results in variations in the measured flux at the few percent level.

An early approach for correcting this effect was proposed by Reach et al. (2005), who decorrelated the measured flux against a quantity called the "pixel phase", defined as:

$$S(x, y) = 1 + \frac{c_0 D}{\sqrt{2\pi}}, \tag{3.1}$$

where D is the distance of the PSF centroid from the center of the corresponding pixel. Reach et al. recommended a value of $c_0 = 0.0535$ for the 3.6μm channel,

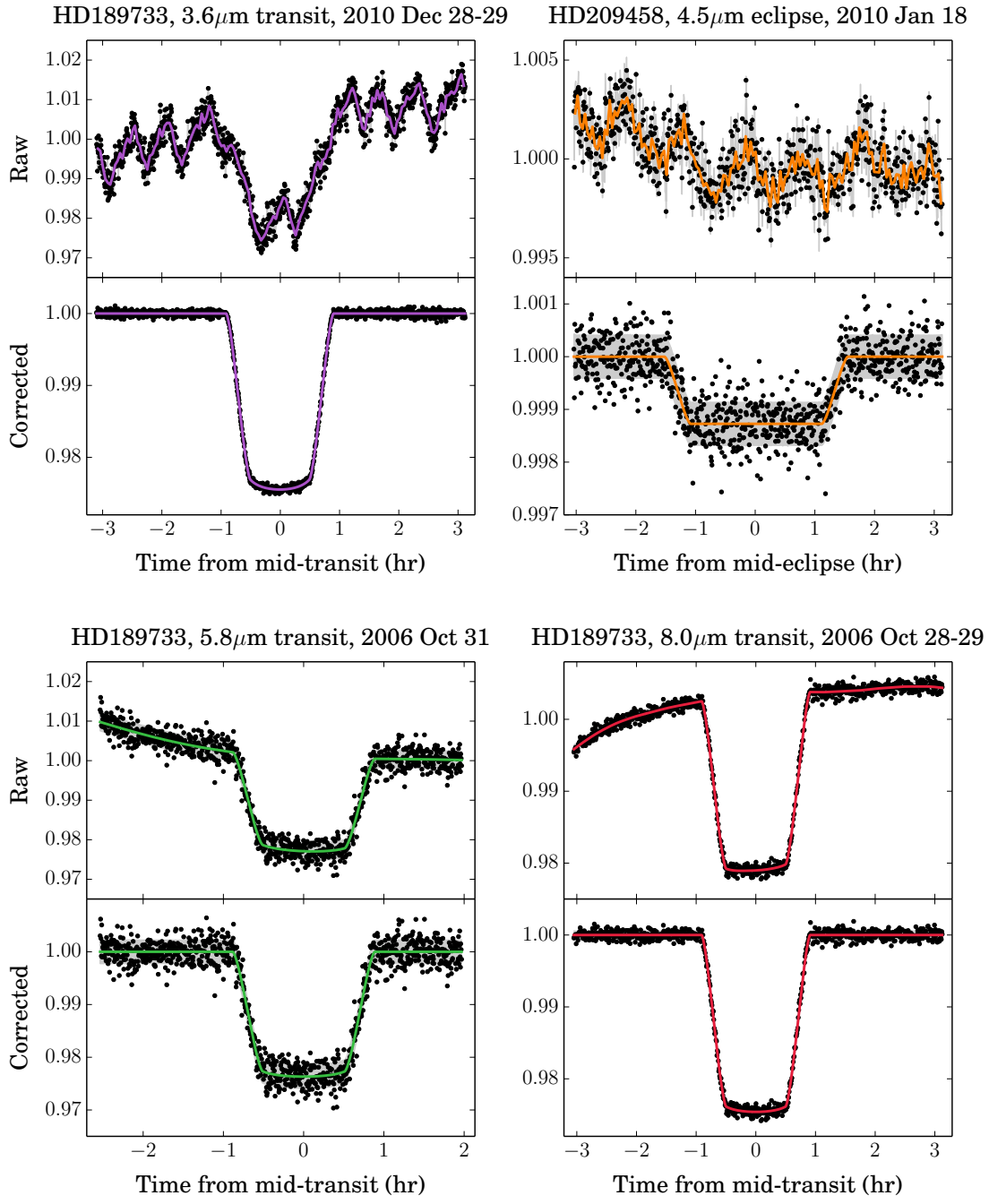

Figure 3.6: Example lightcurves illustrating the characteristic systematics for each of the four IRAC channels. Top panels show the raw lightcurves with best-fit GP means overplotted as solid lines and 1σ uncertainties indicated by grey shading. Bottom panels show the lightcurves once the non-planetary contribution to the GP mean function (i.e. the systematics component) has been divided out.

while Morales-Calderón et al. (2006) found that $c_0 = 0.05$ was suitable for the 4.5μm channel.

It is straightforward to extend the pixel phase correction to a simple polynomial decorrelation against the xy centroid coordinates, with sensitivity function given by:

$$S(x, y) = 1 + \sum_{i=1}^{n} a_i x^i + \sum_{j=1}^{m} b_j y^j \ . \tag{3.2}$$

Additional $x^i y^j$ cross-terms can be added if desired. The coefficients a_i and b_j are treated as free parameters to be optimised at the same time as the free parameters of the transit model. This has perhaps been the most widely used correction method (e.g. Charbonneau et al., 2008; Knutson et al., 2008; Désert et al., 2009, 2011b), as it performs adequately well and is simple to implement. Most authors have found 2nd or 3rd order polynomials to be sufficient, but the basis for this choice is usually somewhat arbitrary.

Ballard et al. (2010) introduced an alternative correction method that avoids the need to specify an explicit functional form for the intrapixel sensitivity variations. Instead, the sensitivity map is constructed from the measured fluxes $\boldsymbol{d}$ according to:

$$S(x_i, y_i) = c_0 \sum_{j \neq i} \exp\left[-\frac{(x_j - x_i)^2}{2\sigma_x^2}\right] \exp\left[-\frac{(y_j - y_i)^2}{2\sigma_y^2}\right] d_j \mathcal{M}_j \ , \tag{3.3}$$

where σ_x and σ_y must either be specified or fit for, c_0 is a normalisation constant, d_j is the jth flux measurement, and $\mathcal{M}$ is a mask that is set to zero for points that do not contribute to the sensitivity map estimation (e.g. those that are far away from the current data point, or those that are in-transit etc). Ballard et al. claimed that using this squared exponential kernel method, they were able to improve their systematics correction by mapping previously unrecognised structure in the intrapixel sensitivity, with a characteristic amplitude of ~100 ppm and length

scale of ~ 0.01 pixels. This method has subsequently been adopted in Knutson et al. (2012), Lewis et al. (2013), and Zellem et al. (2014).

Another approach advocated by Stevenson et al. (2012a) is to use bilinear interpolation on the xy plane to build up a map of the pixel sensitivity variations from the measured fluxes. This requires that a grid of knots on the xy plane be specified to be interpolated between. Decreasing the spacing between knots runs the risk of overfitting the data, while increasing the spacing may cause fine-scale structure to be missed. This raises the question of how to select the most appropriate knot spacing for a given dataset, which has so far been done in a rather ad hoc manner (e.g. Stevenson et al., 2012a,b, 2014; Cubillos et al., 2013; Blecic et al., 2013, 2014).

Furthermore, it has not been conclusively demonstrated that the low-amplitude features identified by Ballard et al. are bona fide xy correlations. Given that points nearby each other in x and y tend to be nearby each other in time as the telescope pointing drifts, it is conceivable that correlations uncovered by the Gaussian kernel and bilinear interpolation methods could actually be proxies for other time-dependent correlations, which may be unrelated to x and y. Indeed, Demory et al. (2011a) mention that from an extensive analysis of *Warm Mission* IRAC datasets, they identified a periodic variation in the effective detector gain of few dozens of ppm with a ~30–60 min period. Such a signal may have the potential to mimic a low-amplitude xy correlation, similar to that noted by Ballard et al.

Finally, pixelation is another effect that introduces correlations between the xy position of the PSF and the measured flux, which was alluded to in Section 3.4.2. Essentially, pixels that partially fall within the photometric aperture can contain a significant amount of the total stellar flux. Whether or not such pixels are counted towards the flux, however, will vary with the exact xy coordinates of

the PSF centroid. By interpolating the native pixel array onto a finer grid before summing the flux within the photometric aperture, the `pyriac` pipeline mitigates against this effect somewhat. A good discussion of pixelation as relevant to IRAC datasets is given in Stevenson et al. (2012a).

3.6.2 Detector ramps

Lightcurves acquired in the 5.8μm and 8.0μm channels exhibit a characteristic ramp behaviour, with the measured flux smoothly increasing or decreasing before levelling off (Figure 3.6). The amplitude of this effect is typically ~1–3%, with the steepest change in flux occurring over the first ~1 hr of observations. It can be attributed to the sensitivity of individual pixels varying as a function of time, where the rate of change depends upon the illumination level of the pixel (e.g. Knutson et al., 2007a).

It has been suggested (e.g. Deming et al., 2007; Agol et al., 2010; Seager & Deming, 2010) that the ramp effect is caused by electrons getting caught in charge traps within the detector. As more photons arrive, the charge traps fill up, resulting in less electrons getting trapped and a higher flux being measured. However, Seager & Deming (2010) acknowledge that the physics of the detectors are poorly understood and this simple picture may be incomplete or wrong. For instance, lightcurves measured in the 5.8μm channel exhibit a ramp-down behaviour (Figures 3.4–3.6), which is not obviously explained by charge trapping.

It has been standard to fit 5.8 μm and 8.0 μm lightcurves by multiplying the transit signal by a parametric model that provides a good approximation to the ramp behaviour. One approach is to remove the initial steep section of the ramp and fit the remainder with a linear or quadratic polynomial in time (e.g. Deming et al., 2007; Beaulieu et al., 2010). However, a more common approach (e.g. Charbonneau et al., 2008; Knutson et al., 2008; Désert et al.,

2009; Machalek et al., 2010) is to model the full baseline with a low-order polynomial in logarithmic time, given by:

$$R(t) \quad = \quad c_0 + \sum_{i=1}^{n} c_i \ln\left(t - b_i\right) , \tag{3.4}$$

as this is generally capable of providing a good match to the ramp. Alternatively, a double-exponential model of the form:

$$R(t) \quad = \quad c_0 + \exp\left(c_1 t - c_2\right) + \exp\left(c_3 t - c_4\right) , \tag{3.5}$$

has been advocated by Agol et al. (2010). Additional linear and quadratic time terms are also sometimes added to supplement Equations 3.4 and 3.5.

The lack of consensus on the most appropriate model to adopt for the ramp can be problematic, as subtle differences between the various ramp functions can potentially bias the inferred planet parameters. For example, the double exponential function (Equation 3.5) tends to retain slightly more curvature than a logarithmic polynomial (Equation 3.4) after the initial steep gradient. This can result in an underestimated transit depth if the ramp is decreasing (i.e. 5.8 μm channel), or vice versa if the ramp is increasing (i.e. 8.0 μm channel). In an attempt to avoid effects such as these, the lightcurve analysis is often performed separately for a number of different ramp functions, and the one that minimises the model residuals is retained. To reduce the risk of overfitting, some authors (e.g. Stevenson et al., 2012a) instead use the Bayesian information criterion (BIC) (Schwarz, 1978) to discriminate between ramp functions.[10]

[10] Under the assumption of independent Gaussian measurement errors, the Bayesian information criterion (BIC) is given by:

$$\mathrm{BIC} = \chi^2 + k \ln n , \tag{3.6}$$

where χ^2 is the chi-squared value of the best-fit model (Equation 2.23), k is the number of tunable model parameters, and n is the number of data points. The χ^2 term rewards a good fit to the data, while the $k \ln n$ term penalises adding free parameters to the model. It can be shown that the BIC is a crude approximation to the negative log of the model evidence (i.e. $p(\boldsymbol{d})$, see Chapter 2) to within an additive constant, assuming that the model evidence can be approximated as a normal distribution and the priors on the model parameters are broad.

3.7 Gaussian process analyses

In this study, the planet signal and systematics are modelled simultaneously using the GP methodology described in Section 2.3. If $\boldsymbol{d}$ is a vector containing the measured fluxes, we can rewrite Equation 2.25 to express the model likelihood as:

$$p(\boldsymbol{d}|\boldsymbol{\alpha},\boldsymbol{\gamma}) \quad = \quad \mathcal{N}\left(\boldsymbol{\mu},\mathbf{K}+\boldsymbol{\Sigma}\right) , \tag{3.7}$$

where $\boldsymbol{\mu}$ is the mean function, $\boldsymbol{\alpha}$ are the mean function parameters, $\boldsymbol{\gamma}$ are the covariance parameters, $\mathbf{K}$ is the covariance matrix, and $\boldsymbol{\Sigma}$ is the diagonal white noise matrix.

3.7.1 Mean functions

The mean function $\boldsymbol{\mu}$ defines the model for the astrophysical signal with well-understood form; namely, a primary transit or secondary eclipse. For this purpose, the analytic transit functions of Mandel & Agol (2002) are adopted (Section 1.4).

The orbital eccentricity is set to zero for both HD 189733b and HD 209458b, as implied by observational constraints (e.g. Pont et al., 2011) and in line with expectations that both orbits should be tidally circularised (Section 1.3). The orbital periods are also fixed to $P = 2.21857567$ days for HD 189733b (Agol et al., 2010) and $P = 3.52474859$ days for HD 209458b (Knutson et al., 2007b).

For the primary transits, the mean function parameters allowed to vary are the radius ratio $R_p/R_\star$, normalised semimajor axis $a/R_\star$, impact parameter $b = a\cos i/R_\star$, and transit mid-time $T_{\rm mid}$ (i.e. $\boldsymbol{\alpha} = \{R_p/R_\star\,,\, a/R_\star\,,\, b\,,\, T_{\rm mid}\}$ in Equation 3.7). Stellar limb darkening is treated using the nonlinear law of Claret (2004) with coefficients fixed to those provided by Sing (2010). The latter were calculated specifically for HD 189733 and HD 209458 in the IRAC bandpasses using 3D stellar models.

For the secondary eclipses, the only mean function parameters that are allowed to vary are the eclipse depths δ_s and eclipse mid-times $T_{\rm mid}$ (i.e. $\boldsymbol{\alpha} = \{\delta_s, T_{\rm mid}\}$ in Equation 3.7). The remaining mean function parameters are fixed to values published in the literature: for HD 189733b, I set $R_p/R_\star = 0.1555$ (Pont et al., 2013), and $a/R_\star = 8.86$ and $b = 0.6620$ (Agol et al., 2010); for HD 209458b, I set $R_p/R_\star = 0.1209$, $a/R_\star = 8.76$ and $b = 0.507$ (Torres et al., 2008).

3.7.2 Covariance kernels

For the 3.6μm and 4.5μm channel lightcurves, the entries of the covariance matrix **K** are parameterised as the sum of two kernels: a squared exponential kernel (Equation 2.27) with the centroid xy coordinates as inputs and a Matérn $\nu = 3/2$ kernel (Equation 2.29) with time t as the input. Writing this out explicitly, the combined kernel is given by:

$$k\left(\boldsymbol{v_i}, \boldsymbol{v_j}\right) = k_{xy} + k_t \,, \tag{3.8}$$

where:

$$k_{xy} = A_{xy}^2 \exp\left(-\left[\left(\frac{x_i - x_j}{L_x}\right)^2 + \left(\frac{y_i - y_j}{L_y}\right)^2\right]\right) , \tag{3.9}$$

$$k_t = A_t^2 \left[1 + \frac{t_i - t_j}{L_t}\sqrt{3}\right] \exp\left(-\frac{\left(t_i - t_j\right)}{L_t}\sqrt{3}\right) , \tag{3.10}$$

such that $\boldsymbol{\gamma} = \{A_{xy}, L_x, L_y, A_t, L_t\}$ in the notation of Equation 3.7. The squared exponential component of this kernel accounts for the smooth spatial variations in pixel sensitivities that dominate the systematics of the 3.6 μm and 4.5 μm channels at the ~ 1 % level (Section 3.6.1), while the Matérn component accounts for any residual correlated noise in the lightcurve. The latter have amplitudes of order ~10–100 ppm, and often exhibit a non-smooth structure, which is why the Matérn kernel is used instead of another squared exponential kernel

(Figures 2.3 and 2.4). As was mentioned in Section 3.6.1, possible sources of these low-amplitude, high-frequency correlations include, but are not limited to: photometric aperture losses due to pointing jitter combined with imperfect centroiding; pixelation effects; fine-scale variations in the intrapixel sensitivity; and time-dependent variations in the detector gain.

For the 5.8μm and 8.0μm channel lightcurves, the dominant systematic is the smooth, time-dependent ramp effect (Section 3.6.2). A squared exponential kernel is used to model this effect, with form given by:

$$k_\tau = A_\tau^2 \exp\left[-\left(\frac{\tau_i - \tau_j}{L_\tau} \right)^2 \right] , \tag{3.11}$$

where τ is a parameterisation of time t. In the simplest case, I adopt $\tau = t$. However, given the nature of the ramp effect – a steep initial gradient followed by a levelling off of the measured flux (Figure 3.6) – I also investigate logarithmic and exponential parameterisations of the form:

$$\tau = c_0 \ln(t - c_1) , \tag{3.12}$$

$$\tau = \exp(c_0 t - c_1) , \tag{3.13}$$

where c_0 and c_1 are parameters that can be inferred from the data (see below). As with the 3.6μm and 4.5μm channels, a time-dependent Matérn $\nu = 3/2$ kernel k_t is also included to account for residual correlations in the lightcurve. Therefore, the final kernel for the 5.8μm and 8.0μm channels is given by:

$$k(\boldsymbol{v_i}, \boldsymbol{v_j}) = k_\tau + k_t , \tag{3.14}$$

with covariance parameters $\boldsymbol{\gamma} = \{A_\tau, L_\tau, c_0, c_1, A_t, L_t\}$.

The covariance kernels outlined above allow for systematics treatments that are at least as versatile as others used in the literature. For instance, the k_{xy} component of Equation 3.8 is similar in concept to the squared exponential correction

developed by Ballard et al. (2010) (Equation 3.3). Similarly, the k_τ component of Equation 3.14 is reminiscent of the linear, exponential, and logarithmic time polynomials used in other published studies (Section 3.6.2). However, by parameterising the covariance rather than systematics signal directly, the GP model is capable of marginalising over a broader range of function space with relatively few tunable parameters. This is achieved within the framework of a self-consistent probabilistic model, allowing for a robust treatment of uncertainty complete with inbuilt Occam's razor (Section 2.3.2).

3.7.3 Lightcurve binning

In practice, GP models become computationally intractable for datasets with $N \gtrsim 1000$, due to the $\mathcal{O}\left(N^3\right)$ cost associated with each likelihood evaluation (Section 2.3.4). For this reason, I bin the fluxes and centroid xy coordinates in time prior to performing the Markov chain Monte Carlo (MCMC) analysis, which is described in the next section, as most of the datasets have $N >> 1000$ (Table 3.2).

Binning factors are chosen according to the format of individual lightcurves, such that the time interval between successive binned points is 15–30 sec and N = 800–1300. To achieve this, lightcurves obtained in full array mode for Programs 461, 30590, and 40732 are binned by a factor of two, giving a cadence of approximately 17 sec. Lightcurves obtained in subarray mode with frame times of 0.4 sec are binned in sets of 64 frames, giving a median cadence of 27 sec. Lightcurves obtained in subarray mode with frame times of 0.1 sec are binned in sets of 128 frames, giving a median cadence of 17 sec.

The exceptions to the above are the lightcurves of Programs 261 and 20523, which were acquired in subarray mode with frame times of 0.1 sec by cycling between each of the wavelength channels (Section 3.3). The resulting lightcurves

for each wavelength channel are thus interrupted by gaps, allowing them to be binned into sets of 64 frames with a median cadence of 8 sec while maintaining a manageable dataset size, i.e. $N = 250$–350.

3.7.4 MCMC analysis

Following Bayes theorem (Equation 2.1), the model posterior distribution is given by:

$$p(\boldsymbol{\alpha}, \boldsymbol{\gamma} | \boldsymbol{d}) \quad \propto \quad p(\boldsymbol{d} | \boldsymbol{\alpha}, \boldsymbol{\gamma})\, p(\boldsymbol{\alpha})\, p(\boldsymbol{\gamma}) \;, \tag{3.15}$$

where $p(\boldsymbol{d}|\boldsymbol{\alpha}, \boldsymbol{\gamma})$ is the GP likelihood given by Equation 3.7, and $p(\boldsymbol{\alpha})$ and $p(\boldsymbol{\gamma})$ are the priors on the mean function and covariance parameters. Uniform priors are adopted for the mean function parameters and covariance length scales. For the covariance amplitudes, Gamma priors[11] of the form $p(A_i) = \mathrm{Gam}(1, 100)$ for $i = \{t, xy, \tau\}$ are adopted. The latter give decreasing probability to increasing covariance amplitudes, encouraging the GP to reduce the covariance amplitude unless justified by the data.

A white noise scaling factor is also included in the model, such that the white noise associated with the ith data point is given by $\sigma_i = \beta \hat{\sigma}_i$, where $\hat{\sigma}_i$ is the photon noise value calculated from Poisson statistics, i.e. the square root of the flux. The ability to inflate the Poisson errorbars gives the model some additional flexibility for dealing with high-frequency noise that may be present in the data, without having to reduce the correlation length scales L_i of the covariance kernels to unreasonably small values. A uniform prior is adopted for β.

The model fitting for each lightcurve proceeds as follows. Values for the model parameters are drawn randomly from the model prior, i.e. $p(\boldsymbol{\alpha})\, p(\boldsymbol{\gamma})$. With this

[11] A Gamma probability distribution over the non-negative variable x is given by:

$$\mathrm{Gam}(q, h) = \frac{x^{q-1} e^{-hx}}{h^{-q} \Gamma(q)} \;,$$

where Γ is the Gamma function, q is the shape parameter, and h is the rate parameter.

as a starting point, Equation 3.15 is optimised using the Nelder-Mead simplex algorithm (Nelder & Mead, 1965) to obtain maximum likelihood estimates (MLEs) for the parameters. A short Metropolis-Hastings MCMC chain of 1000 steps is initiated at the MLE, with step sizes pretuned to give acceptance rates of 20–40 %. The median chain values are then used as the starting location for a second MLE optimisation. In practice, the randomness introduced by the short MCMC chains helps prevent the MLE optimisations getting trapped in local maxima of the likelihood surface, thus increasing the chance of the global likelihood maximum being located. To further increase this probability, the entire process is repeated ten times, each time from a different random starting point.

Before commencing the final MCMC chains, the covariance parameters are fixed to their MLE values, which are reported in Appendix B. This allows the expensive $\mathcal{O}\left(N^3\right)$ covariance matrix inversion in the GP likelihood evaluation (Equation 2.26) to be performed only once at the beginning of the chain, rather than at every step. The disadvantage is that by fixing the covariance parameters $\boldsymbol{\gamma}$, they cannot be marginalised. In effect, this imposes an artificial restriction on the range of systematics functions that can be explored by the GP model. Therefore, there may be degeneracies between the planet signal and systematics that are not properly incorporated into the final uncertainties for the planet parameters $\boldsymbol{\alpha}$, which will be underestimated. For instance, in their re-analysis of the NICMOS transmission spectrum for HD 189733b, Gibson et al. (2012b) found uncertainties that were up to ~ 1.5 larger when the covariance parameters were marginalised over rather than fixed to their MLE values.

It should also be emphasised, however, that fixing the covariance parameters is quite different to fixing the parameters of an explicit functional model for the systematics. Instead, fixing the covariance parameters is somewhat analogous to selecting a parameteric form for the systematics model (e.g. Equations 3.4

and 3.5), as they control the high-level properties of the function space spanned by the GP model. Rather than selecting from a handful of distinct parametric models, the GP model offers access to a continuum of possible functions. By selecting covariance parameters that optimise the GP likelihood, this continuum is narrowed in a principled manner.

Having fixed the covariance parameters, an initial chain of 10^5 steps is run with the planet parameters allowed to vary and step sizes again pretuned to ensure acceptance rates of 20–40 %. The first 5×10^4 steps are discarded as burn-in. An additional four chains are then run for 10^5 steps each, with starting parameter values drawn randomly from normal distributions centered on the mean values of the first chain. The width of the normal distributions are taken to be five times the standard deviation of the first chain, to ensure the starting locations are well-dispersed in parameter space. After discarding the first 5×10^4 burn-in steps of these chains, the Gelman-Rubin statistics for each parameter are calculated. For all lightcurves, these are found to be well within 1 % of unity, suggesting the chains have reached stable states (Section 2.2.1). Finally, the five independent chains are combined into a single chain, giving 2.5×10^5 samples from the posterior distribution.

3.8 Results

3.8.1 HD 189733

Results of the primary transit and secondary eclipse MCMC analyses for HD 189733 are given in Tables 3.3 and 3.4, respectively. For all but two of the 5.8 μm and 8.0 μm lightcurves, the recovered values were fully consistent regardless of the adopted parameterisation for the GP model (i.e. linear, exponential, or logarithmic; see Section 3.7) so only those obtained using the logarithmic parameterisation are quoted. The exceptions were the 5.8 μm 2006 Oct 31 and 8.0 μm

Table 3.3: Results of MCMC primary transit analyses for HD 189733b. Quoted values are the chain medians, and uncertainties give the ranges either side of the medians that contain 34% of the chain samples. Orbital inclination i values are derived from the impact parameter $b = a\cos i/R_\star$ and system scale $a/R_\star$ values. Results for the three time parameterisations are given separately in Table 3.5 for the 5.8 μm 2006 Oct 31 and 8.0 μm 2006 Oct 29 primary transits (see Sections 3.7.2 and 3.8.1.1 for details). Note that the $R_p/R_\star$ values reported here have not been corrected for the effect of unocculted star spots.

Transits						
Channel (μm)	Date	$T_{\rm mid}$ (BJD$_{\rm UTC}$ − 2450 000)	$R_p/R_\star$	$a/R_\star$	b	i (deg)
3.6	2006 Oct 31	$4039.830608^{+0.000080}_{-0.000081}$	$0.15458^{+0.00045}_{-0.00044}$	$9.051^{+0.106}_{-0.101}$	$0.642^{+0.010}_{-0.011}$	$85.931^{+0.115}_{-0.111}$
	2007 Nov 25	$4430.299878^{+0.000038}_{-0.000039}$	$0.15559^{+0.00024}_{-0.00024}$	$8.968^{+0.036}_{-0.037}$	$0.656^{+0.003}_{-0.004}$	$85.807^{+0.039}_{-0.039}$
	2010 Dec 28-29	$5559.554538^{+0.000041}_{-0.000040}$	$0.15458^{+0.00021}_{-0.00022}$	$8.881^{+0.036}_{-0.036}$	$0.664^{+0.003}_{-0.003}$	$85.713^{+0.039}_{-0.040}$
4.5	2007 Nov 23	$4428.081187^{+0.000151}_{-0.000152}$	$0.15604^{+0.00120}_{-0.00122}$	$8.956^{+0.129}_{-0.127}$	$0.648^{+0.012}_{-0.013}$	$85.853^{+0.141}_{-0.138}$
	2009 Dec 23	$5189.052518^{+0.000036}_{-0.000036}$	$0.15601^{+0.00020}_{-0.00020}$	$8.912^{+0.035}_{-0.034}$	$0.659^{+0.003}_{-0.003}$	$85.758^{+0.038}_{-0.037}$
8.0	2007 Nov 23	$4428.081586^{+0.000169}_{-0.000167}$	$0.15599^{+0.00117}_{-0.00119}$	$8.964^{+0.170}_{-0.175}$	$0.663^{+0.015}_{-0.016}$	$85.759^{+0.177}_{-0.182}$

Table 3.4: Same as Table 3.3, but giving the results of MCMC secondary eclipse lightcurve analyses for HD 189733b. Brightness temperatures T_b are derived from the measured eclipse depths δ_s assuming an ATLAS stellar model for HD 189733 and integrating over the IRAC bandpasses, where the brightness temperature is defined as the temperature that the planet would require if it radiated as a blackbody and had the same bandpass-integrated flux.

Eclipses				
Channel (μm)	Date	$T_{\rm mid}$ (BJD$_{\rm UTC}$ − 2 450 000)	δ_s (%)	T_b (K)
3.6	2005 Nov 24	$3699.278937^{+0.001140}_{-0.001213}$	$0.1591^{+0.0235}_{-0.0232}$	1370^{+62}_{-66}
	2010 Dec 27-28	$5558.445098^{+0.001336}_{-0.001027}$	$0.1362^{+0.0241}_{-0.0241}$	1305^{+68}_{-74}
	2010 Dec 30	$5560.668125^{+0.001295}_{-0.001082}$	$0.1676^{+0.0095}_{-0.0097}$	1393^{+25}_{-26}
4.5	2005 Nov 24	$3699.279252^{+0.001040}_{-0.000965}$	$0.1954^{+0.0239}_{-0.0241}$	1211^{+51}_{-53}
	2009 Dec 22	$5187.944127^{+0.000401}_{-0.000393}$	$0.1851^{+0.0045}_{-0.0044}$	1188^{+10}_{-10}
	2009 Dec 24	$5190.160043^{+0.000971}_{-0.000869}$	$0.1695^{+0.0132}_{-0.0128}$	1153^{+30}_{-30}
5.8	2005 Nov 24	$3699.282843^{+0.002196}_{-0.001815}$	$0.3613^{+0.0751}_{-0.0757}$	1385^{+128}_{-136}
8.0	2005 Nov 24	$3699.278644^{+0.002251}_{-0.002058}$	$0.2568^{+0.0768}_{-0.0789}$	1074^{+146}_{-162}
	2006 Oct 30	$4038.723078^{+0.000302}_{-0.000322}$	$0.3362^{+0.0067}_{-0.0066}$	1225^{+12}_{-12}

2006 Oct 29 primary transit lightcurves, for which the recovered $R_p/R_\star$ values differed by up to $\sim 2\sigma$ depending on the adopted time parameterisation. Results for these latter two lightcurves are reported separately in Table 3.5 for each time parameterisation, and described in more detail in Section 3.8.1.1. Corrected transit and eclipse lightcurves[1] are shown in Figure 3.7.

3.8.1.1 Transmission

The top two rows of Figure 3.8 compare the radius ratio $R_p/R_\star$ values obtained obtained for HD 189733 in the current study with those previously reported in the literature for the same lightcurves. Overall, the results of the current study are in fairly good agreement with the latter. The most significant exception is the 3.6 μm 2006 Oct 31 lightcurve, for which the following values

[1] To remove the systematics component for plotting purposes, I have simply divided the $\boldsymbol{k}_\star\,(\mathbf{K}+\boldsymbol{\Sigma})^{-1}\,\boldsymbol{r}$ term of the best-fit GP predictive mean (Equation 2.32) out of the data.

Table 3.5: MCMC primary transit results for the HD 189733b 5.8μm 2006 Oct 31 and 8.0μm 2006 Oct 29 lightcurves. Inferred planet parameters for these two lightcurves were more sensitive to the choice of time parameterisation than the other 5.8μm and 8.0μm lightcurves analysed in this study. The logarithmic time parameterisation is favoured for the reasons given in Section 3.8.1.1, but results for the linear and exponential parameterisations are also reported for completeness. Note that the $R_p/R_\star$ values reported here have not been corrected for the effect of unocculted star spots.

5.8μm, 2006 Oct 31					
Time input	$T_{\rm mid}$ (BJD$_{\rm UTC}-2\,450\,000$)	$R_p/R_\star$	$a/R_\star$	b	i (deg)
linear	$4039.830604^{+0.000129}_{-0.000128}$	$0.15238^{+0.00104}_{-0.00103}$	$9.223^{+0.154}_{-0.152}$	$0.639^{+0.015}_{-0.016}$	$86.028^{+0.161}_{-0.159}$
exponential	$4039.830553^{+0.000117}_{-0.000118}$	$0.15299^{+0.00065}_{-0.00066}$	$9.204^{+0.147}_{-0.143}$	$0.640^{+0.014}_{-0.015}$	$86.013^{+0.155}_{-0.150}$
logarithmic	$4039.830504^{+0.000113}_{-0.000114}$	$0.15455^{+0.00062}_{-0.00062}$	$9.158^{+0.139}_{-0.148}$	$0.642^{+0.015}_{-0.014}$	$85.983^{+0.145}_{-0.158}$
8.0μm, 2006 Oct 28-29					
Time input	$T_{\rm mid}$ (BJD$_{\rm UTC}-2\,450\,000$)	$R_p/R_\star$	$a/R_\star$	b	i (deg)
linear	$4037.611995^{+0.000047}_{-0.000047}$	$0.15582^{+0.00060}_{-0.00060}$	$8.885^{+0.047}_{-0.044}$	$0.663^{+0.004}_{-0.004}$	$85.722^{+0.050}_{-0.047}$
exponential	$4037.611994^{+0.000058}_{-0.000057}$	$0.15532^{+0.00083}_{-0.00081}$	$8.885^{+0.046}_{-0.047}$	$0.664^{+0.004}_{-0.004}$	$85.715^{+0.049}_{-0.051}$
logarithmic	$4037.612015^{+0.000049}_{-0.000050}$	$0.15467^{+0.00035}_{-0.00035}$	$8.890^{+0.045}_{-0.044}$	$0.663^{+0.004}_{-0.004}$	$85.723^{+0.049}_{-0.047}$

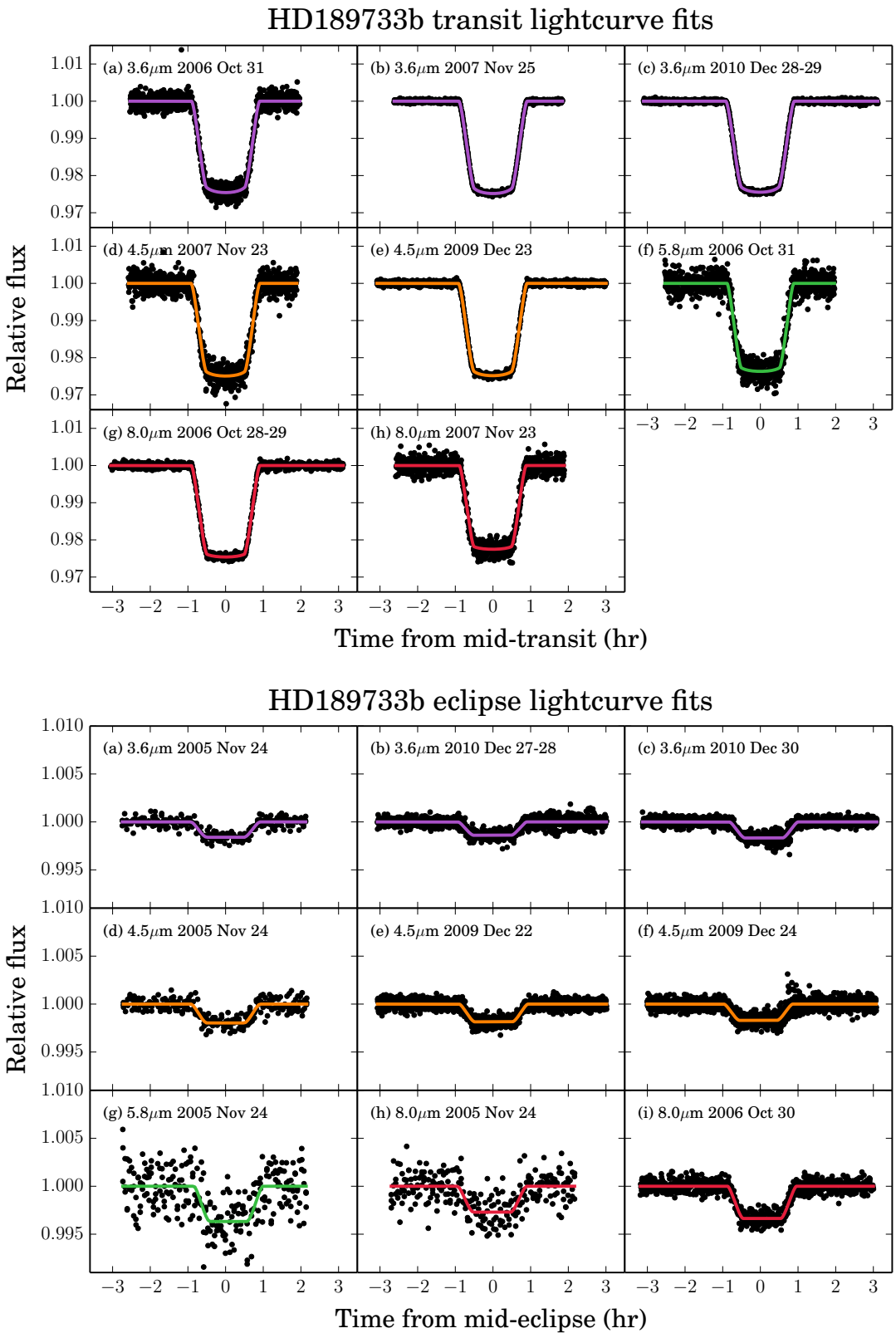

Figure 3.7: Corrected primary transit (top) and secondary eclipse (bottom) lightcurves for HD 189733b, obtained by dividing the raw lightcurves by the non-planetary component of the best-fit GP mean function. Note that these are the binned lightcurves used for the fitting, as described in Section 3.7.

are obtained: $R_p/R_\star = 0.15458^{+0.00045}_{-0.00044}$ (current study), $R_p/R_\star = 0.15285 \pm 0.0003$ (Beaulieu et al., 2008), and $R_p/R_\star = 0.1545 \pm 0.0003$ (Désert et al., 2009). My value is consistent with that of Désert et al., but differs from that of Beaulieu et al. by $> 3\sigma$. I also note that my uncertainty is ~ 1.5 times larger than that reported by both Beaulieu et al. and Désert et al. Indeed, for all lightcurves, my uncertainties are ~1–5 times larger than those reported in the literature. I believe that my uncertainties are the more realistic, as the GP model used in the current study allows for greater flexibility in the systematics treatment while at the same time avoiding problems of overfitting (Section 2.3.2).

There is also some discrepancy between the $R_p/R_\star$ values recovered for the three 3.6 μm lightcurves. However, as has been pointed out by Désert et al. (2011b), these differences can be largely attributed to variations in the fractional star spot coverage between the epochs – the effective radius inferred for the planet increases as the fractional coverage of unocculted spots increases, and vice versa. In Table 3.6, I list the $R_p/R_\star$ values for all channels after applying corrections for this effect provided by Pont et al. (2013), which were determined by performing a GP interpolation of ground-based photometric data aquired as part of the long-term monitoring program of Henry (1999). The resulting spot-corrected 3.6 μm values are consistent with the mean value at the $\sim 1\sigma$ level.

As was noted above, the $R_p/R_\star$ values recovered for the 5.8 μm 2006 Oct 31 and 8.0 μm 2006 Oct 29 lightcurves varied somewhat according to the adopted time parameterisation of the GP model (Table 3.5). The $R_p/R_\star$ results for all three parameterisations are shown in Figure 3.8, and the differences between the lightcurve fits are shown in Figure 3.9. From now on, I adopt the values obtained with the logarithmic parameterisation, as these give the best agreement with those reported in the literature. A more satisfactory solution, however, would be

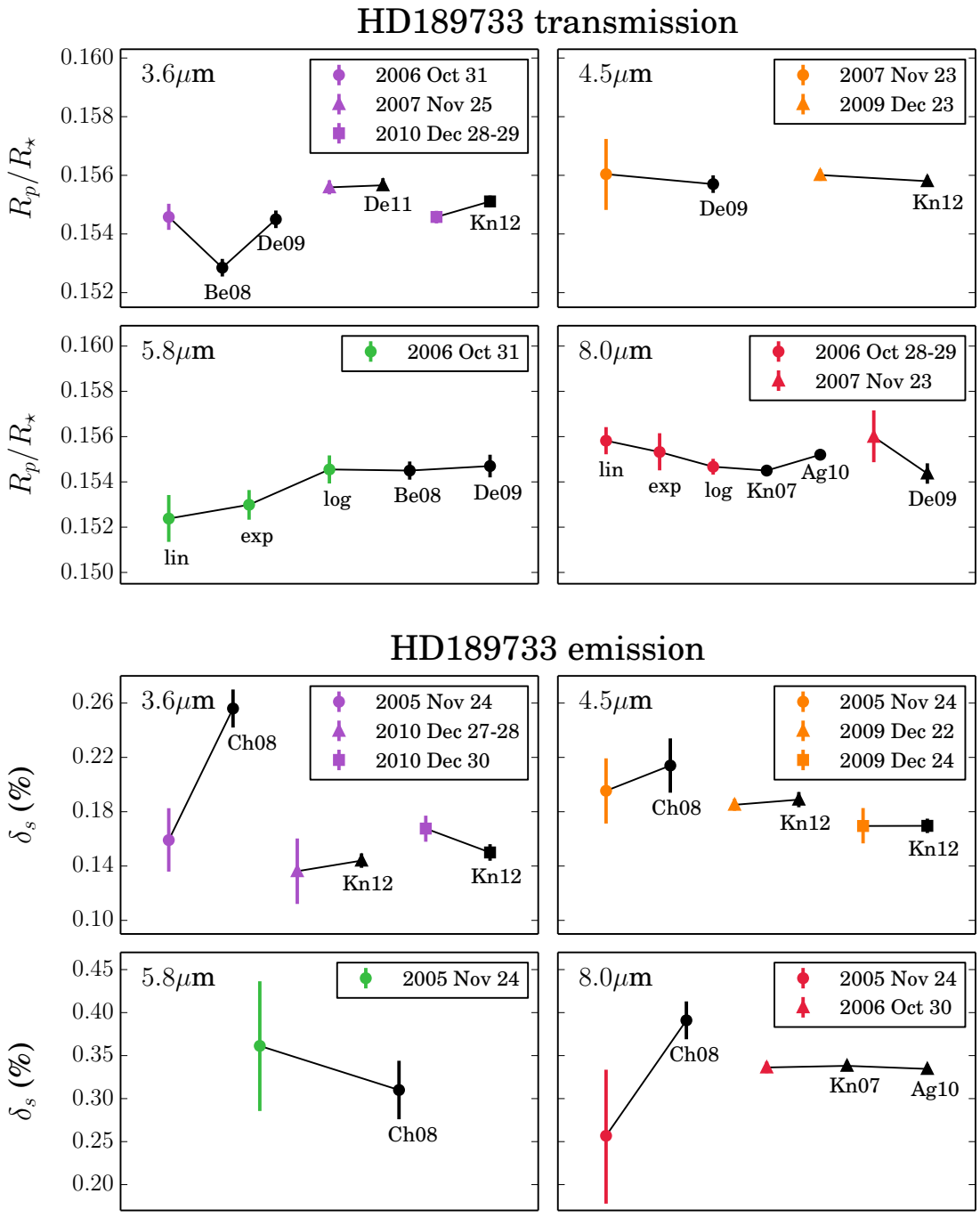

Figure 3.8: Comparison of transmission $R_p/R_\star$ (top) and emission δ_s (bottom) results for HD 189733b obtained in the present study with those published in the literature. Wavelength channels are indicated in the top left corner of each axis. Independent analyses of the same dataset are linked by solid black lines: coloured symbols show the values obtained in the current study and black symbols show values obtained by other authors. The latter are labelled using the same abbreviations adopted in Table 3.1. Unless indicated otherwise, results of the present study shown for the 5.8 μm and 8.0 μm datasets are those obtained using the logarithmic time parameterisation, as these were fully consistent with both the linear and exponential time parameterisations (see Section 3.7.2 for details). The exceptions are the 5.8 μm 2007 Nov 23 and 8.0 μm 2006 Oct 29 transmission datasets, for which all three time parameterisations are shown, as these exhibited a larger spread in the recovered $R_p/R_\star$ values.

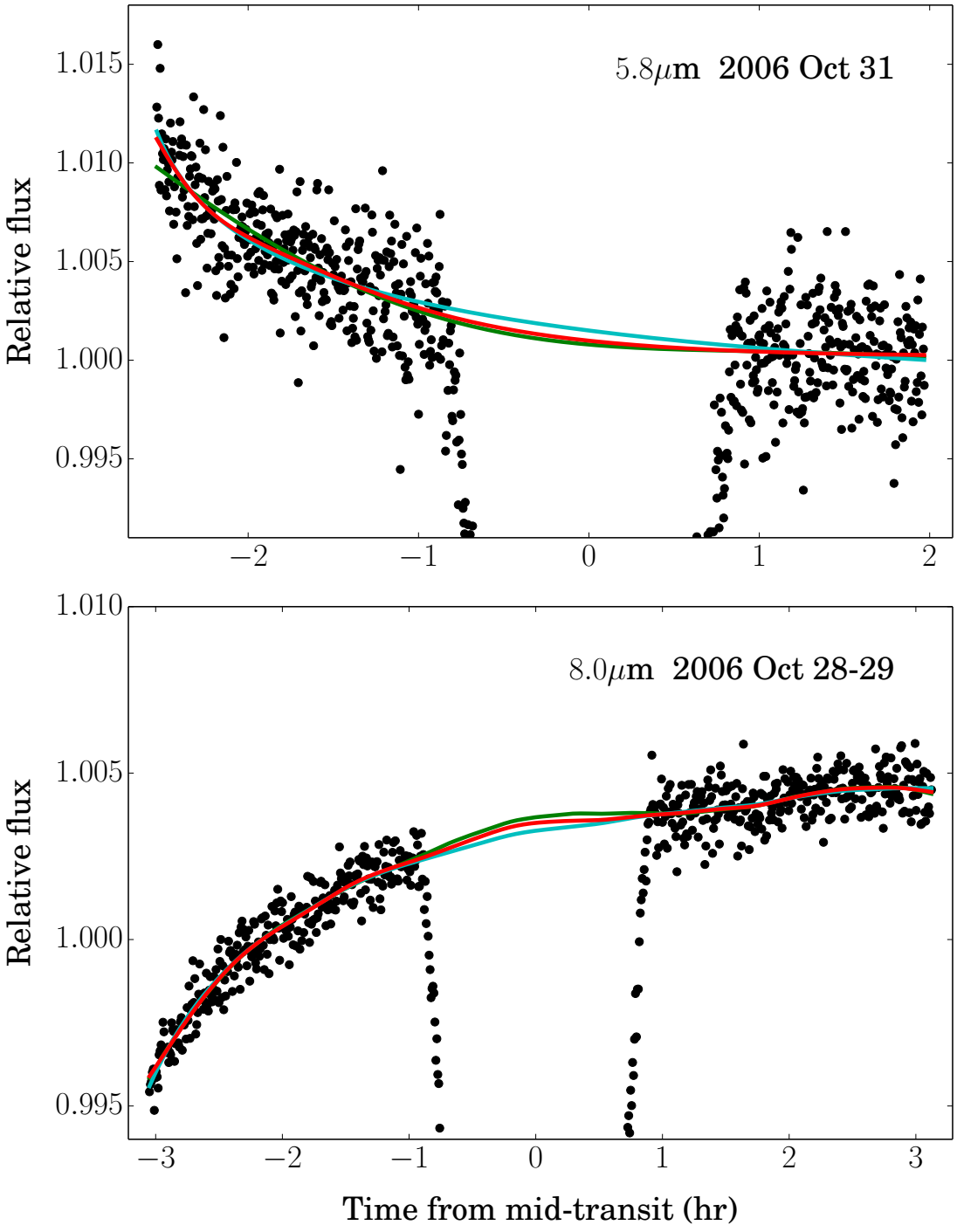

Figure 3.9: HD 189733b primary transit lightcurves for 5.8μm 2006 Oct 31 (top) and 8.0μm 2006 Oct 29 (bottom), illustrating the ramp fits obtained using the three different time parameterisations (see Section 3.7.2 for details). Solid lines show the best-fit GP means after dividing through by the planetary components, with colours corresponding to the different time parameterisations: logarithmic (light blue), linear (green), and exponential (red). The logarithmic parameterisation exhibits less curvature, resulting in larger $R_p/R_\star$ values inferred for the 5.8μm decreasing ramp and smaller $R_p/R_\star$ values inferred for the 8.0μm increasing ramp compared to the linear and exponential parameterisations, as shown in Figure 3.8.

Table 3.6: HD 189733b $R_p/R_\star$ results corrected for the effect of varying star spot coverages, following Pont et al. (2013). Note that no corrections have been applied to the 8.0μm values, as the effect of spots is negligible in this channel.

Channel (μm)	Date	$T_{\rm mid}$ (BJD$_{\rm UTC}$ − 2450000)	$R_p/R_\star$
3.6	2006 Oct 31	$4039.830608^{+0.000080}_{-0.000081}$	$0.15556^{+0.00053}_{-0.00052}$
	2007 Nov 25	$4430.299878^{+0.000038}_{-0.000039}$	$0.15464^{+0.00047}_{-0.00047}$
	2010 Dec 28-29	$5559.554538^{+0.000041}_{-0.000040}$	$0.15399^{+0.00059}_{-0.00060}$
4.5	2007 Nov 23	$4428.081187^{+0.000151}_{-0.000152}$	$0.15577^{+0.00126}_{-0.00129}$
	2009 Dec 23	$5189.052518^{+0.000036}_{-0.000036}$	$0.15564^{+0.00049}_{-0.00049}$
5.8	2006 Oct 31	$4039.830581^{+0.000131}_{-0.000130}$	$0.15465^{+0.00074}_{-0.00074}$
8.0	2006 Oct 28-29	$4037.612015^{+0.000049}_{-0.000050}$	$0.15467^{+0.00035}_{-0.00035}$
	2007 Nov 23	$4428.081586^{+0.000169}_{-0.000167}$	$0.15599^{+0.00117}_{-0.00119}$

to use a method such as nested sampling (Section 2.2.2) to determine which of the three parameterisations has the highest model evidence value (Equation 2.4). Such an approach will be investigated for future analyses.

3.8.1.2 Emission

The bottom two rows of Figure 3.8 compare the secondary eclipse depth δ_s values obtained for HD 189733 in the current study with those previously reported in the literature. As with the $R_p/R_\star$ values, I generally find good agreement with the published results, but with larger uncertainties. However, for the 3.6μm 2005 Nov 24 lightcurve, my value of $\delta_s = 0.1591^{+0.0235}_{-0.0232}$% is 3.5$\sigma$ lower than the value of $\delta_s = 0.256 \pm 0.014$% published by Charbonneau et al. (2008). This revision greatly improves the agreement with the 2010 Dec 27 and 30 lightcurves, for which I obtain $\delta_s = 0.1362^{+0.0241}_{-0.0241}$% and $\delta_s = 0.1676^{+0.0095}_{-0.0097}$%, respectively.

For the 8.0μm 2006 Oct 30 lightcurve, the value of $\delta_s = 0.3362^{+0.0067}_{-0.0066}$% obtained in the current study is consistent with the values of $\delta_s = 0.3381 \pm 0.0055$% and $\delta_s = 0.3345 \pm 0.0057$% reported by Knutson et al. (2007a) and Agol et al. (2010), respectively. My revised value of $\delta_s = 0.2568^{+0.0768}_{-0.0789}$% for the 2005

Nov 24 lightcurve is also consistent with these results, unlike the value of $\delta_s = 0.391 \pm 0.022\,\%$ originally reported by Charbonneau et al. (2008).

3.8.2 HD 209458

Results of the primary transit and secondary eclipse MCMC analyses for HD 209458 are given in Tables 3.7 and 3.8, respectively. As with the HD 189733 results, values quoted for the 5.8 μm and 8.0 μm lightcurves are those obtained using the logarithmic time parameterisation for the GP model, although fully consistent results were obtained using the linear and exponential parameterisations. Corrected transit and eclipse lightcurves are shown in Figure 3.10.

Note that Table 3.7 also provides the results for the fit to the transit lightcurve taken as part of the full-phase observation over 2011 Jan 12–16. However, a large amount of noise is observed during the second half of the transit, which is not corrected for by the systematics treatment (Figure 3.10). This lightcurve has been classed as a failed observation by the *Spitzer* Science Center, so it is discarded from further consideration.

3.8.2.1 Transmission

The top two rows of Figure 3.11 compare the radius ratio $R_p/R_\star$ values obtained for HD 209458 in the current study with those that have previously been published. As for HD 189733, my uncertainties are comparable to, or larger than, the literature uncertainties.

For the 3.6 μm 2007 Dec 31 lightcurve, I find $R_p/R_\star = 0.12058^{+0.00065}_{-0.00065}$, which is close to the value of $R_p/R_\star = 0.120835 \pm 0.00054$ reported by Beaulieu et al. (2010). However, I find $R_p/R_\star = 0.12225^{+0.00066}_{-0.00066}$ for the 2008 Jul 19 lightcurve, which is $\sim 1.8\sigma$ higher. This result was confirmed by an independent fit to the same lightcurve performed by N. Gibson (2014, private communication). The discrepancy is unlikely to be due to variations in star spot coverage, as HD 209458

Table 3.7: Same as Table 3.3, but for HD 209458b.

Transits						
Channel (μm)	Date	$T_{\rm mid}$ (BJD$_{\rm UTC}-2\,450\,000$)	$R_p/R_\star$	$a/R_\star$	b	i (deg)
3.6	2007 Dec 31	$4465.636997^{+0.000238}_{-0.000241}$	$0.12058^{+0.00065}_{-0.00065}$	$8.809^{+0.261}_{-0.247}$	$0.516^{+0.038}_{-0.044}$	$86.639^{+0.375}_{-0.349}$
	2008 Jul 18-19	$4666.547400^{+0.000218}_{-0.000217}$	$0.12225^{+0.00066}_{-0.00066}$	$8.759^{+0.257}_{-0.284}$	$0.508^{+0.043}_{-0.045}$	$86.675^{+0.381}_{-0.406}$
	2011 Jan 14	$5575.932594^{+0.000444}_{-0.000447}$	$0.11007^{+0.00388}_{-0.00398}$	$8.719^{+0.286}_{-0.263}$	$0.534^{+0.039}_{-0.047}$	$86.491^{+0.411}_{-0.372}$
4.5	2008 Jul 22	$4670.072697^{+0.000291}_{-0.000291}$	$0.12151^{+0.00092}_{-0.00091}$	$9.225^{+0.300}_{-0.316}$	$0.436^{+0.057}_{-0.065}$	$87.292^{+0.476}_{-0.464}$
	2010 Jan 19-20	$5216.405623^{+0.000071}_{-0.000071}$	$0.12104^{+0.00027}_{-0.00026}$	$8.886^{+0.063}_{-0.062}$	$0.495^{+0.011}_{-0.011}$	$86.809^{+0.092}_{-0.091}$
5.8	2007 Dec 31	$4465.636784^{+0.000281}_{-0.000281}$	$0.12188^{+0.00123}_{-0.00121}$	$9.271^{+0.309}_{-0.353}$	$0.412^{+0.070}_{-0.075}$	$87.451^{+0.529}_{-0.547}$
	2008 Jul 18-19	$4666.547224^{+0.000300}_{-0.000299}$	$0.12138^{+0.00135}_{-0.00133}$	$9.151^{+0.375}_{-0.388}$	$0.453^{+0.068}_{-0.081}$	$87.163^{+0.595}_{-0.573}$
8.0	2007 Dec 23-24	$4458.587581^{+0.000192}_{-0.000189}$	$0.11963^{+0.00105}_{-0.00105}$	$8.699^{+0.137}_{-0.136}$	$0.527^{+0.020}_{-0.022}$	$86.525^{+0.197}_{-0.189}$
	2008 Jul 22	$4670.072255^{+0.000293}_{-0.000294}$	$0.11928^{+0.00249}_{-0.00246}$	$8.745^{+0.221}_{-0.210}$	$0.511^{+0.033}_{-0.039}$	$86.649^{+0.331}_{-0.303}$

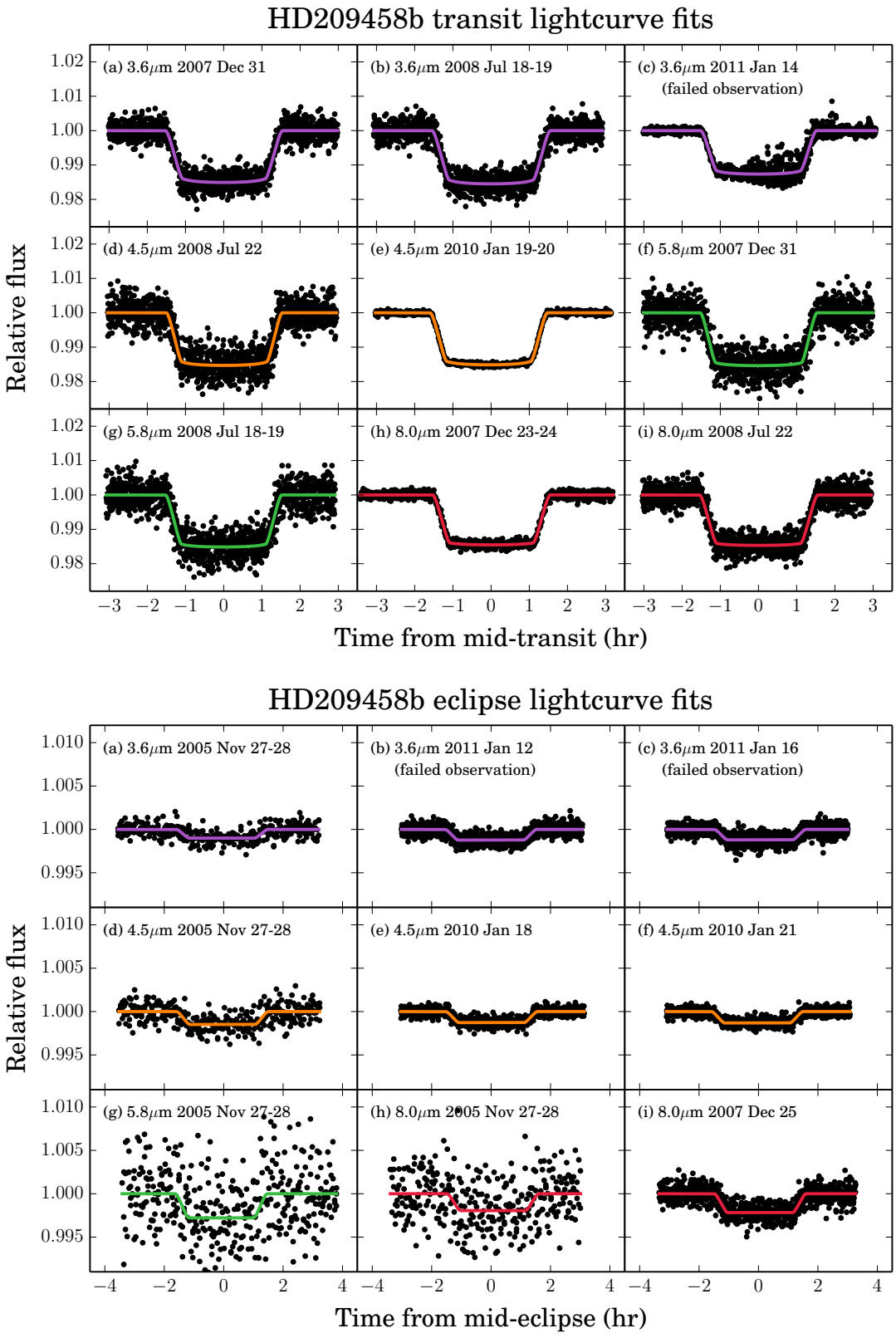

Figure 3.10: Same as Figure 3.7, but for HD 209458b, showing the corrected primary transit (top) and secondary eclipse (bottom) lightcurves.

Table 3.8: Same as Table 3.4, but for HD 209458b. Brightness temperatures T_b are derived from the measured eclipse depths δ_s assuming an ATLAS stellar model for HD 209458 and integrating over the IRAC bandpasses.

Eclipses				
Channel (μm)	Date	$T_{\rm mid}$ (BJD$_{\rm UTC}$ − 2 450 000)	δ_s (%)	T_b (K)
3.6	2005 Nov 27-28	$3702.526095^{+0.002478}_{-0.002303}$	$0.0996^{+0.0158}_{-0.0155}$	1481^{+77}_{-81}
	2011 Jan 12	$5574.169632^{+0.001802}_{-0.001734}$	$0.1229^{+0.0097}_{-0.0097}$	1594^{+44}_{-46}
	2011 Jan 16	$5577.697057^{+0.002533}_{-0.002346}$	$0.1190^{+0.0165}_{-0.0169}$	1576^{+76}_{-82}
4.5	2005 Nov 27-28	$3702.526761^{+0.002258}_{-0.002375}$	$0.1481^{+0.0184}_{-0.0183}$	1488^{+74}_{-76}
	2010 Jan 18	$5214.646591^{+0.001139}_{-0.001021}$	$0.1268^{+0.0128}_{-0.0129}$	1399^{+54}_{-56}
	2010 Jan 21	$5218.168947^{+0.001644}_{-0.001470}$	$0.1308^{+0.0122}_{-0.0124}$	1416^{+51}_{-53}
5.8	2005 Nov 27-28	$3702.524604^{+0.004003}_{-0.004075}$	$0.2759^{+0.0671}_{-0.0659}$	1786^{+226}_{-232}
8.0	2005 Nov 27-28	$3702.531964^{+0.004480}_{-0.004434}$	$0.1902^{+0.0407}_{-0.0399}$	1309^{+145}_{-148}
	2007 Dec 25	$4460.351677^{+0.001010}_{-0.001112}$	$0.2176^{+0.0104}_{-0.0102}$	1407^{+37}_{-36}

is known to be a photometrically quiet star (e.g. Knutson et al., 2010). Indeed, Beaulieu et al. (2010) obtained a value $R_p/R_\star = 0.120387 \pm 0.00053$ for the 2008 Jul 19 lightcurve, which is in good agreement with their value for the 2007 Dec 31 lightcurve. It remains unclear why the current study did not obtain such consistent results between the two epochs.

For the 5.8 μm 2007 Dec 31 and 2008 Jul 19 lightcurves, I find $R_p/R_\star = 0.12188^{+0.00123}_{-0.00121}$ and $R_p/R_\star = 0.12138^{+0.00135}_{-0.00133}$, respectively. Although these values are consistent with each other, they are significantly lower than the corresponding values of $R_p/R_\star = 0.1246 \pm 0.00095$ and $R_p/R_\star = 0.1244 \pm 0.00059$ obtained by Beaulieu et al. (2010).

I also find good agreement between the 8.0 μm 2007 Dec 24 and 2008 Jul 22 lightcurves, with $R_p/R_\star = 0.11963^{+0.00105}_{-0.00105}$ and $R_p/R_\star = 0.11928^{+0.00249}_{-0.00246}$. My value for the second epoch is significantly lower than the value of $R_p/R_\star = 0.1240 \pm 0.00046$ found by Beaulieu et al. (2010), while the first epoch lightcurve

was unpublished at the time of writing.

Meanwhile, my results for the 4.5 μm lightcurves are in good agreement both with each other and with the values previously published by Beaulieu et al. (2010) and Zellem et al. (2014).

3.8.2.2 Emission

The bottom two rows of Figure 3.11 compare the secondary eclipse depth δ_s values obtained for HD 209458 in the current study with those previously reported in the literature. I find consistent results across epochs for all wavelength channels, and once again I mostly find good agreement with previously published values, but with larger uncertainties. The only exception is for the 4.5 μm 2005 Nov 28 lightcurve, for which I obtain $\delta_s = 0.1481^{+0.0184}_{-0.0183}$%, which is $\sim 2.7\sigma$ lower than the value of $\delta_s = 0.213 \pm 0.015$% published by Knutson et al. (2008). However, my revision brings the value into good agreement with those obtained in the current study and by Zellem et al. (2014) for the two eclipses measured as part of the 2010 Jan 17–21 full-phase observation.

3.8.3 Orbit parameters

The normalised semimajor axis $a/R_\star$ and orbital inclination i values recovered from the primary lightcurve analyses for HD 189733 and HD 209458 are shown in Figure 3.12. Note the clear correlation between both parameters, i.e. higher values of $a/R_\star$ are associated with higher values of i, and vice versa. This reflects the fact that both parameters exert opposing influences on the transit duration, assuming a fixed orbital period: as i increases towards 90 deg the transit duration increases, but as $a/R_\star$ increases the transit duration decreases.

In theory, $a/R_\star$ and i are constant in time and across the different wavelength channels. Combining the measurements for HD 189733, I obtain weighted means of $a/R_\star = 8.92 \pm 0.02$ and $i = 85.76 \pm 0.02$ deg. Both of these values are

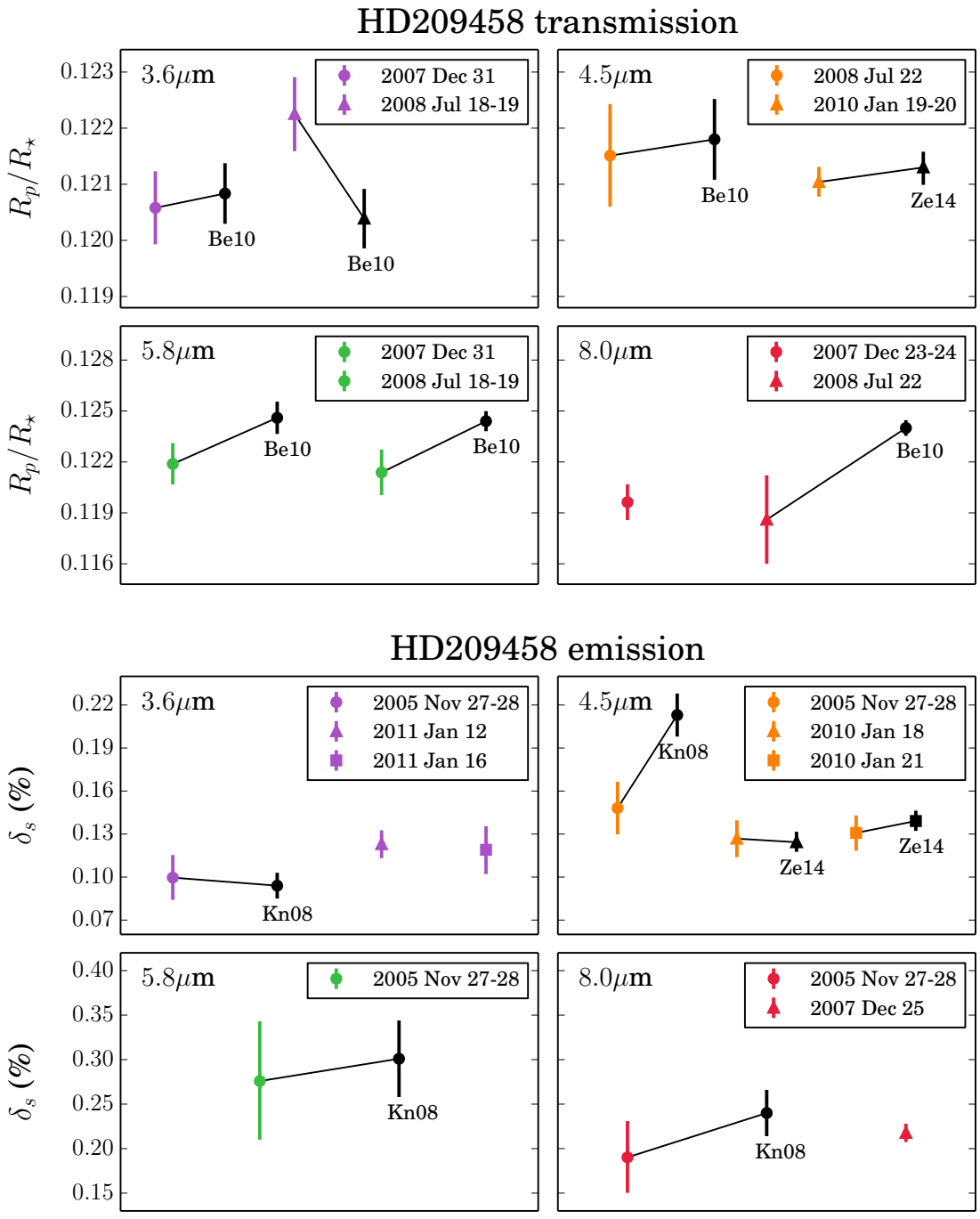

Figure 3.11: Same as Figure 3.8, but for HD 209458b, comparing transmission $R_p/R_\star$ (top) and emission δ_s (bottom) results obtained in the present study with those published in the literature.

somewhat higher than those obtained by Knutson et al. (2012), who found $a/R_\star = 8.879 \pm 0.022$ and $i = 85.701 \pm 0.025$ deg, and Agol et al. (2010), who found $a/R_\star = 8.863 \pm 0.020$ and $i = 85.710 \pm 0.024$ deg. However, my values are consistent if comparisons with the Knutson et al. and Agol et al. values are restricted to the subset of lightcurves analysed both in the current study and by those authors:

- 3.6μm 2010 Dec 29: I obtain $a/R_\star = 8.881^{+0.036}_{-0.036}$ and $i = 85.713^{+0.039}_{-0.040}$ deg, compared with $a/R_\star = 8.858 \pm 0.031$ and $i = 85.671 \pm 0.034$ deg obtained by Knutson et al.

- 4.5μm 2009 Dec 23: I obtain $a/R_\star = 8.912^{+0.035}_{-0.034}$ and $i = 85.758^{+0.038}_{-0.037}$ deg, compared with $a/R_\star = 8.902 \pm 0.032$ and $i = 85.735 \pm 0.036$ deg obtained by Knutson et al.

- 8.0μm 2006 Oct 29: I obtain $a/R_\star = 8.890^{+0.045}_{-0.044}$ and $i = 85.723^{+0.049}_{-0.047}$ deg, compared with $a/R_\star = 8.83 \pm 0.05$ and $i = 85.72 \pm 0.07$ deg obtained by Agol et al.

Therefore, results for the lightcurves that were not analysed by Knutson et al. and Agol et al. are responsible for the different weighted mean values that I obtain for $a/R_\star$ and i.

For HD 209458, my weighted means are $a/R_\star = 8.87 \pm 0.05$ and $i = 86.78 \pm 0.07$ deg. These are both consistent at the $\sim 1\sigma$ level with the values of $a/R_\star = 8.77 \pm 0.07$ and $i = 86.76 \pm 0.10$ deg obtained by Beaulieu et al. (2010), and $a/R_\star = 8.810^{+0.064}_{-0.069}$ and $i = 86.69^{+0.09}_{-0.10}$ deg obtained by Zellem et al. (2014).

Furthermore, almost all of the values obtained in this study for both HD 189733 and HD 209458 are consistent with the weighted means at the $\sim 1\sigma$ level or better, as is illustrated in Figure 3.12. The only notable exception is the single 5.8 μm measurement for HD 189733, where both the $a/R_\star$ and i values differ from the

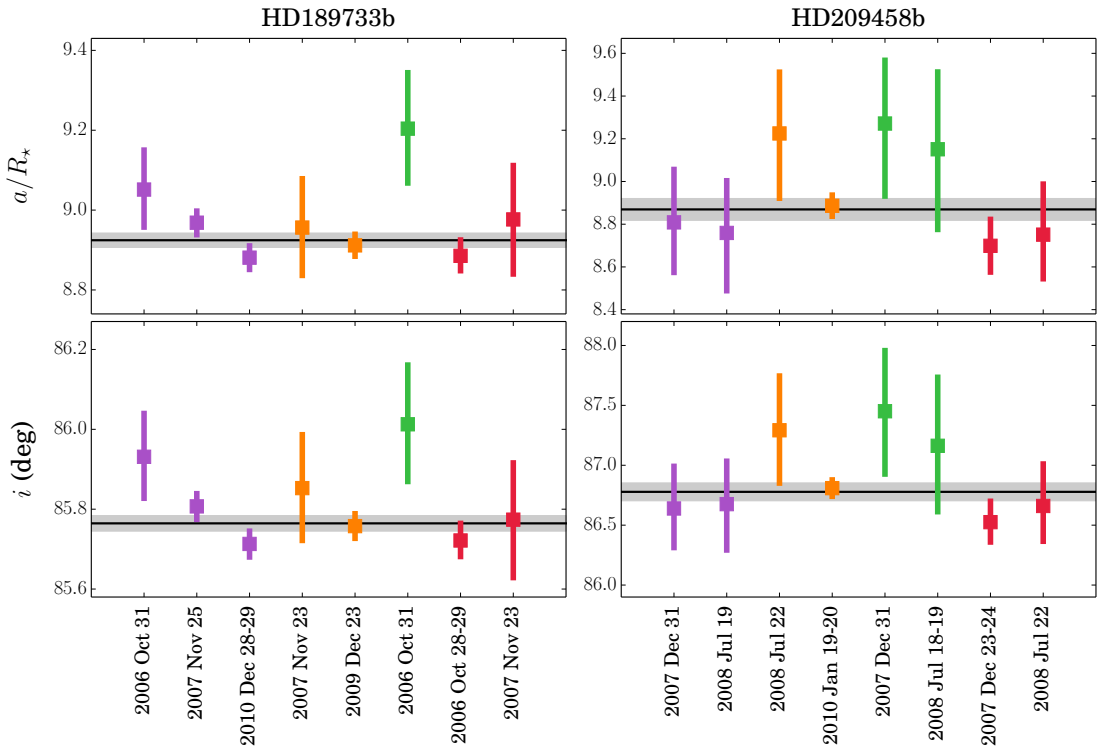

Figure 3.12: Normalised semimajor axis $a/R_\star$ and orbital inclination i values obtained from the MCMC analyses for HD 189733b (left) and HD 209458b (right). Lightcurve epochs are labelled along the horizontal axis, and colours indicate the wavelength channel: 3.6μm (purple), 4.5μm (orange), 5.8μm (green), and 8.0μm (red). Black horizontal lines show the weighted arithmetic means and grey shaded regions show the associated 1σ values.

weighted mean by more like $\sim 2\sigma$. However, my values of $a/R_\star = 9.158^{+0.139}_{-0.148}$ and $b = 0.642^{+0.015}_{-0.014}$ are only 0.6σ and 0.8σ away, respectively, from the corresponding values of $a/R_\star = 9.257 \pm 0.107$ and $b = 0.6279 \pm 0.0111$ determined by Désert et al. (2009) for the same lightcurve. Furthermore, the quoted uncertainties on the weighted means are likely to be underestimated, as they are calculated by combining multiple measurements under the assumption that the errorbars are normally distributed, which is not necessarily true. It is also possible that variations in the star spot coverage between epochs is responsible for a significant amount of the scatter in the measured $a/R_\star$ and i values for HD 189733b. This is supported by the fact that measurements made close in time for HD 189733b are consistent with each other to within 1σ.

Overall, the values obtained for $a/R_\star$ and i in the present study are more consistent across wavelengths and epochs than previously published values. For instance, the $a/R_\star$ and b values obtained by Désert et al. (2009, 2011b) for HD 189733b in the 3.6 μm channel disagree at 3.4σ and 2.6σ, respectively, between the 2006 Oct 31 and 2007 Nov 25 lightcurves. The $a/R_\star$ and b values measured by Désert et al. (2009) for HD 189733b in the 8.0 μm channel also differ from the values they measure in different channels by up to 2.9σ. Meanwhile, for HD 209458b, the $a/R_\star$ and i values reported by Beaulieu et al. (2010) for the 8.0 μm channel differ from most of the values they obtain in the other channels by 3–5σ. The relative consistency of the orbit parameters derived in the current study therefore offers further evidence that the GP modelling approach is doing a better job of accounting for the lightcurve systematics and providing realistic parameter uncertainties.

3.9 Discussion

The results outlined in Sections 3.8.1–3.8.3 demonstrate the effectiveness of the GP modelling approach for handling systematics in IRAC lightcurves. Compared with those previously published in the literature, the planet properties inferred from the GP analyses are overall more consistent across different epochs and, for the wavelength-independent properties, across the different wavelength channels. In a number of cases, this is due to the GP analysis giving uncertainty estimates that are up to ~ 5 times larger than those reported by other authors. However, for the reasons given in Chapter 2, I argue that the GP uncertainties provide a realistic reflection of our ignorance. This is primarily because the GP models offer greater flexibility for handling systematics that do not have a well-understood functional form, compared with the simple parametric approximations used widely in the literature. Marginalisation of the GP model posterior distributions therefore allows us to more exhaustively explore possible degeneracies between the planet signal and systematics, and incorporate these into the uncertainties associated with the inferred planet properties.

In Sections 3.9.1 and 3.9.2, I discuss the implications of my results for the atmospheres of HD 189733b and HD 209458b, respectively.

3.9.1 HD 189733b

To recap the results for HD 189733 given in Section 3.8.1 and illustrated in Figure 3.8, the main contributions of the current study are the following:

- Confirmation of the planet-star radius ratio at 3.6 μm determined by Désert et al. (2009) for the 2006 Oct 31 lightcurve, casting further doubt on the value provided by Beaulieu et al. (2008).

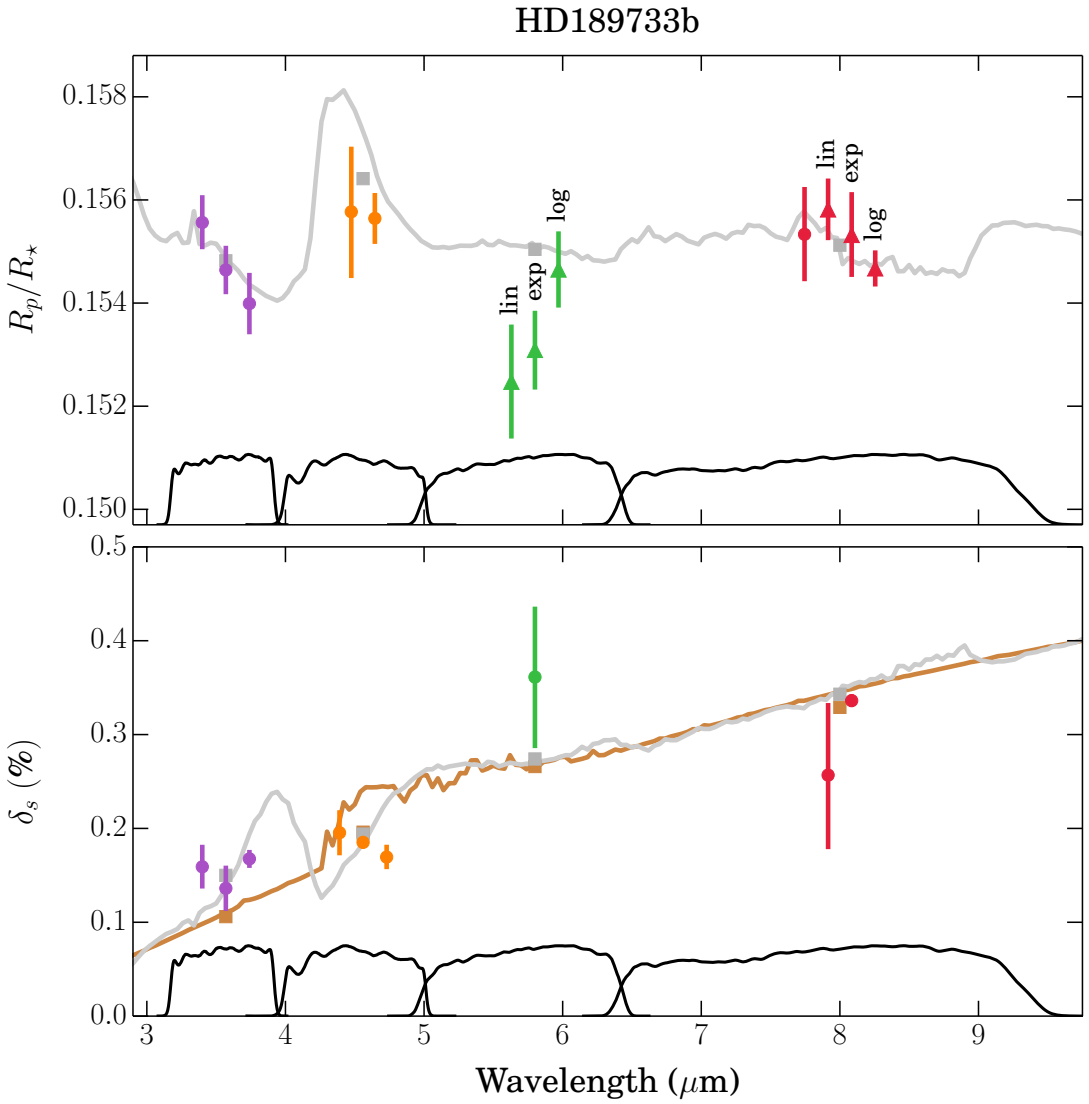

Figure 3.13: Transmission $R_p/R_\star$ (top) and emission δ_s (bottom) IRAC measurements for HD 189733b obtained in this study, with colours indicating the wavelength channels: 3.6μm (purple), 4.5μm (orange), 5.8μm (green), and 8.0μm (red). Values shown for $R_p/R_\star$ are those given in Table 3.6, which have been corrected for variations in the star spot coverage between epochs. Bandpasses are indicated at the bottom of each axis as solid black lines. For measurements obtained in the same wavelength channel, small horizontal offsets have been applied for clarity. In general, circle symbols show the measured transit and eclipse values. The exceptions to this, however, are the two transits shown as triangle symbols: one in the 5.8μm channel and the other in the 8.0μm channel. For these latter cases there are three points per transit, each of which corresponds to the result obtained using one of the three time parameterisations (linear, logarithmic, exponential) for the GP model, as indicated by the labels (see Section 3.7.2 for details). The grey lines in both panels show NEMESIS models produced by J. Barstow (2014, private communication), which are described in Section 3.9.1. The brown line in the bottom panel shows the emission spectrum obtained using the ATLAS stellar model for HD 189733 and assuming the planet radiates as a blackbody with a temperature of 1212 K. Square symbols give model values integrated over photometric bandpasses.

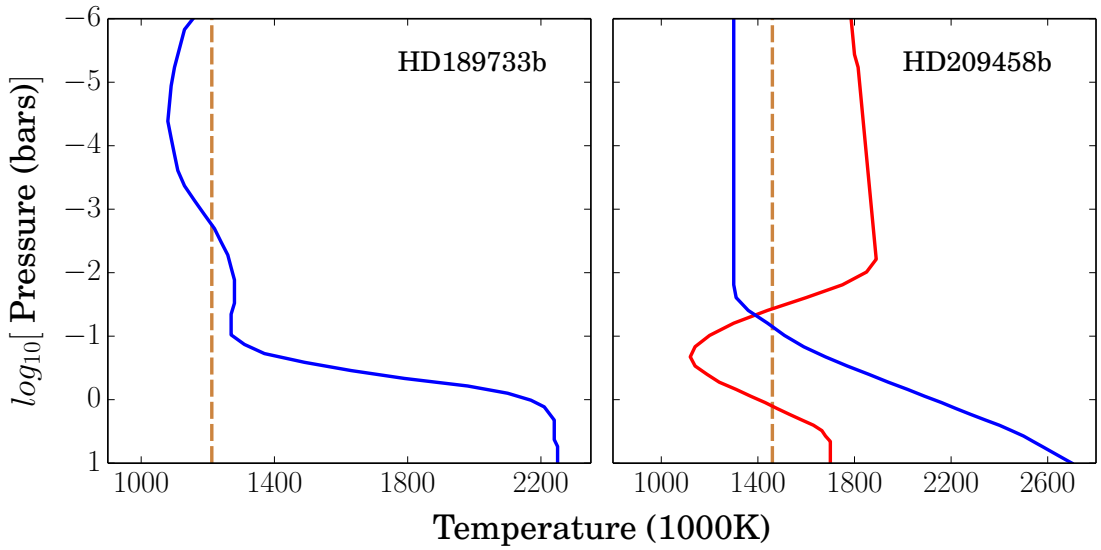

Figure 3.14: Pressure-temperature profiles for HD 189733b (left) and HD 209458b (right) used to generate the models shown in Figures 3.2, 3.3, 3.13, and 3.15. For HD 189733b, the blue line shows the profile derived by Barstow et al. (2014) by performing an atmospheric retrieval on available emission data. For HD 209458b, the blue line shows a non-inverted profile adjusted to give a good fit to the IRAC data presented in this study, and the red line shows the inverted profile derived by Madhusudhan & Seager (2010) by fitting to the IRAC data originally published by Knutson et al. (2008). In both panels, dashed brown lines indicate the temperatures obtained under the assumption that the planets radiate as simple blackbodies, i.e. 1212 K for HD 189733b (Section 3.9.1) and 1460 K for HD 209458b (Section 3.9.2).

- Downward revision of the 3.6 μm eclipse depth for the 2005 Nov 24 lightcurve from the value originally reported by Charbonneau et al. (2008), making it consistent with the values obtained for the 2010 Dec 27 and 30 lightcurves in the current study and by Knutson et al. (2012).

- Downward revision of the 8.0 μm eclipse depth for the 2005 Nov 24 lightcurve from the value originally reported by Charbonneau et al. (2008), which, combined with a larger uncertainty estimate, makes it consistent with the value obtained for the 2006 Oct 30 lightcurve in the current study and by Knutson et al. (2007a) and Agol et al. (2010).

Figure 3.2 shows the IRAC transmission and emission data obtained in this study for HD 189733, along with data from the literature at other wavelengths, while Figure 3.13 shows zoom-ins on the IRAC data points. Also shown in both

Table 3.9: Volume mixing ratios of major gas species for HD 189733b and HD 209458b obtained from best-fit NEMESIS retrievals performed by J. Barstow (2014, private communication). These values should be interpreted as an approximate guide only, as the allowed ranges are likely to span 1–2 orders of magnitude. The abundances derived from the transmission data should be treated with particular caution, as the presence of haze is likely to influence the results (see Section 3.9.1) and is not included in the simple model treatment presented here.

				Mixing ratios (ppmv)			
Target	Spectrum	Dataset [a]	PT profile	H_2O	CO_2	CO	CH_4
HD 189733	transmission	this study	non-inverted	Retrieval not attempted			
	emission	this study	non-inverted	1120	660	258	9
HD 209458	transmission	this study	non-inverted	2	< 1	3	2
	emission	this study	non-inverted	20	10	10	4
	emission	Kn08	inverted	30	< 1	7×10^4	< 1

Notes:

[a] IRAC emission data from the current study was used to constrain the retrievals, along with additional data published at other wavelengths. The only exception was the HD 209458b case with a thermal inversion, for which the IRAC emission measurements of Kn08 (Knutson et al., 2008) were used to constrain the retrieval. See Section 3.9.2 for more details.

figures are model spectra computed by J. Barstow (2014, private communication) using the Oxford NEMESIS code, which was originally written to analyse data for solar system planets (Irwin et al., 2008) and has since been adapted for exoplanets (e.g. Lee et al., 2012; Barstow et al., 2013, 2014). The emission model was generated by first fixing the pressure-temperature (PT) profile to that recovered by the retrieval analysis of Barstow et al. (2014), which is shown in the lefthand panel of Figure 3.14. The NEMESIS code was then used to determine optimal mixing ratios for the four main molecules expected to provide opacity at infrared wavelengths in hot Jupiter atmospheres: H_2O, CO_2, CO, and CH_4. Uniform abundances were assumed for each molecule throughout the atmosphere. In addition to the IRAC eclipse depths presented in the current work, published emission data at other wavelengths were used to constrain the optimisation, most of which are shown in Figure 3.2: namely, a re-analysed *HST*/NICMOS spectrum (Barstow et al., 2014), a *Spitzer*/IRS spectrum (Grillmair et al., 2008), and *Spitzer*/MIPS 24 μm broadband photometry (Charbonneau et al., 2008). Molecular abundances recovered from the retrieval are listed in Table 3.9. No attempt has been made to estimate uncertainties for the abundances, which would require a full retrieval analysis to properly explore the degeneracies of the model parameter space. However, we note that the abundances given in Table 3.9 are consistent with the allowed ranges found by Barstow et al. (2014) at the 1σ level or better for H_2O, CO_2, and CH_4, and at 1.5σ for CO. Meanwhile, a retrieval was not performed for the transmission data due to the complicating effect of haze which is clearly evident at visible wavelengths and perhaps also at longer wavelengths (Figure 3.2). Instead, once the emission model had been determined, the best-fit PT profile and molecular abundances were used to generate a corresponding transmission model without fitting to the data, for the purpose of plotting and to check consistency with the data.

As Figure 3.13 shows, the IRAC transmission and emission spectra are both fit fairly well by these models. Indeed, as can be seen in Figure 3.2, the transmission model provides a reasonable match to all data at wavelengths longward of $2\,\mu$m, consistent with the haze responsible for the enhanced opacity at in the visible range (Section 3.2.1) becoming transparent at these wavelengths. This might be expected, for instance, if the high-altitude haze is dominated by sub-μm grains such that the scattering obeys a λ^{-4} Rayleigh law. In particular, the two $4.5\,\mu$m data points are marginally consistent with an enhanced opacity, which if real, is presumably due to the strong CO absorption feature centered at this wavelength (e.g. Désert et al., 2009). However, the errorbars are too large and the bandpasses too broad to rule out a flat, featureless spectrum for wavelengths $> 1\,\mu$m, as would be expected if the haze dominates the transmission at these wavelengths and mutes the molecular absorption features (e.g. see Pont et al., 2013). The infrared transmission data is therefore inconclusive.

For the emission, although the IRAC and MIPS data are well-explained by the retrieved model, there are large discrepancies for many of the points in the IRS 5–$14\,\mu$m spectrum (Figure 3.2). However, taking the uncertainties at face value, the spectrum exhibits a fine-scale structure that cannot be reproduced by standard atmosphere models. It therefore seems possible that the quoted uncertainties have been underestimated. Similarly, the current study has shown that uncertainties are often significantly underestimated for broadband photometry, which may also help explain the poor fit to the IRS $16\,\mu$m point.

Aside from the IRS points, the NEMESIS model provides a better fit to the emission data than a basic model that treats the planet as an isothermal blackbody, which is also shown in Figures 3.2 and 3.13. To generate the latter, an ATLAS spectrum (Kurucz, 1979, 1993) computed specifically for HD 189733[1] was used

[1] Available at `http://kurucz.harvard.edu/stars/hd189733`

for the stellar emission and the radius ratio was fixed to $R_p/R_\star = 0.155$. Assuming a Planck spectrum for the planet, a temperature of 1212 K was then found to give the best fit to the IRAC data, which is close to the upper atmosphere temperature derived from the NEMESIS retrieval (Figure 3.14). However, the 3.6 μm 2005 Nov 24 and 2010 Dec 30 are clearly inconsistent with this simple isothermal model (Figure 3.13).

The inability of an isothermal model to explain the emission data at 3.6μm, and to a lesser extent at 4.5μm, suggests that different altitudes of the atmosphere are being probed as a function of wavelength. Pont et al. (2013) have suggested that the haze observed in transmission could extend to an optically thick layer across the entire dayside hemisphere. Under this scenario, the haze would scatter most of the incident stellar photons high in the atmosphere, generating a thermal inversion. In addition, the photosphere would be located higher in the atmosphere for shorter wavelengths, assuming that the scattering efficiency of the haze increases with decreasing wavelength as it does for the transmission spectrum. Combined, these two elements – the thermal inversion and the increasing photosphere altitude with decreasing wavelength – would result in the 3.6 μm channel probing a hotter layer of the atmosphere than the 4.6 μm channel. This could perhaps explain why the measured thermal emission in the 3.6 μm channel is higher than expected for the simple isothermal model shown in Figure 3.13.

Alternatively, the haze-free NEMESIS model is capable of reproducing the enhanced emission in the 3.6 μm channel due to an opacity window centered at $\sim 3.5\,\mu$m, with H_2O and CO absorption suppressing the emission from deeper, hotter layers at wavelengths either side (Figure 3.13). Indeed, robust detections of CO and H_2O absorption have recently been made for HD 189733b using high-resolution spectroscopy (de Kok et al., 2013; Birkby et al., 2013). These

measurements appear to disfavour the haze-dominated picture put forward by Pont et al. In addition, Birkby et al. noted that H_2O is likely to be more abundant than the major carbon molecules (CO, CO_2, CH_4) based on the relative strengths of the measured absorption features. This is consistent with the abundances retrieved from the emission data given in Table 3.9.

Nonetheless, there are a number of observations that remain difficult to reconcile with a haze-free atmosphere, which are more easily explained if an optically thick haze is present. Perhaps most significantly, the amplitude of the planetary phase curve measured in the 3.6 μm channel is higher than that measured for the 4.5 μm channel (Knutson et al., 2012). This is puzzling, because the strong CO absorption feature centered on the 4.5 μm channel should result in a higher opacity at these wavelengths relative to the 3.6 μm channel. As such, the 4.5 μm emission should originate from higher in the atmosphere, where radiation timescales are shorter (e.g. Showman et al., 2009). This would allow more heat to be re-radiated back to space at 4.5 μm before it could be circulated to the nightside, giving a higher day-night contrast relative to 3.6 μm, which is the opposite of what is observed. However, phase curve measurements are particularly challenging for HD 189733b, owing to the variability of the host star. There is a significant risk that stellar flux variations on timescales similar to the planet orbital period could confuse the interpretation of the phase curve data, despite best efforts to correct for such effects. An extensive discussion of these issues and others can be found in Pont et al. (2013).

3.9.2 HD 209458b

The main contributions of the current study for HD 209458b as outlined in Section 3.8.2 and illustrated in Figure 3.11 are the following:

- Downward revision of the planet-star radius ratios inferred for the 5.8 μm

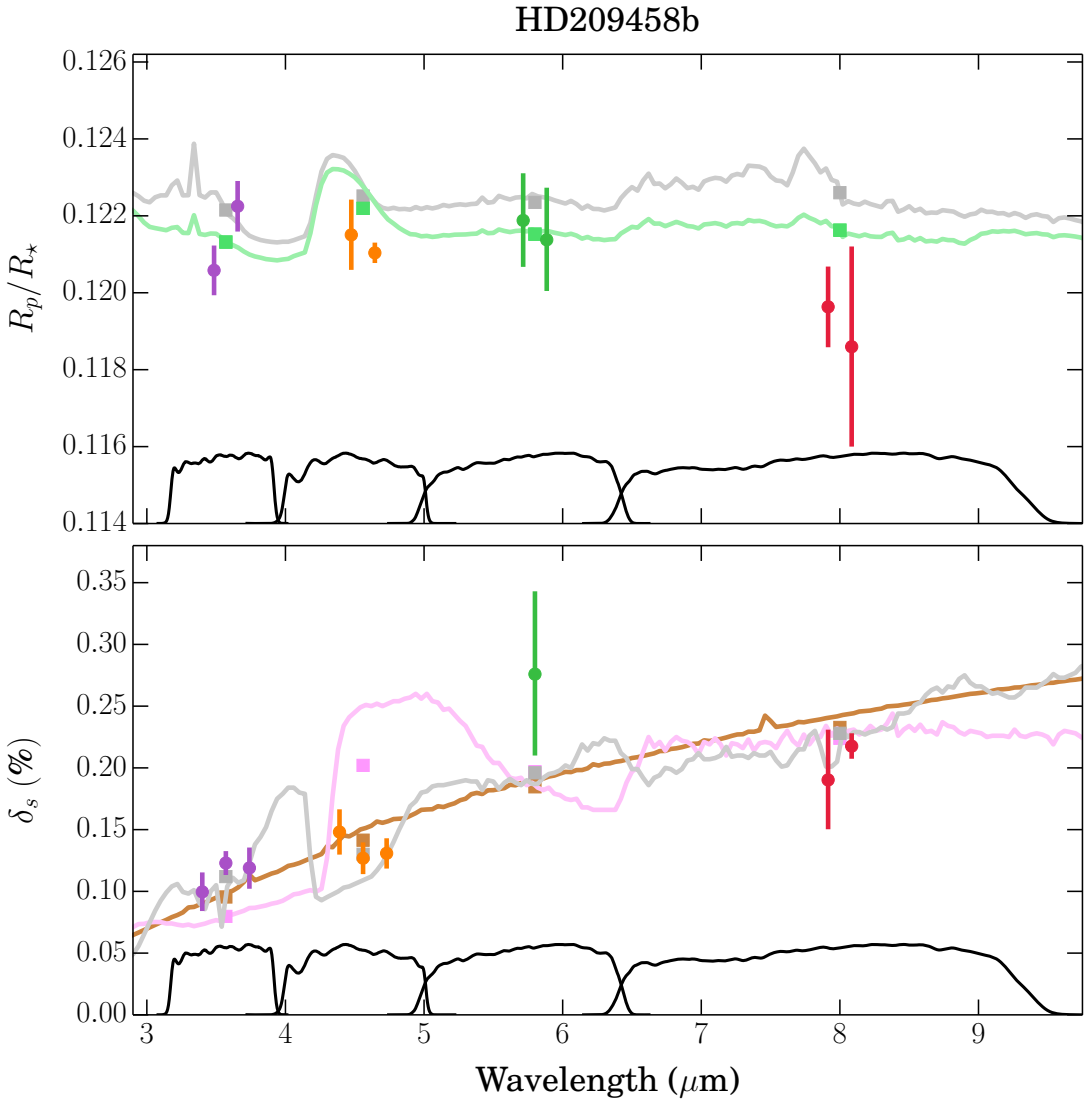

Figure 3.15: Same as Figure 3.13 but for HD 209458b, showing the transmission $R_p/R_\star$ (top) and emission δ_s (bottom) IRAC measurements obtained in this study. Grey lines in both panels show NEMESIS models produced by J. Barstow (2014, private communication), which are described in Section 3.9.2, using the non-inverted PT profile shown in Figure 3.14 and with the retrieved molecular abundances given in Table 3.9. In the top panel, the green line shows the transmission model that has been rescaled to match the amplitude of the WFC3 H_2O absorption feature in order to simulate the effect of a grey opacity source (see Section 3.9.2 for details). In the bottom panel, the brown line shows the emission spectrum obtained using the ATLAS stellar model for HD 209458 and assuming the planet radiates as a blackbody with a temperature of 1460 K, and the pink line shows the best-fit NEMESIS model obtained with the inverted temperature profile. Square symbols give model values integrated over photometric bandpasses.

and 8.0 μm channels from those originally published by Beaulieu et al. (2010) and interpreted as evidence of H_2O absorption.

- Downward revision of the 4.5 μm eclipse depth for the 2005 Nov 28 lightcurve from the value reported by Knutson et al. (2008), making it consistent with both eclipse depths inferred for the 2010 Jan 18 and 21 lightcurves in the current study and by Zellem et al. (2014).

- Fits to previously unpublished 8.0 μm 2007 Dec 24 transit lightcurve, 3.6 μm 2011 Jan 12 and 16 eclipse lightcurves, and 8.0 μm 2007 Dec 25 eclipse lightcurve.

As was noted in Section 3.2, previously published IRAC secondary eclipse measurements for HD 209458b have been interpreted as evidence for a temperature inversion in the upper atmosphere (Knutson et al., 2008). To investigate this claim, models were generated for the emission spectrum, both with and without a temperature inversion. This was done using the NEMESIS code by J. Barstow (2014, private communication) as follows. Firstly, the two pressure-temperature profiles shown in the righthand panel of Figure 3.14 were adopted – one with and one without a temperature inversion. The upper atmosphere temperature of the non-inverted profile was fixed to 1300 K, as this was found to give a good fit to the data presented in the current study, and the inverted pressure-temperature profile was taken directly from by Madhusudhan & Seager (2010). The latter was derived using published emission data, including the IRAC data of Knutson et al. (2008). As with HD 189733b, mixing ratios were then retrieved for the four major molecular species (H_2O, CO, CO_2, CH_4) assuming uniform abundances throughout the atmosphere. The IRAC emission data of the current study, along with the MIPS 24 μm photometric point of Crossfield et al. (2012b), were used

to constrain the retrieval.[2] The model without a temperature inversion, which is shown in Figures 3.3 and 3.15, was found to give a significantly better fit than the model with an inversion, which is also shown in Figure 3.15. This is in contrast to the finding of Knutson et al. (2008), and is primarily due to the revised eclipse depth for the 4.5 μm 2005 Nov 28 lightcurve obtained in the current study, which has been confirmed for the 2010 Jan 18 and 21 lightcurves (Figures 3.11 and 3.15). The retrieved molecular abundances are given in Table 3.9.

As a check, the retrieval was repeated using the eclipse depths of Knutson et al. (2008) and the inverted model was found to gave a much better fit to those data. This model is also shown in Figures 3.2 and 3.13, to emphasise the effect of the revised 4.5μm eclipse depth on the interpretation of the atmosphere. Retrieved molecular abundances are listed in Table 3.9. In particular, note the strong feature in the model spectrum between ~4.5–5.5 μm due to CO and H_2O in emission, which allows the model to fit the relatively high emission value measured by Knutson et al. (2008) for the 4.5 μm channel (Figure 3.11). Meanwhile, it is clearly inconsistent with the three lower values obtained in the current study.

As with HD 189733b, a simple isothermal blackbody model was also fit to the IRAC emission data. Using an ATLAS model computed specifically for HD 209458,[3] a temperature of 1460 K was found to give the best fit for the planet, which is within 200 K of the upper atmosphere temperature of the non-inverted model shown in Figure 3.14. Recall from above, however, that the latter was obtained by Madhusudhan & Seager (2010) by fitting to the original IRAC data of Knutson et al. (2008), so a direct comparison is not strictly appropriate. As can be

[2]Following other retrieval studies (e.g. Madhusudhan & Seager, 2009) we did not include the IRS 8–15 μm data, as two inconsistent spectra have been published (Richardson et al., 2007; Swain et al., 2008a, see Section 3.2), and we were unable to satisfactorily model either of these jointly with the IRAC and MIPS data.

[3]Available at `http://kurucz.harvard.edu/stars/hd209458`

seen in Figures 3.2 and 3.13, the isothermal model provides a quality of fit to the emission data that is comparable to the more complex NEMESIS model, which includes molecular opacity sources and a non-constant pressure-temperature profile. Apart from the IRS spectrum, only the 3.6 μm 2011 Jan 12 data point is significantly discrepant at the level of the uncertainties.[4]

I note that Snellen et al. (2010a) have detected CO absorption lines in the planetary limb of HD 209458b using high-resolution transit spectroscopy, and reported an estimated mixing ratio of 1000–3000 ppmv. This is substantially higher than the value of 10 ppmv retrieved from the broadband emission spectrum in the current study (Table 3.9). Of course, a direct comparison of these two estimates is not possible, as the latter does not have an associated uncertainty due to the preliminary nature of the estimate. Also, the emission probes the composition of the dayside hemisphere, whereas the transmission probes the planetary limb, which in theory could have a different composition. Furthermore, Snellen et al. have since identified a bug in their analysis code, and note that it caused them to overestimate the CO mixing ratio (I. Snellen, 2014, private communication). Although their revised estimate has not yet been made public, it is quite possible that it will be closer to the value obtained in the current study based on the emission data. Alternatively, if the absorption detected by Snellen et al. is restricted to the relatively narrow line cores, it might get washed out when integrated over the broader IRAC bandpasses analysed here. Such a situation could arise, for instance, if there is a haze obscuring the broader wings of the absorption profiles.

The IRAC transmission data presented in the current study may also lend some support to the hazy atmosphere hypothesis (e.g. Charbonneau et al., 2002;

[4]Note that this eclipse was extracted from the 2011 Jan 12–16 full phase observation, which was classed as "failed" by the *Spitzer* Science Cente. However, although the transit was clearly corrupted, the two eclipses were not obviously affected (Figure 3.10).

Deming et al., 2013, see Section 3.2.2). These data are shown in Figures 3.3 and 3.15 along with two illustrative models. The first of these, indicated by the grey line, shows a NEMESIS retrieval assuming the non-inverted temperature-pressure profile (Figure 3.14), constrained by the IRAC transmission measurements of the current study as well as the WFC3 1.0–1.7 μm spectrum (Deming et al., 2013) and MIPS 24 μm photometric point (Crossfield et al., 2012b). As has been reported in other studies (e.g. Deming et al., 2013), a decent fit to the STIS data of Knutson et al. (2007b) was not possible, so they were excluded them from the final retrieval. The recovered molecular abundances are listed in Table 3.9, and are much lower than those inferred from the dayside emission data. For this reason, a second model was generated, shown by the green line, following a similar approach to that taken by Deming et al. (2013). Firstly, the non-inverted pressure-temperature profile was adopted and the molecular mixing ratios were fixed to those recovered from the dayside emission. Then, a transmission model was generated and adjusted to give an optimal fit to the WFC3 data, as these are the most constraining available, clearly mapping out the 1.4 μm water absorption feature. This was done by allowing the transmission model to be shifted and rescaled in the vertical direction. A reduction of the model absorption amplitude by a factor of 0.4 was found to give the best fit to measured water feature.

Both transmission models give good fits to the WFC3 and MIPS data. The rescaled model matches the average level of the STIS data better than the depleted abundance model, but both fail to provide a satisfactory fit at the level of the uncertainties. Overall, the rescaled model gives the best fit to the IRAC data. However, both models overpredict the transit depth for the 4.5 μm 2010 Jan 19 lightcurve by $> 4\sigma$. The lower measured transit depth implies less absorption by CO, which is the major opacity source expected in the 4.5μm channel for a haze-free atmosphere. The other notable ($> 1.5\sigma$) discrepancies

are: the depleted abundance model overpredicts the 3.6 μm 2007 Dec 31 and 8.0 μm 2007 Dec 24 absorption by 2.4σ and 2.8σ, respectively; and the rescaled model also overpredicts the 8.0 μm 2007 Dec 24 absorption by 1.9σ. Therefore, neither of these models provide a particularly satisfactory fit to the IRAC data.

In fact, the IRAC and MIPS data are more consistent with a featureless transmission profile, i.e. a horizontal line. This could indicate, for instance, that the opacity source believed to be responsible for muting the H_2O feature at 1.4 μm has an even stronger effect at longer wavelengths, muting molecular absorption features. The 4.5 μm 2010 Jan 19 lightcurve gives the most support to this picture, as it constrains the CO absorption to have a much lower amplitude than predicted by models without such an opacity source (Figures 3.2 and 3.13).

Meanwhile, the revised IRAC results presented in this study are inconsistent with the H_2O absorption reported by Beaulieu et al. (2010), which was mentioned in Section 3.2.2. Specifically, Beaulieu et al. measured larger effective radii for the planet in the 5.8 μm and 8.0 μm channels, which coincide with an H_2O band, relative to the 3.6 μm and 4.5 μm channels (Figure 3.11). The difference between the results obtained in the current study and those of Beaulieu et al. are likely due to the two approaches used for treating the ramp systematics in these channels (Section 3.6.2). For the 5.8 μm channel, Beaulieu et al. truncated the first section of the lightcurve and fit a linear trend in time to the remainder, and for the 8.0 μm channel they decorrelated the ramp using a quadratic polynomial in logarithmic time. The GP model adopted in the current study should be capable of replicating both these explicit functional forms, and indeed, allows marginalisation over an even broader function space (Section 2.3). I also confirmed that consistent results were obtained with the GP model when sections of varying duration were truncated from the start of the lightcurve. For these reasons, I argue that the revised results presented here are more robust than those originally published by

3.10 Conclusion

This chapter has presented a re-analysis of most of the primary transits and secondary eclipses that have been observed using IRAC for HD 189733b and HD 209458b. By binning the lightcurves in time, it was possible to perform the lightcurve analyses using GP models. A number of significant revisions have been made to previously published results, and in many cases the uncertainties for inferred planet parameters have been increased by factors of $\sim$ 1–5. This can largely be attributed to the flexibility of the GP models, which allow complex correlations to be handled with a small number of free parameters. The latter point is important, as it means that marginalisation over the model parameter space remains computationally tractable, allowing uncertainties that realistically quantify the degeneracies between the planet signal and instrumental systematics to be derived. The revised GP analyses presented here draw into question a number of claims that have previously been made in the literature for these datasets, including the detection of water absorption in the transmission spectra of HD 189733b and HD 209458b, and the inference of a thermal inversion in the atmosphere of HD 209458b.

Chapter 4

The visible-wavelength albedo of the hot Jupiter HD 189733b

The detection of radiation emitted by transiting exoplanets has until now been predominantly carried out at wavelengths longward of $\sim 1\mu$m. This is due to the favourable contrast ratios afforded in the infrared, where thermal fluxes from hot Jupiters peak and the fluxes of FGK dwarf hosts drop significantly. As a result, the change in system brightness that occurs when the planet disappears behind the host star during secondary eclipse is typically ~ 1000 ppm at mid-infrared wavelengths, which is readily detectable with instruments such as *Spitzer*/IRAC. At shorter wavelengths, the spectra of FGK dwarfs peak, while thermal emission from hot Jupiters drops sharply, even for the hottest planets with effective temperatures above 2000 K. Reflected stellar radiation can therefore constitute the bulk of the emission from the planet. However, for a typical hot Jupiter the corresponding eclipse depth will typically be ~ 100 ppm or less, making the signal more difficult to measure than the thermal emission at longer wavelengths.

This chapter presents the first ever multi-wavelength detection of an exoplanet secondary eclipse in reflected light. The measurement was made for HD 189733b using *HST*/STIS across the wavelength range 290–570 nm, and has been published in Evans et al. (2013). Section 4.1 sets the context by providing an overview of exoplanet emission observations at visible wavelengths that have

been published previously. Section 4.2 outlines the motivations for attempting to measure the reflection signal of HD 189733b in particular. Section 4.4 gives details of the STIS observations and data reduction. Section 4.5 presents the lightcurve analysis and Section 4.6 discusses the results.

4.1 Previous measurements

One of the earliest reported attempts to detect reflected stellar radiation from an exoplanet atmosphere was made by Charbonneau et al. (1999), who monitored the non-transiting hot Jupiter τ Boo b with the *Keck Telescope* High Resolution Echelle Spectrometer across the 465–500 nm wavelength range. Using the method outlined in an earlier paper (Charbonneau et al., 1998), the authors sought to detect distortions in the spectral line bisectors caused by the low-amplitude reflection signal of the planet moving in anti-phase with the stellar motion. Although no signal was detected, an upper limit for the geometric albedo of $A_g < 0.5$ was obtained at 99% confidence.[1] Soon afterwards, Collier Cameron et al. (1999) used the same method with the *William Herschel Telescope* Utrecht echelle spectrograph and claimed to make a detection of reflected light from τ Boo b at 4.7σ significance over the 456–524 nm wavelength range. However, this claim was later retracted by the authors, based on follow-up observations that failed to confirm the initial signal (Leigh et al., 2003; Collier-Cameron et al., 2004).

The first truly constraining result for the reflection signal of an exoplanet was made by Rowe et al. (2008), who used the *Microvariability and Oscillations of STars* (*MOST*) satellite to observe a total of 16 secondary eclipses for HD 209458b. No drop in the system flux was detected over the 400–700 nm bandpass, allowing an upper limit of $A_g < 0.17$ to be derived at 3σ confidence. In the same year,

[1] This constraint assumes a value of $R_p = 1.2R_J$ for the unknown planetary radius.

Winn et al. (2008) failed to detect the secondary eclipse of TrES-3b using ground-based photometry, and were able to place upper limits on the albedo of $A_g < 0.30$ and $A_g < 0.60$ in the *Sloan* i (700–850 nm) and z (850–1000 nm) bandpasses, respectively, at 99% confidence.

In addition to measuring the eclipse, exoplanet reflection signals can be measured by monitoring the polarisation of light from the system at different orbital phases. This technique works because scattering by the planetary atmosphere linearly polarises the incident stellar radiation, which is initially unpolarised. Variations in the polarisation signature during the orbit can thus reveal information about the scattering particles, such as their composition and size (e.g. Seager et al., 2000; Madhusudhan & Burrows, 2012).

The first claimed detection of an exoplanet polarisation signature was made by Berdyugina et al. (2008), who observed HD 189733b in the B bandpass (400–490 nm) using a 60 cm ground-based telescope. Assuming a uniform, isotropically scattering atmosphere, the signal measured by the authors implied an albedo of $A_g = 1.13 \pm 0.34$, higher than the theoretical maximum value of $A_g = 2/3$. Wiktorowicz (2009) subsequently published an upper limit for the polarisation signal over the 400–675 nm wavelength range, which appeared to contradict the value obtained by Berdyugina et al., although the bandpasses were not identical. Meanwhile, Lucas et al. (2009) provided upper limits on the polarisation signals for the τ Boo and 55 Cnc systems, and also noted that the signal claimed by Berdyugina et al. for HD 189733b significantly exceeded the maximum value expected for a hot Jupiter atmosphere with multiple scattering. Berdyugina et al. (2011) acquired new data for HD 189733b using the *Nordic Optical Telescope* (*NOT*) TurPol instrument, and revised the B band albedo value to $A_g = 0.69 \pm 0.08$, and also provided measurements of $A_g = 0.70 \pm 0.30$ in the U bandpass (330–

400 nm) and $A_g = 0.28 \pm 0.16$ in the V bandpass (500–600 nm).[2] Berdyugina et al. responded to Lucas et al. by suggesting that the high polarisation signal in the U and B bandpasses could indicate that most photons at these wavelengths were scattered only once, resulting in a higher-amplitude polarisation signal. They also argued that their multi-bandpass measurements were consistent with the upper limit of Wiktorowicz. More polarisation measurements are required, however, to provide a definitive resolution, as there remains significant skepticism within the community.

Important contributions have also been made by the space-based photometers, *CoRoT* and *Kepler*. Using data taken in the *CoRoT* red channel (~560–900 nm), Snellen et al. (2009) were the first to measure the phase curve and secondary eclipse of an exoplanet (CoRoT-1b) at visible wavelengths. Due to the high temperature of CoRoT-1b (~ 2400 K), the detected signal was almost entirely comprised of thermal emission. Secondary eclipses have since been detected in *CoRoT* lightcurves for CoRoT-2b by Alonso et al. (2009) and Snellen et al. (2010b), and for CoRoT-6b, CoRoT-11b, and CoRoT-15b by Parviainen et al. (2013). For these planets, as with CoRoT-1b, the thermal radiation contribution dominates over the reflection signal, owing to their high temperatures (> 1000 K) and the broad *CoRoT* bandpass extending to the near-infrared.

The *Kepler* satellite, which observes in a single bandpass spanning 400–900 nm, has proven particularly useful for measuring secondary eclipses and phase curves of hot Jupiters. Indeed, the first published demonstration of the instrument's photometric performance was a phase curve measured for HAT-P-7b (Borucki et al., 2009), which was subsequently refined by others as more data were accumulated (e.g. Christiansen et al., 2010; Welsh et al., 2010; Jackson

[2]Note that the albedo values quoted here have been adjusted according to the most up-to-date radius values for HD 189733b given in Pont et al. (2013), and as such are slightly different to those in Table 2 of Berdyugina et al. (2011), but consistent within the uncertainties.

et al., 2012; Morris et al., 2013). Kipping & Bakos (2011) detected the secondary eclipse of the hot Jupiter Kepler-7b at 3.5σ confidence, and placed useful upper limits on the eclipse depths for three others (Kepler-5b, Kepler-6b, Kepler-8b). Additional constraints on secondary eclipses and phase curves have been obtained by Désert et al. (2011a), Demory et al. (2011b), Fortney et al. (2011), Kipping & Spiegel (2011), Santerne et al. (2011), Coughlin & López-Morales (2012), Esteves et al. (2013), Heng & Demory (2013), and Quintana et al. (2013). In particular, Heng & Demory (2013) presented a homogeneous analysis for the eleven hot Jupiters that had been confirmed at the time, with the aim of constraining their geometric albedos. For nine of the sample, they found $A_g < 0.2$ at $> 3\sigma$ confidence, with the two exceptions being Kepler-7b ($A_g = 0.352 \pm 0.023$) and HAT-P-7b ($A_g = 0.225 \pm 0.004$).

The low albedos measured for most hot Jupiters to date are broadly consistent with expectations from clear atmosphere models (i.e. without cloud/haze), which predict that absorption by Na and K at visible wavelengths will suppress the reflection signal (Sudarsky et al., 2000; Burrows et al., 2008). Models also predict, however, that silicates (e.g. $MgSiO_3$, $MgSiO_4$), corundum (Al_2O_3), and Fe could condense in the uppermost layers of some atmospheres, potentially raising the albedo significantly (Marley et al., 1999; Sudarsky et al., 2000). Condensates may explain, for instance, the relatively high albedos measured for Kepler-7b and HAT-P-7b. Evidence for atmospheric condensates has also been uncovered in a number of hot Jupiter transmission spectra (e.g. Gibson et al., 2013b; Sing et al., 2013; Deming et al., 2013), including HD 189733b (Pont et al., 2013, and references therein). Reflection measurements for planets such as these will play an important role in teasing out the differences between cloudy and cloud-free hot Jupiters.

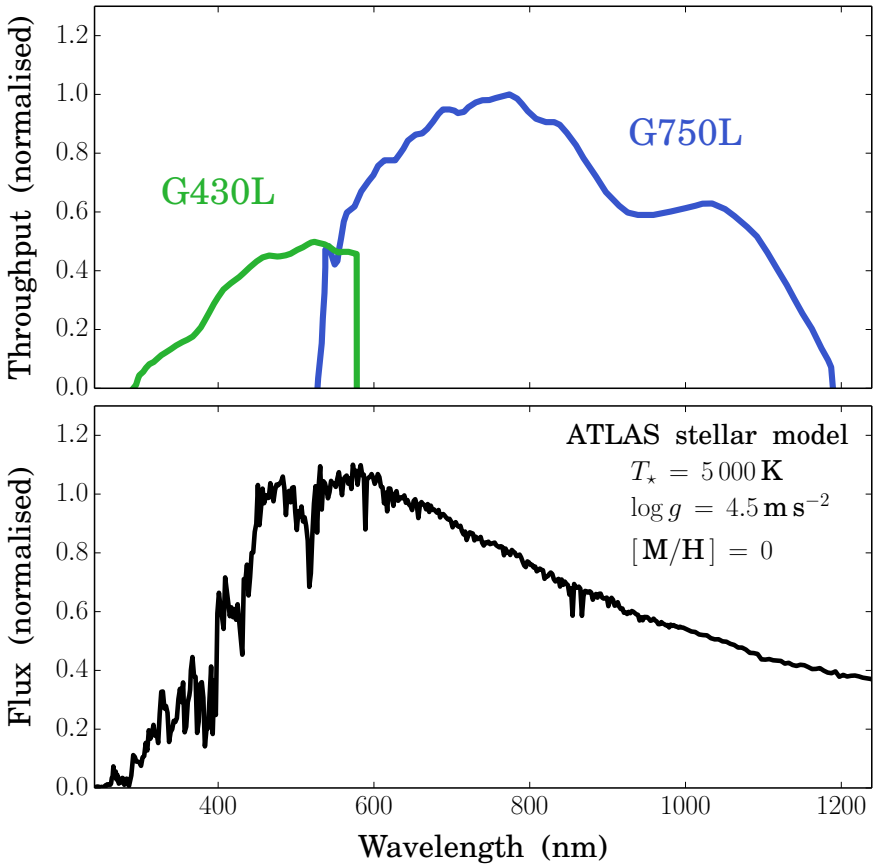

Figure 4.1: (Top) Relative throughputs of the STIS G430L and G750L gratings with arbitrary normalisation. (Bottom) An ATLAS model spectrum (Kurucz, 1979, 1993) computed specifically for HD189733. Approximately 20% of the stellar flux falls within the G430L bandpass, and a further 50% falls within the G750L bandpass.

4.2 Aim of the current measurement

The specific goal of measuring an eclipse for HD 189733b at visible wavelengths was to gauge the role of cloud/haze in the atmosphere. Motivation came from the observed transmission spectrum (Figure 3.2), which slopes downwards from 290 nm out to 1 μm (Pont et al., 2008a; Sing et al., 2011b) and possibly further into the infrared (Sing et al., 2009; Gibson et al., 2012a). As was mentioned in Section 3.2.1, a likely explanation for this feature is Rayleigh scattering by a high altitude haze of dust (Lecavelier Des Etangs et al., 2008). Furthermore, Huitson et al. (2012) detected the narrow core of the Na 589 nm doublet in transmission, but not the pressure-broadened wings that would be expected

in a clear atmosphere (e.g. Fortney et al., 2010). Indirect evidence of cloud on the dayside hemisphere has also been identified by Pont et al. (2013), who speculated that the albedo of HD 189733b might be high as a result.

All of the hot Jupiter emission measurements at visible wavelengths described in the previous section were made over broad photometric bandpasses, which makes the problem of physical interpretation highly degenerate. The most productive instruments, *CoRoT* and *Kepler*, have particularly broad bandpasses that extend to the near-infrared. Instead, we sought to measure the *wavelength-dependent* reflection signal for HD 189733b, with the hope of extracting basic spectral information for the key absorbers and scattering particles in the atmosphere.

One possibility would be to attempt this measurement using the multi-object spectroscopy (MOS) technique that has been used in recent years to measure exoplanet transmission spectra from the ground (see Section 1.4.1.1 and Chapter 5). Although MOS observations for eclipses have been attempted at visible wavelengths,[3] none have yet been published. This might reflect the fact that intrumental systematics due to effects such as time-variable slit losses remain a significant challenge for ground-based MOS observations. Nonetheless, it is an approach that deserves serious consideration going forward, especially as increasing numbers of favourable targets are discovered with suitable nearby comparison stars.

The alternative is to make the observation from a space-based platform, to avoid problems caused by the distorting effect of the Earth's atmosphere. Currently, *HST*/STIS is the only such instrument with spectroscopic capabilities in the visible, and has become established as the premiere instrument for measur-

[3]For example, visible wavelength eclipse observations have been attempted with the *Gemini*/GMOS instrument for the hot Jupiters WASP-12b, WASP-19b, Kepler-13b, and WASP-33b as part of Programs GN-2011B-Q-46 (PI: de Mooij), GN-2013B-Q-52 (PI: de Mooij), and GS-2012A-Q-24 (PI: de Mooij) over the course of 2011-2013.

ing hot Jupiter transmission spectra over the 0.3–1μm wavelength range (e.g. Sing et al., 2008a, 2011b, 2013; Huitson et al., 2012; Nikolov et al., 2014). Encouraged by this success, we applied to use STIS to observe a secondary eclipse of HD 189733b with the G430L grating, which covers the 290–570 nm wavelength range (Figure 4.1), where the transmission spectrum implies the scattering signature of the putative cloud/haze is strongest. Substituting values of $R_p/R_\star = 0.1555$ (Pont et al., 2013) and $a/R_\star = 8.86$ (Agol et al., 2010) into Equation 1.4, the expected eclipse depth is $\delta_s = A_g \times 310$ ppm. Based on our groups' previous STIS observations of the transmission spectrum (Sing et al., 2011b), the anticipated level of precision for the system flux measurement was ~ 30 ppm. Therefore, an isotropically, non-absorbing atmosphere with an albedo of $A_g = 2/3$ would produce a $\sim 7\sigma$ detection. Alternatively, if the atmosphere has a low albedo $A_g \approx 0$, then an upper limit of $A_g < 0.3$ could be obtained at 3σ confidence. Our proposal was successful, and we were awarded time to observe a secondary eclipse for HD 189733b as part of Program GO-13006 with Frédéric Pont as Principal Investigator.

4.3 Instrument description

The *Hubble Space Telescope* (*HST*) is located in a low-Earth orbit at an altitude of 570 km, with an orbital period of 96 min. Most targets, including HD 189733b, are occulted by the Earth for approximately half of each orbit. The mirror diameter is 2.4 m, and currently there are five operational instruments providing photometric and spectroscopic capabilities across the ultraviolet to near-infrared wavelength range: Wide Field Camera 3 (WFC3), the Cosmic Origins Spectrograph (COS), the Advanced Camera for Surveys (ACS), Near Infrared Camera and Multi-Object Spectrometer (NICMOS), and the Space Telescope Imaging Spectrograph (STIS).

STIS is a single slit spectrograph with a coverage across the 114–1027 nm wavelength range. First order gratings for long slit spectroscopy provide coverage across this full range, with echelle gratings also available at ultraviolet wavelengths (114–315 nm). An additional spectral mode is provided by an objective prism, which disperses light at near-ultraviolet wavelengths (115–362 nm). The detectors consist of two multianode microchannel arrays for ultraviolet observations, and a CCD for visible to near-infrared observations. Further details of the STIS instrument design and operating modes can be found in Woodgate et al. (1998) and the instrument handbook.[4]

4.4 Observations

The observations described in this section were designed and scheduled by Frédéric Pont.

One secondary eclipse of HD 189733b was observed over four consecutive *HST* orbits with the G430L grating on UT 2012 December 20. Spectra taken during the first *HST* orbit exhibited much larger systematics than those taken in subsequent orbits due to settling of the telescope into its new pointing position. This is a well-known phenomenon (e.g. Brown et al., 2001; Sing et al., 2011b), and all data taken during this settling orbit were discarded from further analysis. Of the remaining three orbits, the first and third sampled the out-of-eclipse flux while the planet was close to full phase (star plus the dayside hemisphere of the planet), and the second sampled the in-eclipse flux while the planet was fully obscured by the star (star only).

We used the widest available slit ($52'' \times 2''$) to minimise time-varying slit losses caused by pointing drifts and reduced overheads by only reading out the 1024×128 pixel subarray containing the target spectrum. Based on previous experience

[4] Available at `http://www.stsci.edu/hst/stis`

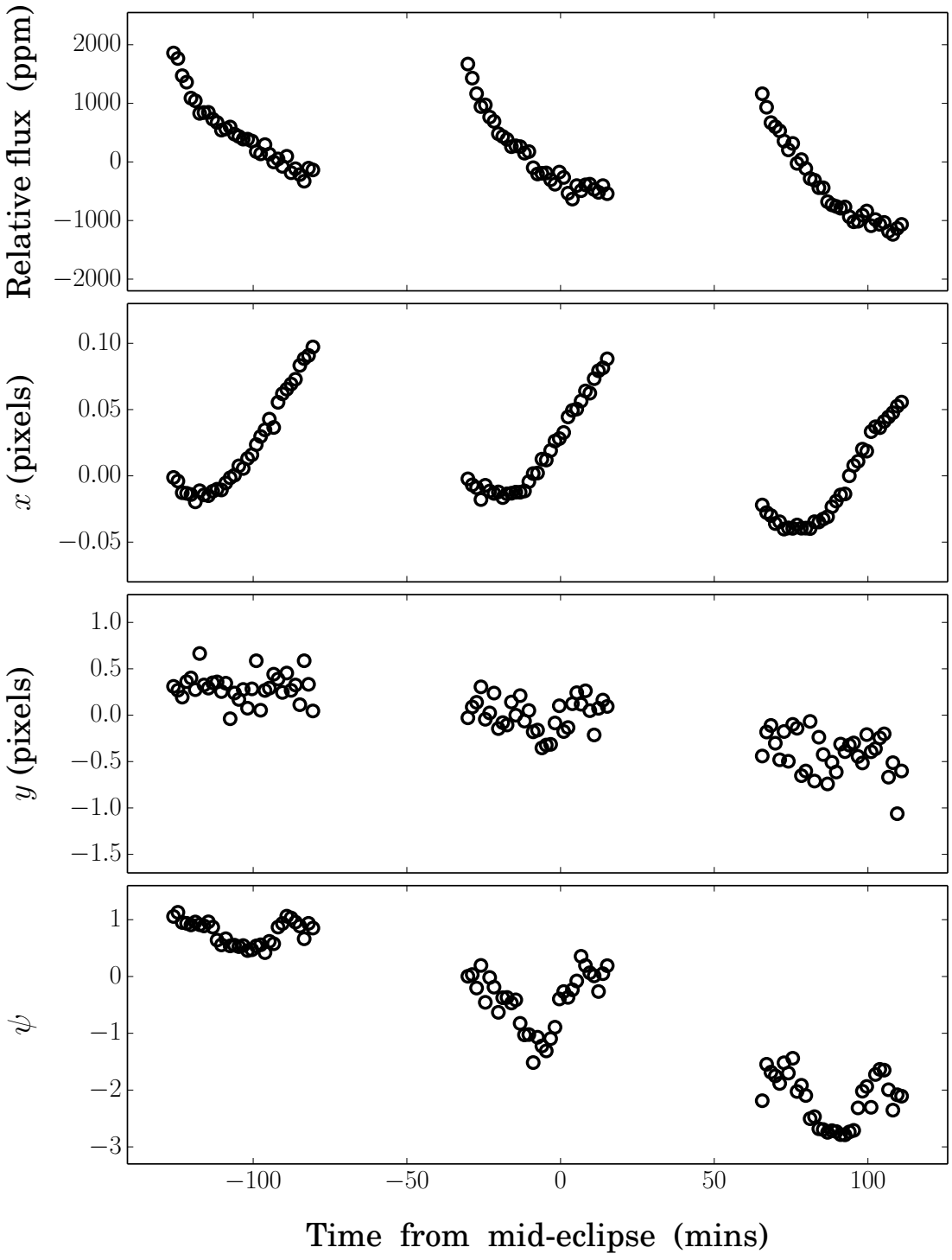

Figure 4.2: Time series for variables extracted from the HD 189733b STIS dataset. Rows from top to bottom are: (1) median-normalised stellar flux integrated over the dispersion axis; (2) dispersion shifts x; (3) cross-dispersion shifts y; (4) spectral trace slope ψ. All variables, except for ψ have been median-subtracted. Note that although ψ is strictly speaking dimensionless (see Equation 4.1), it is quoted here as the number of pixels by which the trace cross-dispersion centroid changes over the full length of the 1024 pixel dispersion axis.

with STIS observations, we expected the first exposure of each *HST* orbit to have systematically lower counts than the exposures immediately following. For this reason, a dummy 1 sec exposure was taken at the beginning of each orbit, followed by a series of 35 science exposures with integration times of 64 sec. Unfortunately, the first science exposure of each *HST* orbit still exhibited systematically lower flux levels, so we discarded these as well. Our final dataset thus consisted of 102 spectra taken over 237 min.

4.4.1 Data reduction

The data reduction described in this section was performed by David Sing.

The CALSTIS v2.40 pipeline[5] was used to reduce the images, clean cosmic rays, and obtain the instrument wavelength solution. Spectra were extracted using the IRAF (Tody, 1986) `apall` routine. This involved fitting the spectral trace in each frame with a linear model of the form:

$$\hat{y} \;=\; y + \frac{\psi}{1024}\,\hat{x}\,, \tag{4.1}$$

where $\hat{y}$ denotes the cross-dispersion coordinate, y is an offset along the cross-dispersion axis, ψ is the trace slope,[6] and $\hat{x}$ is the dispersion coordinate. Fluxes were then obtained by summing pixels within a 13 pixel-wide aperture centered on the trace solution. Background subtraction was not performed, as the background contribution was negligible. Spectra were Doppler-corrected to the heliocentric rest frame, corresponding to shifts of ~1 pixel along the dispersion axis. In addition, frame-to-frame shifts of the spectral trace along the dispersion axis x were computed using the cross-correlation of each individual spectrum with the median. Figure 4.2 shows the resulting time series for the wavelength-integrated flux, along with the auxiliary variables x, y, and ψ. Gaps in the time

[5] `http://www.stsci.edu/hst/stis/software/analyzing`

[6] Normalised by 1024, ψ corresponds to the number of pixels that the cross-dispersion centroid changes by across the full length of the 1024 pixel dispersion axis.

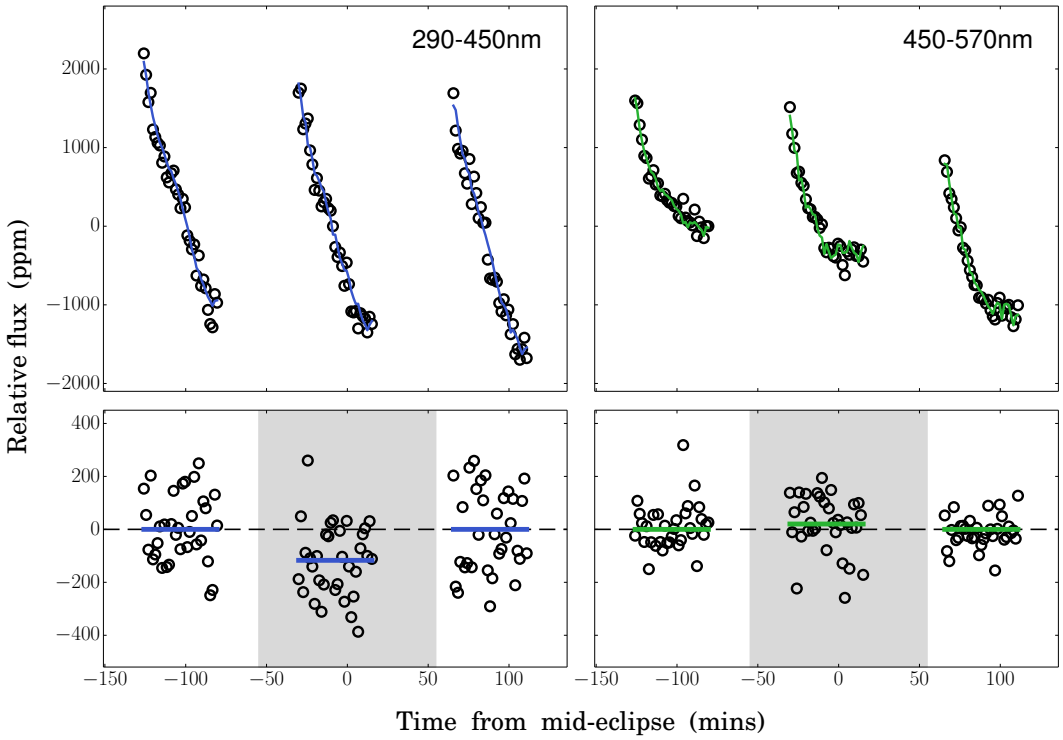

Figure 4.3: (Top panel) Raw lightcurve binned into 290–450nm and 450–570 nm wavelength channels, in units of relative flux about the median. Solid lines show the best-fit GP mean functions that incorporate both the planet and systematics signals. (Bottom panel) The corrected lightcurves, obtained by dividing the raw lightcurves by the non-planetary component the best-fit GP mean function, leaving only the planet signal. Times between start-of-ingress and end-of-egress are shaded grey. Solid lines show the MLE boxcar fit for the planet signal.

series correspond to periods during the *HST* orbit when the target was occulted by the Earth.

4.5 Lightcurve analysis

Broadband spectral lightcurves were generated by summing the flux from each individual frame across the dispersion axis within different wavelength channels. Separate analyses were performed for two-channel and six-channel binnings, with bin widths of ~ 160 nm and ~ 50 nm, respectively. Raw lightcurves for the two-channel binning are shown in the top panel of Figure 4.3, and for the six-channel binning in Figure 4.4. For all channels, the measured flux is dominated by an approximately repeatable decrease of ~ 1500–3000 ppm over each *HST*

orbit. Smaller-amplitude correlations on shorter time scales are also evident, as well as a more gradual decrease in the baseline flux level. These systematics are believed to be primarily caused by the thermal cycle of the *HST* and the drift of the spectral trace across the detector (eg. Brown et al., 2001; Sing et al., 2011b; Huitson et al., 2012).

Before proceeding to the final Gaussian process (GP) lightcurve fitting, a number of tests were performed using linear basis models, which are commonly used to decorrelate STIS lightcurves (e.g. Sing et al., 2011b; Huitson et al., 2012). The purpose of these tests, which are reported in Appendix C, was to get a feel for which input variables were correlated with the measured flux, and to evaluate the reliability of the inferred eclipse depths. In summary, the key results are:

- the measured flux is most strongly correlated with the *HST* orbital phase ϕ and the slope of the spectral trace on the detector ψ;
- the inclusion of a linear trend in time t seems to systematically bias the inferred eclipse depths δ_s to lower values;
- consistent eclipse depths are recovered when different subsets of the data are analysed.

Following these tests, the lightcurves were modelled as GPs using a similar approach to that outlined for the *Spitzer*/IRAC lightcurve analyses in Chapter 3. The model likelihood thus took the form $p(\boldsymbol{d}|\boldsymbol{\alpha},\boldsymbol{\gamma}) = \mathcal{N}\left(\boldsymbol{\mu}, \mathbf{K}+\boldsymbol{\Sigma}\right)$ (i.e. Equation 3.7). Since the ingress and egress were not sampled, a simple boxcar signal multiplied by a linear trend in time was used for the mean function, such that:

$$\mu(t, \boldsymbol{\alpha}) \quad = \quad \left(f_0 + g\,t\right)\left[1 - \delta_s B(t)\right] \ , \tag{4.2}$$

where f_0 is the constant baseline flux level, δ_s is the fractional flux change during

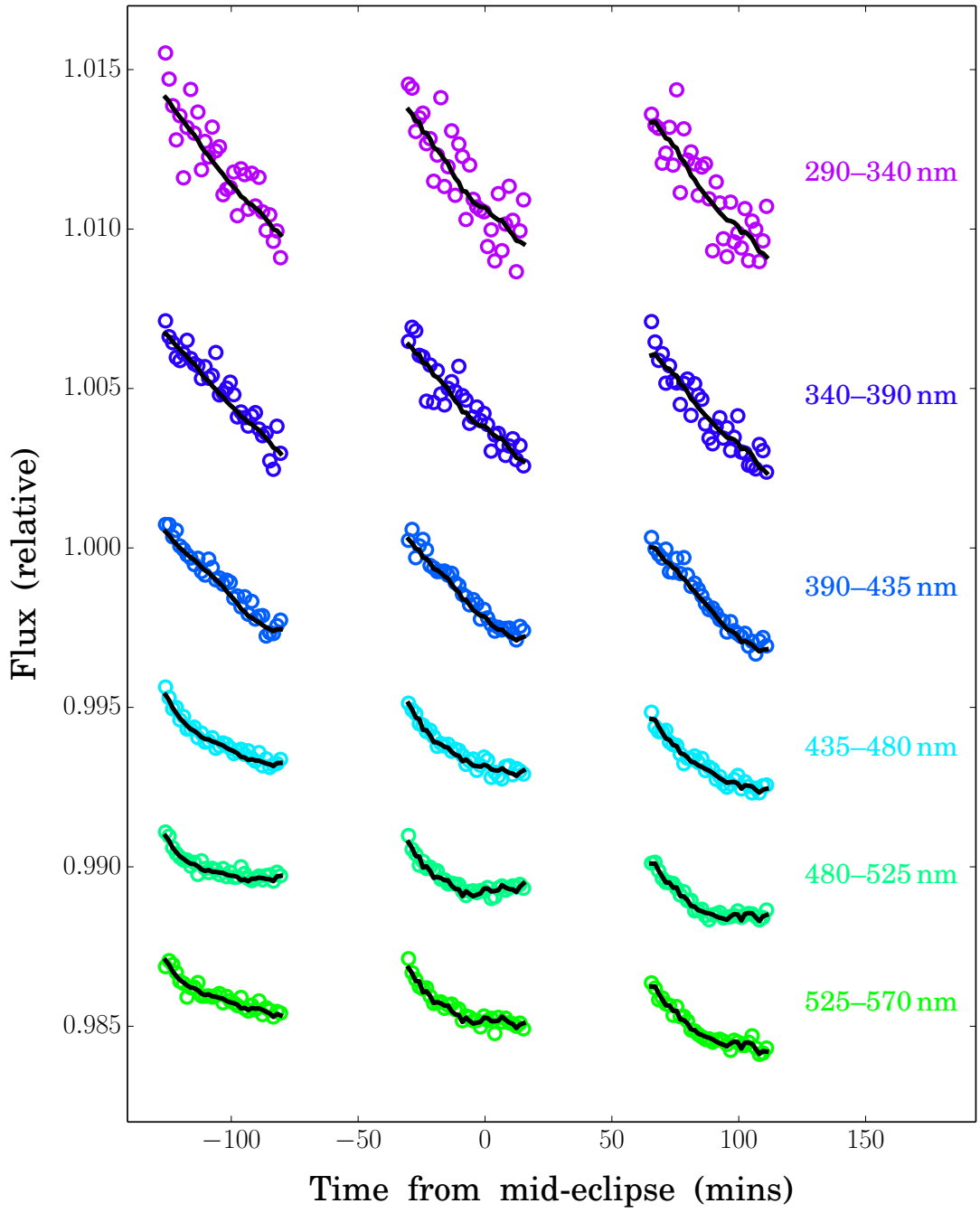

Figure 4.4: Raw lightcurve binned into six wavelength channels across the 290–570 nm wavelength range. Solid black lines show the best-fit GP mean functions, which incorporate both the planet and systematics signals.

eclipse, and:

$$B(t) = \begin{cases} 0 & \text{for } t \text{ during the 1st and 3rd } \textit{HST} \text{ orbits ,} \\ 1 & \text{for } t \text{ during the 2nd } \textit{HST} \text{ orbit .} \end{cases} \tag{4.3}$$

A Matérn $\nu = 3/2$ kernel function (Equation 2.29) was used to populate the entries of the covariance matrix $\mathbf{K}$.[7] To start with, ϕ, ψ, x, and y were used as input variables to the covariance kernel, each with a corresponding length scale. The model therefore had three mean function parameters $\boldsymbol{\alpha} = \{\delta_s, f_0, g\}$ and six covariance parameters $\boldsymbol{\gamma} = \{\sigma_w, A, L_\phi, L_\psi, L_x, L_y\}$, where σ_w was the white noise, A was the covariance amplitude, and the L_i terms were the correlation length scales. Uniform priors were assumed for all parameters, with the exception of f_0 which was constrained to be close to 1 by a normal prior $p(f_0) = \mathcal{N}(1, 0.01^2)$. The eclipse depth δ_s was allowed to assume positive and negative values. The Nelder-Mead simplex algorithm used to find the combination of parameters that optimised the posterior likelihood for each wavelength channel separately. This process was repeated for multiple (~ 20) well-dispersed starting points that were chosen randomly to reduce the risk of the optimiser getting caught in a local maximum. The highest-likelihood solutions are listed in Table 4.1 for each wavelength channel.

I found, however, that as with the linear decorrelations reported in Appendix C, the inclusion of a linear trend in time seemed to bias the inferred eclipse depths to lower values than those obtained when no linear time trend was used. For example, the best-fit eclipse depth for the 290–450 nm channel is $\delta_s = -87$ ppm. I also found that the same behaviour occurred when time was included as an input variable to the GP instead of using a linear trend. For this reason, the MLE analyses were repeated without the linear trend in time (i.e. setting $g = 0$) and

[7] The analysis was also repeated using a squared exponential kernel function (Equation 2.27), and gave near-identical results to those obtained with the Matérn $\nu = 3/2$ kernel.

Table 4.1: MLE values for GP model with ϕ, ψ, x, and y as inputs to a Matérn $\nu = 3/2$ kernel plus a linear trend in t. Note that values for the length scales L_i are given in units of the range Δ_i spanned by the corresponding input variable, e.g. $L_x = 50$ implies the length scale is 50 times the range spanned by x. See Figure 4.2 for the ranges spanned by each input variable.

Parameter	f_0	g	δ_s (ppm)	σ_w (ppm)	A (ppm)	L_ϕ (Δ_ϕ)	L_ψ (Δ_ψ)	L_x (Δ_x)	L_y (Δ_y)
290–450	1.016483	−0.002307	93	146	22795	2.06	50.58	2017.50	99.90
450–570	0.997969	−0.007600	−87	83	5206	1.48	$> 10^5$	44.24	0.72
290–340	1.000272	−0.001977	−22	789	3859	0.78	6.12	$> 10^5$	$> 10^5$
340–390	1.003751	−0.002320	−1	399	10872	2.48	20.97	37.29	23.59
390–435	1.009476	−0.005062	62	205	15767	3.61	$> 10^5$	$> 10^5$	$> 10^5$
435–480	1.000166	−0.010953	−28	139	1261	0.73	1.63	$> 10^5$	58.27
480–525	0.985160	−0.007444	−125	113	11208	2.12	1.15	1.57	32.40
525–570	1.004451	−0.000575	−3	118	8164	1.77	1.78	1.88	$> 10^5$

Table 4.2: Same as Table 4.1, but for GP model with ϕ, ψ, x, and y as inputs.

Parameter	f_0	δ_s (ppm)	σ_w (ppm)	A (ppm)	L_ϕ (Δ_ϕ)	L_ψ (Δ_ψ)	L_x (Δ_x)	L_y (Δ_y)
290–450	6418	113	144	7886	1.17	3.22	$> 10^5$	$> 10^5$
450–570	1996	40	87	7779	1.31	4.44	3.60	$> 10^5$
290–340	−1745	45	766	4597	2.32	5.42	$> 10^5$	$> 10^5$
340–390	5860	151	453	6937	3.74	4.18	$> 10^5$	$> 10^5$
390–435	1259	174	189	4163	1.06	2.38	1.08	$> 10^5$
435–480	−489	58	148	5443	0.99	3.40	0.50	$> 10^5$
480–525	14704	10	112	8793	1.50	5.51	3.00	$> 10^5$
525–570	10495	−12	127	8757	1.51	0.26	$> 10^5$	$> 10^5$

Table 4.3: Same as Table 4.1, but for GP model with ϕ, ψ, and x as inputs.

Parameter	f_0	δ_s (ppm)	σ_w (ppm)	A (ppm)	L_ϕ (Δ_ϕ)	L_ψ (Δ_ψ)	L_x (Δ_x)
290–450	1.007756	120	148	10959	0.72	5.18	$> 10^5$
450–570	1.008292	26	86	15179	0.95	61.56	1.79
290–340	1.006106	139	781	11045	2.24	33.55	$> 10^5$
340–390	1.005437	137	430	10285	2.64	23.85	$> 10^5$
390–435	1.004384	100	198	5739	0.68	22.50	$> 10^5$
435–480	1.006713	53	143	10759	0.77	4.25	$> 10^5$
480–525	1.008635	14	114	16054	0.99	10.14	0.44
525–570	1.008911	9	126	17670	1.02	37.44	$> 10^5$

the results are reported in Table 4.2. As expected, the eclipse depths are shifted to higher values across all channels.

In terms of the other input variables, very large length scales are obtained for the y input variable; specifically, $L_y > 10^5 \Delta y$, where $\Delta y = y_{\max} - y_{\min}$ was the range of y values spanned over the course of the observations (~ 1 pixel; see Figure 4.2). These large length scales imply that y is not needed to explain the correlations in the dataset. Discarding y from the input variables, I again repeated the analyses and report the results in Table 4.3. For the majority of cases, I found that x also had a very large length scale ($L_x > 10^5 \Delta x$) and could be discarded as an input variable. The only exceptions were the 450–570 nm and 480–525 nm wavelength channels, for which smaller L_x values were obtained.

These results indicate that apart from the dominant phase-dependent (ϕ) systematic, the measured fluxes were most strongly correlated with the trace angle on the detector (ψ), and to a lesser extent, the jitter in the trace position along the dispersion axis (x). The correlations with ψ and x seem perfectly reasonable, as their values determine the amount of flux that falls within a given

Table 4.4: Same as Table 4.1, but for GP model with ϕ and ψ as inputs.

Parameter	f_0	δ_s (ppm)	σ_w (ppm)	A (ppm)	L_ϕ (Δ_ϕ)	L_ψ (Δ_ψ)
290–450	1.005406	116	145	6951	0.55	33.57
450–570	1.007431	−20	91	10757	0.76	49.74
290–340	1.006315	150	782	10757	1.92	47.38
340–390	1.005702	139	432	11721	2.64	59.73
390–435	1.005021	89	194	6826	0.63	41.93
435–480	1.006542	43	143	8068	0.64	37.14
480–525	1.007823	−53	119	11547	0.77	51.67
525–570	1.007054	−8	127	11026	0.80	51.49

wavelength channel. The lack of correlation with the cross-dispersion jitter (y) is also unsurprising, as the flux for each wavelength channel is summed along the cross-dispersion axis within the spectral aperture. An inspection of Figure 4.2 suggests that the overall downward drift in the ψ variable correlates with the similar trend observed for the measured flux. As an aside, I suspect that the linear trend in time t discussed above was attempting to model this downward drift in the flux level, and the fact that it can be explained by a correlation with the physically-interpretable ψ instead provides justification for excluding t as a decorrelation variable.

Given that large length scales L_x were recovered for most channels, the GP analyses were also repeated without x as an input variable, using only ϕ and ψ. The results are reported in Table 4.4. The MLE models are plotted for each wavelength channel in Figures 4.3 and 4.4. Specifically, MLE models using $\{\phi, \psi, x\}$ as inputs are shown for the 450–570 nm and 480–525 nm wavelength channels, while MLE models using $\{\phi, \psi\}$ are shown for the rest.

As with the IRAC lightcurve analyses described in Section 3.7.4, the posterior

distributions were marginalised using MCMC with Metropolis-Hastings sampling. This was done separately for two cases: one with $\{\phi, \psi\}$ and the other with $\{\phi, \psi, x\}$ as GP inputs. To assist convergence of the MCMC chains, steps were taken in *inverse* length scales (i.e. $1/L_\phi$, $1/L_\psi$, $1/L_x$), which were given priors of the form:

$$p(1/L_i) \quad = \quad \exp\left[-10\Delta_i\left(\frac{1}{L_i}\right)\right] , \tag{4.4}$$

where $\Delta_i = i_{\max} - i_{\min}$ was the range spanned by the variable. Equation 4.4 defines a Gamma distribution over $1/L_i$ with a shape parameter of 1 and a rate parameter of $10\Delta_i$, i.e. $p(1/L_i) = \mathrm{Gam}(1, 10\Delta_i)$, preferencing longer length scales. Different priors were tested, including uniform distributions and gamma distributions with different rate parameters. Consistent results were obtained in all cases. However, the use of gamma priors of the form given by Equation 4.4 allowed the MCMC chains to converge in fewer steps. Similarly, a Gamma prior of the form $p(A) = \mathrm{Gam}(1, 10)$ was used for the covariance amplitude, in order to preference smaller covariance amplitudes and aid convergence of the MCMC chains. For the baseline flux level f_0, I adopted the same normal prior as was used for the MLE analysis (see above), and uniform priors for the remaining parameters.

Step sizes were pretuned to give an acceptance rate of 20–40 % using the same procedure described in Section 3.7.4 for the IRAC lightcurve analyses. Five independent chains of 1.5×10^5 steps each were initiated at well-dispersed locations $\sim 5\sigma$ from the MLE solution. The first 5×10^4 steps were removed from each chain as burn-in and the Gelman-Rubin values were found to be within 1% of unity for each parameter, indicating good mixing (Section 2.2.1). The five chains were then combined into a single chain of 50×10^4 samples from the posterior distribution.

Table 4.5: MCMC output for GP model with ϕ, ψ, and x as inputs. Quoted values are the chain medians, and uncertainties give the ranges either side of the medians that encompass 34% of the chain samples. As with Tables 4.1–4.4, values for the length scales L_i are given in units of the range Δ_i spanned by the corresponding input variable.

Channel (nm)	f_0 (ppm)	σ_w (ppm)	A (ppm)	L_ϕ (Δ_ϕ)	L_ψ (Δ_ψ)	L_x (Δ_x)	δ_s (ppm)
290–450	$1.005336^{+0.006926}_{-0.005206}$	146^{+13}_{-11}	8343^{+5923}_{-3500}	$0.60^{+1.89}_{-2.56}$	$29.41^{+41.08}_{-62.66}$	$1.43^{+0.93}_{-1.95}$	126^{+49}_{-51}
450–570	$1.005079^{+0.007127}_{-0.005755}$	87^{+7}_{-7}	9506^{+6529}_{-3783}	$0.76^{+2.64}_{-3.39}$	$29.05^{+44.58}_{-67.91}$	$0.83^{+1.07}_{-1.65}$	40^{+42}_{-37}
290–340	$1.005248^{+0.007764}_{-0.006429}$	795^{+62}_{-55}	10577^{+7041}_{-4258}	$1.56^{+2.48}_{-3.80}$	$32.14^{+35.71}_{-62.86}$	$1.82^{+1.13}_{-2.46}$	139^{+185}_{-183}
340–390	$1.004689^{+0.007526}_{-0.006204}$	428^{+36}_{-33}	10295^{+7348}_{-4253}	$1.27^{+1.19}_{-2.42}$	$36.69^{+41.04}_{-72.69}$	$1.79^{+1.03}_{-2.38}$	150^{+107}_{-108}
390–435	$1.004480^{+0.006568}_{-0.004911}$	197^{+17}_{-15}	7491^{+5513}_{-3181}	$0.66^{+1.84}_{-2.39}$	$31.49^{+38.82}_{-63.80}$	$1.78^{+1.09}_{-2.42}$	100^{+58}_{-57}
435–480	$1.005801^{+0.006816}_{-0.005627}$	143^{+13}_{-11}	8678^{+5806}_{-3563}	$0.69^{+2.14}_{-2.71}$	$28.69^{+41.19}_{-63.23}$	$0.99^{+0.85}_{-1.44}$	68^{+51}_{-50}
480–525	$1.005022^{+0.007217}_{-0.005679}$	114^{+10}_{-9}	9578^{+6698}_{-3852}	$0.80^{+2.55}_{-3.18}$	$31.39^{+46.54}_{-74.40}$	$0.66^{+0.72}_{-1.24}$	28^{+49}_{-44}
525–570	$1.005316^{+0.007256}_{-0.006051}$	130^{+11}_{-10}	10054^{+6911}_{-4102}	$0.81^{+2.51}_{-3.16}$	$36.68^{+50.96}_{-84.59}$	$1.56^{+1.27}_{-2.30}$	29^{+48}_{-44}

Table 4.6: Same as Table 4.5, except for the GP model with only ϕ and ψ as inputs.

Channel (nm)	f_0 (ppm)	σ_w (ppm)	A (ppm)	L_ϕ (Δ_ϕ)	L_ψ (Δ_ψ)	δ_s (ppm)
290–450	$1.007423^{+0.008339}_{-0.006967}$	147^{+12}_{-11}	14342^{+15631}_{-6961}	$0.72^{+2.00}_{-2.53}$	$43.16^{+50.11}_{-78.01}$	114^{+39}_{-41}
450–570	$1.008237^{+0.008900}_{-0.008031}$	92^{+8}_{-7}	21110^{+22873}_{-10493}	$0.98^{+2.68}_{-3.56}$	$60.75^{+62.11}_{-113.50}$	-16^{+27}_{-24}
290–340	$1.007652^{+0.009147}_{-0.008294}$	794^{+62}_{-55}	23662^{+29813}_{-12277}	$2.81^{+3.17}_{-5.39}$	$53.00^{+40.54}_{-89.41}$	143^{+178}_{-179}
340–390	$1.007235^{+0.008996}_{-0.008104}$	430^{+36}_{-33}	22266^{+28139}_{-11872}	$2.58^{+1.28}_{-4.31}$	$61.01^{+50.84}_{-103.57}$	140^{+98}_{-99}
390–435	$1.006984^{+0.008278}_{-0.006874}$	197^{+18}_{-15}	14610^{+17461}_{-7435}	$0.85^{+1.97}_{-2.59}$	$51.77^{+48.45}_{-89.24}$	87^{+47}_{-49}
435–480	$1.008080^{+0.008444}_{-0.007447}$	145^{+12}_{-11}	17384^{+20431}_{-8727}	$0.86^{+2.20}_{-2.88}$	$51.51^{+57.35}_{-93.34}$	44^{+36}_{-38}
480–525	$1.008485^{+0.009009}_{-0.008185}$	120^{+10}_{-9}	22994^{+27442}_{-11509}	$0.98^{+2.56}_{-3.38}$	$66.20^{+66.00}_{-122.88}$	-48^{+32}_{-30}
525–570	$1.008122^{+0.008886}_{-0.008035}$	130^{+11}_{-9}	22746^{+26146}_{-11785}	$1.04^{+2.68}_{-3.64}$	$69.71^{+65.33}_{-131.66}$	-6^{+33}_{-32}

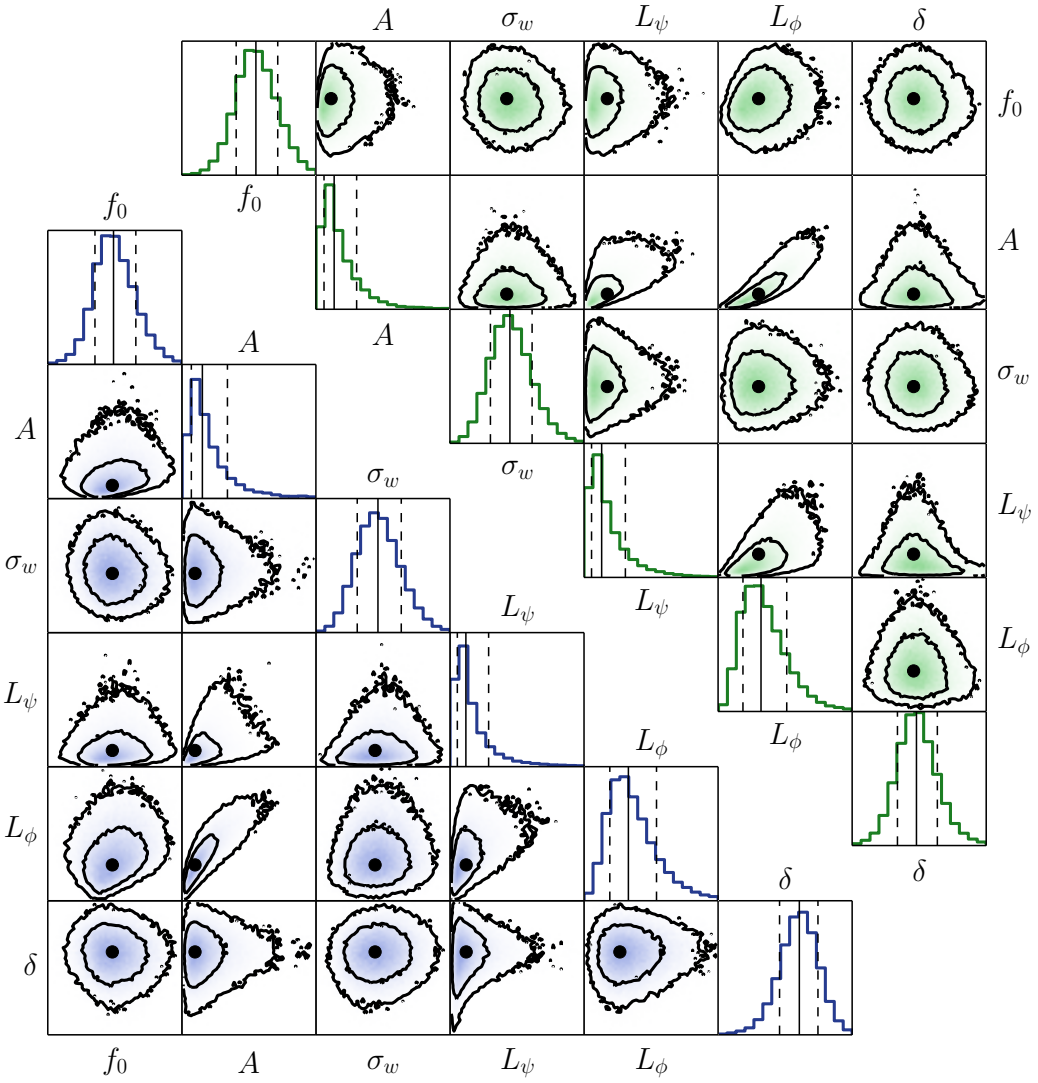

Figure 4.5: MCMC chain histograms for the 290–450 nm (bottom left) and 450–570 nm (top right) wavelength channels. 2D histograms show the joint distributions for parameter pairs, with solid lines indicating the ranges containing 68% and 95% of the samples, and black circles showing the MLE solutions. 1D histograms show the distributions for each parameter marginalised over all other parameters, with solid vertical lines indicating the median and dashed lines indicating ranges either side of the median that each contain 34% of the samples. Corresponding numerical values can be found in Table 4.4 and 4.6, with the purpose of this figure being to display the shape of the posterior distribution.

Medians for each parameter are listed in Tables 4.5 and 4.6 for the cases with $\{\phi, \psi, x\}$ and $\{\phi, \psi\}$ as GP inputs, respectively. Also reported are the ranges below and above the median that each included 34% of the samples. Marginalised posterior distributions are plotted in Figure 4.5 for the two-channel lightcurves, including the correlations for each pair of parameters. Similar distributions were obtained for the six-channel lightcurves, but are not plotted.

For both sets of inputs, the inferred values for δ_s, f_0, and σ_w are consistent within $\sim 1\sigma$. The main effect of excluding x from the GP inputs is to shrink the uncertainties on the inferred eclipse depths by factors ranging between ~ 1.05–1.55. This is probably due to the fact that ψ and x are degenerate with each other. Therefore, I favour the smaller uncertainties of the $\{\phi, \psi\}$ GP, except for the two wavelength channels (i.e. 450–570 nm and 480–525 nm) that exhibit evidence for relatively short correlation length scales L_x (Table 4.5).

4.6 Discussion

4.6.1 Albedo constraints

Table 4.7 lists the values for the geometric albedo A_g that are obtained from Equation 1.3 using the measured δ_s values listed in Tables 4.5 and 4.6, with $R_p/R_\star = 0.155 \pm 0.003$ (Pont et al., 2013) and $a/R_\star = 8.86 \pm 0.02$ (Agol et al., 2010). The most striking result is that a nonzero albedo value of $A_g = 0.37^{+0.12}_{-0.13}$ is measured for the 290–450 nm channel at 2.8σ signficance, while the albedo constraint for the 450–570 nm channel of $A_g = 0.13^{+0.13}_{-0.12}$ is consistent with zero to within $\sim 1\sigma$. Propagating the uncertainties in quadrature, the albedo is higher in the 290–450 nm channel than it is in the 450–570 nm channel, but only at 1.3σ significance. This increases to a 2.3σ difference if the measured albedo for the 290–450 nm channel is instead compared with that measured for the narrower 525–570 nm channel ($A_g = -0.02^{+0.10}_{-0.11}$). The broad trend of decreasing albedo

Table 4.7: Geometric albedos A_g calculated from the measured eclipse depths in Tables 4.5 and 4.6 using Equation 1.3. Values are reported for the GP models with $\{\phi, \psi, x\}$ and $\{\phi, \psi\}$ as input variables. Quoted uncertainties have been calculated by propagating the uncertainties on δ_s, $a/R_\star$, and $R_p/R_\star$ in quadrature. Flux-weighted central wavelengths are also given for each channel.

		A_g	
Channel (nm)	Central λ (nm)	ϕ, ψ, x	ϕ, ψ
290–450	413	$0.41^{+0.16}_{-0.16}$	$0.37^{+0.12}_{-0.13}$
450–570	510	$0.13^{+0.13}_{-0.12}$	$-0.05^{+0.09}_{-0.08}$
290–340	325	$0.45^{+0.60}_{-0.59}$	$0.46^{+0.57}_{-0.58}$
340–390	368	$0.48^{+0.35}_{-0.35}$	$0.45^{+0.32}_{-0.32}$
390–435	416	$0.32^{+0.19}_{-0.18}$	$0.28^{+0.15}_{-0.16}$
435–480	459	$0.22^{+0.16}_{-0.16}$	$0.14^{+0.12}_{-0.12}$
480–525	502	$0.09^{+0.16}_{-0.14}$	$-0.16^{+0.10}_{-0.10}$
525–570	547	$0.09^{+0.15}_{-0.14}$	$-0.02^{+0.11}_{-0.10}$

with increasing wavelength is also recovered from the six-channel analysis, but the uncertainties are too large to decipher any detail.

Comparing the albedos listed in Table 4.7 to those obtained by Berdyugina et al. (2011) using polarimetry (see Section 4.1), the latter are systematically higher. In particular, Berdyugina et al. obtain values of $A_g \sim 0.7$ in both the U (330–400 nm) and B (400–490 nm) bands, compared with the value of $A_g = 0.37^{+0.12}_{-0.13}$ over the 290–450 nm wavelength range obtained in the current study. For the six-channel wavelength analysis, I found $A_g = 0.28^{+0.15}_{-0.16}$ over 390–435 nm and $A_g = 0.14^{+0.12}_{-0.12}$ over 435–480 nm, which together approximately span the B band. These latter values are, respectively, 2.3σ and 3.8σ lower than the B band value of $A_g = 0.69 \pm 0.08$ found by Berdyugina et al. Additional polarisation and photometry measurements are required to resolve this tension. Meanwhile, at longer wavelengths, the value of $A_g = 0.13^{+0.13}_{-0.12}$ acquired for the 450–570 nm range in the current study is consistent with the Berdyugina et al. value of

$A_g = 0.28 \pm 0.16$ in the V (500–600 nm) band at the 1σ level.

4.6.2 Stellar variability

In principle, variations in the brightness of the star itself, rather than the planetary eclipse, could be responsible for the measured change in the system brightness. This is particularly pertinent for HD 189733, which is known to be an active K dwarf with star spots that have estimated temperatures $\sim$ 500–1000 K cooler than the surrounding stellar photosphere (Pont et al., 2008a). Expressing this quantitatively, the measured eclipse depth δ'_s will be related to the true eclipse depth δ_s according to:

$$\delta'_s(\lambda) = \frac{\delta_s(\lambda)\, F_\star(\lambda)}{(1-\alpha)\, F_\star(\lambda) + \alpha\, F_\circ(\lambda)}\,, \tag{4.5}$$

where α is the fractional spot coverage of the stellar disc, $F_\star$ is the surface brightness of the un-spotted stellar photosphere, and $F_\circ$ is the surface brightness of the spotted regions. If we assume that the true eclipse depth is wavelength independent over the 290–570 nm range, such that $\delta_s(\lambda) = \delta_0$, we obtain the relationship:

$$\frac{\delta'_{s,1}}{\delta'_{s,2}} = \frac{1-\alpha+\alpha\eta_2}{1-\alpha+\alpha\eta_1} \tag{4.6}$$

where $\delta'_{s,1}$ and $\delta'_{s,2}$ are the measured eclipse depths in two bandpasses $\Delta\lambda_1$ and $\Delta\lambda_2$, respectively, and:

$$\eta_i = \frac{\int_{\Delta\lambda_i} F_\circ(\lambda)\, d\lambda}{\int_{\Delta\lambda_i} F_\star(\lambda)\, d\lambda}\,. \tag{4.7}$$

To estimate the magnitude of this effect, I evaluated Equation 4.6 by multiplying the transmission function of the G430L grating (Figure 4.1) with an ATLAS spectrum (Kurucz, 1979, 1993) computed specifically for HD 189733[8]

[8] Available at `http://kurucz.harvard.edu/stars/hd189733`

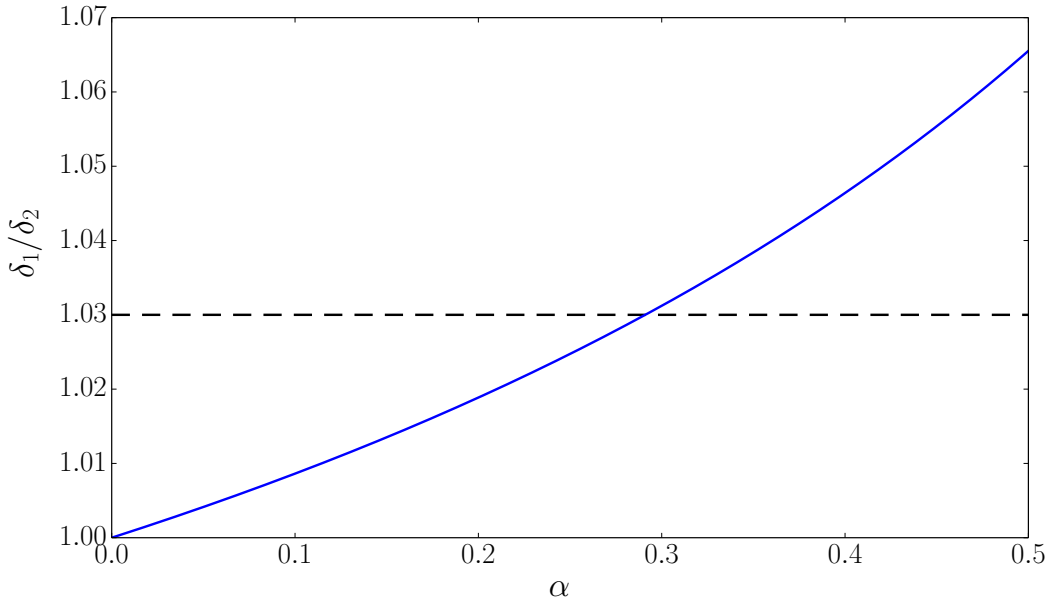

Figure 4.6: Relative flux changes measured for the $\Delta\lambda_1$ = 290–450 nm and $\Delta\lambda_2$ = 525–570 nm wavelength channels due to star spots as a function of the fractional spot coverage α. Specifically, if a measured flux change $\delta_{s,1}$ is due to variations in star spot coverage at wavelengths $\Delta\lambda_1$, then $\delta_{s,2}$ is the flux variation that would be expected at wavelengths $\Delta\lambda_2$. These values were computed using Equation 4.6 and assuming ATLAS models for the star and spot spectra (see Section 4.6.2 for details). As the dashed horizontal line shows, the relative flux change between the two wavelength channels is $\delta_{s,1}/\delta_{s,2} < 1.03$ for spot coverages up to $\sim 20\%$ (i.e. $\alpha \sim 0.2$).

and an ATLAS spectrum with $T_{\rm eff,\star} = 4250$ K, $\log_{10} g = -4.5$ cgs, and [M/H] = 0 for the spots. The results are plotted in Figure 4.6 for the $\Delta\lambda_1$ = 290–450 nm and $\Delta\lambda_2$ = 525–570 nm bandpasses (Table 4.7). They show that $\delta'_{s,1}/\delta'_{s,2} < 1.03$ for spot coverages up to $\sim 20\%$ ($\alpha \sim 0.2$). Assuming then that the inferred albedo of $A_g = -0.02^{+0.12}_{-0.10}$ for the 525–570 nm channel is due to star spots, the albedo that would mistakenly be inferred for the 290–450 nm channel is therefore $A_g \lesssim 0.1$ at 1σ confidence. This implies that measured value of $A_g = 0.37^{+0.12}_{-0.13}$ for the 290–450 nm channel would have to be a $\sim 2.3\sigma$ outlier if the observed wavelength-dependence is caused by star spots.

The characteristic amplitude of flux variations due to stellar activity can also be estimated using the fit made by McQuillan et al. (2012) to the power spectra of the brightest K dwarfs in the Q1 *Kepler* dataset. Specifically, McQuillan et al.

modelled the power spectra of the stellar lightcurves as:

$$\text{Power}(\upsilon) \quad = \quad \frac{G}{1+(H\upsilon)^C} \, , \tag{4.8}$$

where υ is the flux variation frequency, and the parameters G, H, and C were inferred by fitting to the *Kepler* data. Thus, the characteristic amplitudes of flux variations at two frequencies υ_1 and υ_2 are related by:

$$\frac{\text{Amp}(\upsilon_1)}{\text{Amp}(\upsilon_2)} \quad = \quad \sqrt{\frac{1+(H\upsilon_2)^C}{1+(H\upsilon_1)^C}} \, . \tag{4.9}$$

Long-term photometric monitoring data shows that the characteristic amplitude of variations on ~ 12 day timescales is ~ 7000 ppm (S. Aigrain, 2013, private communication). Substituting this in to Equation 4.9, along with the values of $H = 2.8 \times 10^5$ and $C = 2.3$ obtained by McQuillan et al., gives a characteristic amplitude of ~ 95 ppm on 90 min timescales corresponding to the *HST* orbital period. This is close to the variations measured in the 290–450 nm wavelength channels, raising the possibility that the measured signal is in fact due to stellar flux variations rather than the planetary eclipse. However, it should be stressed that this calculation is very approximate – in particular, the value of H was poorly constrained by the study of McQuillan et al. Also, as was argued above, the wavelength-dependence of the measured signal is stronger than would be expected if the flux change was caused by changes in the star spot coverage. Therefore, although it is not possible to definitively rule out stellar variability as the cause of the measured signal, the planetary eclipse interpretation is favoured by the available evidence.

4.6.3 Implications for the atmosphere

To first order, the reflection spectrum of a hot Jupiter at visible wavelengths will depend on whether or not there are reflective clouds in the atmosphere, and the altitude at which they become optically thick (e.g. Seager et al., 2000;

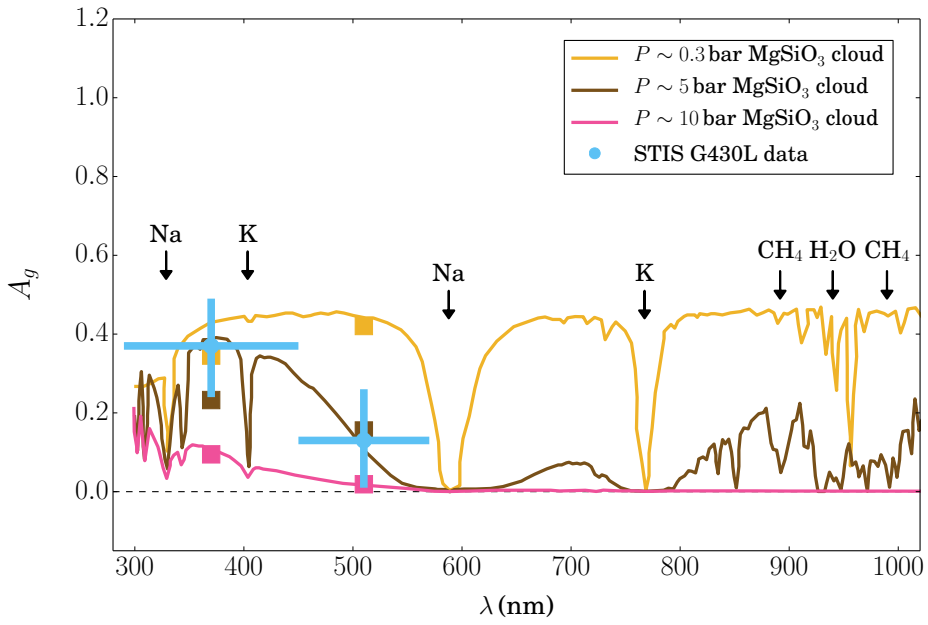

Figure 4.7: Albedo measurements for the 290–450 nm and 450–570 nm wavelength channels. Solid lines show model hot Jupiter reflection spectra taken from Sudarsky et al. (2000). The main difference between each model is the pressure level, and hence altitude, at which a layer of $MgSiO_3$ cloud is located, as indicated in the legend. Note that the thermal emission contribution is not included in the Sudarsky et al. models. However, for the 290–570 nm wavelength range probed by the STIS measurements, thermal emission from HD 189733b is negligible.

Sudarsky et al., 2000). This is illustrated in Figure 4.7, which shows the measured two-channel albedos and three model reflection spectra produced by Sudarsky et al. (2000). The latter have been taken directly from the original paper, and have not been fit to the data in any way. It should also be noted that these models were produced for notional hot Jupiter atmospheres with simplified pressure-temperature profiles and the assumption of chemical equilibrium with solar abundances. They are reproduced here primarily to illustrate the trade-off between alkali absorption and cloud reflection, and how this varies with the cloud altitude. Specifically, the yellow line shows a model with a relatively high-altitude layer of $MgSiO_3$ cloud, which produces a correspondingly high albedo throughout the visible wavelength range, interrupted only by deep absorption features due to Na and K. The brown and pink lines show the emission for similar model atmospheres, with cloud layers located progressively lower in the atmosphere.

The decrease in cloud altitude translates to an increase in the column density of atmosphere that incident stellar radiation must escape through once it has been scattered by the cloud layer. As such, absorption by the broad wings of the Na and K features increases significantly, lowering the albedo. Similar effects are observed for the H_2O and CH_4 features at infrared wavelengths.

The relatively high albedo measured for the 290–450 nm channel compared to the 450–570 nm channel disfavours the presence of a high-altitude, thick cloud layer in the atmosphere of HD 189733b, as has been suggested by Pont et al. (2013). Under such a scenario we would expect the albedos to be similarly high in both channels, due to the masking of the Na and K absorption wings, as the yellow line in Figure 4.7 shows. Nonetheless, the measured albedos are still consistent with a layer of reflective cloud located deeper in the atmosphere, which is illustrated by the good agreement between the data and the brown line in Figure 4.7.

If the measured reflection signal is due to clouds, an interesting question is whether or not it is caused by the same scattering species that produces the Rayleigh profile in the transmission spectrum. Given the strong atmospheric circulation expected for hot Jupiters, this seems plausible (e.g. Showman et al., 2009; Perna et al., 2012; Parmentier et al., 2013). As has been mentioned previously, $MgSiO_3$ grains are likely candidates, being transparent in the visible and formed of atoms abundant in hot atmospheres (Lecavelier Des Etangs et al., 2008).

The limited precision of the current data prevents more intricate scenarios from being ruled out. For example, a high altitude tenuous haze could account for the transmission signal while being transparent at zenith geometry, with a lower, denser layer of clouds producing the reflection signal. The absorption in the 450–570 nm channel could also be due to an absorber other than Na, yet to be

identified. For instance, TiO is an efficient absorber in the 450–570 nm channel, although it is expected to have rained out of the atmosphere of HD 189733b (Fortney et al., 2008).

Aside from clouds, scattering by H_2 molecules could be responsible for the reflection signal measured for the 290–450 nm channel. This possibility has been explored by Barstow et al. (2014), who carried out atmosphere retrievals (Section 1.4.2) using the NEMESIS code (Irwin et al., 2008) with the STIS albedos as input. The authors found that H_2 scattering in a cloud-free atmosphere was capable of reproducing the measured albedos, without needing to resort to more complicated cloud treatments. Therefore, the albedo measurements presented here do not provide definitive proof that optically thick clouds are present on the dayside hemisphere of HD 189733b. The strongest statement that can be made is that Na absorption is very likely to be responsible for the lower albedo in the 450–570 nm channel. This in turn rules out a high-altitude layer of reflective cloud that dominates the dayside hemisphere.

If indeed Na absorption is responsible for the lower albedo measured in the 450–570 nm channel, it would be suggestive of a low Bond albedo[9] for the planet. This follows from the simple argument that if the Na 589 nm doublet is not entirely masked by clouds, the situation could be similar at wavelengths beyond 570 nm where theory predicts significant absorption by the K 770 nm doublet and molecules such as CH_4 and H_2O (Figure 4.7). However, our observations do not provide a model-independent constraint on the Bond albedo (Section 1.4), as only ~5% and ~15% of the stellar flux is emitted in the 290–450 nm and 450–570 nm channels, respectively (see Figure 4.1). A more meaningful constraint could be placed on the Bond albedo by acquiring another eclipse measurement with the G750L grating, as its bandpass accounts for a further

[9]Recall from Section 1.4 that the Bond albedo is the fraction of incident stellar radiation that is reflected back to space integrated over all solid angles and all wavelengths.

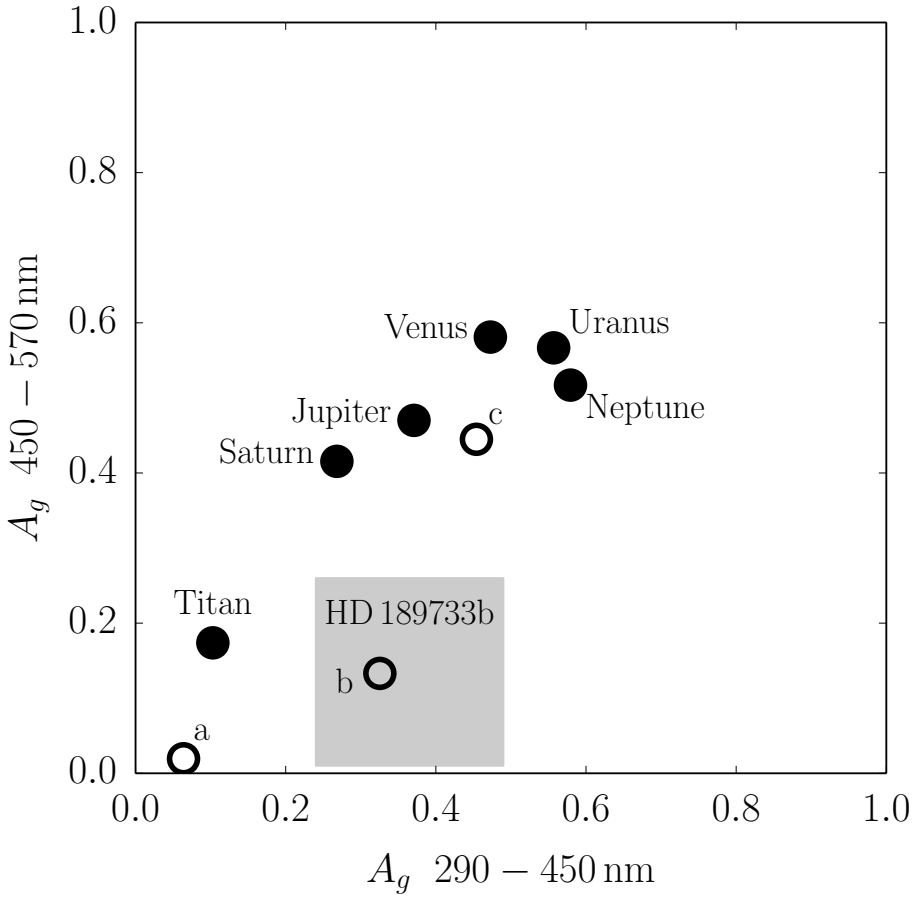

Figure 4.8: Colour-colour plot with filled black circles showing geometric albedos for various solar system bodies and grey box showing the 1σ probability range for HD 189733b derived in this study. Solar system albedos are from Karkoschka (1994), except for the Venus albedo which comes from Titov et al. (2007). Also shown as open circles are the three Sudarsky et al. models from Figure 4.7, with clouds located at different pressure levels in the atmosphere: (a) $P \sim 10$ bar, (b) $P \sim 5$ bar, and (c) $P \sim 0.3$ bar.

~ 50% of the incident stellar flux. We did apply for such an observation during *HST* Proposal Cycle 22, but were unfortunately unsuccessful.

Finally, Figure 4.8 shows HD 189733b on a colour-colour plane with a selection of solar system bodies that also have thick atmospheres, and predicted albedos from the three Sudarsky et al. models plotted in Figure 4.7. The albedo constraints for the 290–450 nm and 450–570 nm channels presented in this chapter suggest that HD 189733b is a deep, dark blue, quite distinct from the atmosphere colours seen in our solar system.

Chapter 5

Ground-based transmission spectroscopy

This chapter presents two primary transit observations for the hot Jupiter HAT-P-32b that were made using the 2.5 m *Nordic Optical Telescope* (*NOT*) at the Observatorio del Roque de los Muchachos on La Palma, Canary Islands, Spain. The aim of this study is to characterise the transmission spectrum of the planetary atmosphere over the 414–702 nm wavelength range using multi-object spectroscopy (MOS).

Section 5.1 outlines the basic principles of the MOS approach and Section 5.2 describes the HAT-P-32 system. Details of the *NOT* observations are given in Section 5.3, followed in Section 5.4 by a description of the data reduction process used to produce time series spectra from the raw data frames. The issue of spectral trace drift across the detector, which is believed to be responsible for systematic trends observed in the final lightcurves, is addressed in Section 5.5. Section 5.6 explains how the lightcurves are generated by binning the timeseries spectra in wavelength, and the subsequent model fitting. Section 5.7 discusses the results, focusing on the implications for the transmission spectrum, and outlines future work.

5.1 Multi-object spectroscopy

Transit observations made from the ground are complicated by the time-variable nature of the Earth's atmosphere. Phenomena such as turbulent air flows, humidity fluctuations, and atmospheric refraction give rise to correlated noise in the lightcurves, with timescales ranging from minutes to hours. For this reason, it is crucial to obtain simultaneous observations for one or more nearby comparison stars. Systematics that affect all stars in the FOV can then be removed from the target by dividing its flux by the combined fluxes of the comparisons, which are assumed to be intrinsically constant.

During recent years, such an approach has been implemented using MOS to measure exoplanet transmission spectra (e.g. Bean et al., 2010, 2011; Gibson et al., 2013a,b). This has been facilitated by the discovery of suitable targets with large atmospheric scale heights in fields rich with nearby comparison stars, as part of ground-based surveys (e.g. HATNet, WASP; Section 1.2). Most of these MOS observations have been performed using custom-designed masks cut with slitlets that coincide with the target and comparison stars, allowing non-overlapping stellar spectra to be projected on to the detector. In other cases, the target and a comparison star have been aligned within the single aperture of a conventional long-slit spectrograph (e.g. Sing et al., 2012; Murgas et al., 2014).

A key feature of successful MOS observations made to date has been the use of wide (> 10 arcsec) slit apertures. Under typical ~ 1 arcsec seeing, this greatly reduces systematics caused by time-variable slit losses. Otherwise, slit losses will tend to vary for each star due to slight mask misalignment, making an accurate differential correction of the target flux effectively impossible. Meanwhile, the seeing-limited spectral resolution afforded by wide slits combined with a medium-resolution grating is amply sufficient for characterising the broad features of

exoplanet transmission spectra. The working rule of thumb is therefore: the wider the slits, the better. A practical upper limit to the slit widths is imposed, however, by the requirement that the spectra of neighbouring stars do not impinge upon those of the main target and comparisons on the detector.

Our group at Oxford has used the Gemini Multi Object Spectrograph (GMOS) instruments on the *Gemini-South* and *Gemini-North* telescopes to measure transmission spectra for WASP-29b (Gibson et al., 2013a) and HAT-P-32b (Gibson et al., 2013b). In both cases, we were able to rule out strong absorption signals due to Na and TiO/VO over the 500–700 nm wavelength range, suggesting that these features are probably masked by high-altitude haze/clouds.

The success of these GMOS observations encouraged us to acquire the additional two transits for HAT-P-32b that are presented in this chapter. Our primary goals were to: (1) demonstrate the feasibility of MOS observations on the 2.5 m *NOT*; (2) search for evidence of Rayleigh scattering by either H_2 or haze by extending the transmission spectrum to wavelengths as short as 400 nm; and (3) verify the overall lack of absorption features across the 500–700 nm wavelength range, including the Na doublet at 589 nm.

5.2 The HAT-P-32 system

The hot Jupiter HAT-P-32b was discovered by Hartman et al. (2011) as part of the HATNet survey (Section 1.2). With a mass of $0.860 \pm 0.164\,M_{\rm J}$ and radius of $1.789 \pm 0.025\,R_{\rm J}$, it orbits a late-F-to-early-G dwarf star once every $2.1500085 \pm 2 \times 10^{-7}$ days. Measurements of the Rossiter-McLaughlin effect made by Albrecht et al. (2012) have shown that the stellar rotation axis lies close to the orbital plane of the system, with a projected obliquity of $\lambda_{\rm obliq} = 85.0 \pm 1.5$ deg. Meanwhile, Hartman et al. noted that the orbital eccentricity was poorly constrained by the RV data presented in the discovery paper, but that a circular orbit fit was favoured

Table 5.1: Basic stellar and planetary parameters for the HAT-P-32 system, taken from Hartman et al. (2011), Albrecht et al. (2012), and Gibson et al. (2013b).

Stellar parameters		
Parameter	Unit	Value
Right ascension, RA	hh mm ss	02 04 10.28
Declination, Dec	dd mm ss	+46 41 16.2
Magnitude, V	mag	11.4
Spectral type	—	F/G V
Effective temperature, $T_{\mathrm{eff},\star}$	K	6207 ± 88
Mass, $M_\star$	$M_\odot$	1.16 ± 0.04
Radius, $R_\star$	$R_\odot$	1.22 ± 0.02
Gravity, $\log g_\star$	cgs	4.33 ± 0.01
Metallicity, [Fe/H]	dex	−0.04 ± 0.08
Rotation speed, $v_\star \, sini$	$\mathrm{km\,s^{-1}}$	21.0 ± 0.5
Ca II index, $\log R'_{HK}$	dex	−4.62

Planet parameters		
Parameter	Unit	Value
Period, P	day	$2.1500085 \pm 2 \times 10^{-7}$
Semimajor axis, a	AU	0.0343 ± 0.0004
Mass, M_p	M_J	0.860 ± 0.164
Radius, R_p	R_J	1.789 ± 0.025
Density, ρ_p	$\mathrm{g\,cm^{-3}}$	0.19 ± 0.04
Inclination, i	deg	88.70 ± 0.6
Eccentricity, e	—	0
Projected obliquity, λ_obliq	deg	85.0 ± 1.5
Equilibrium temperature, T_eq	K	1 786 ± 26
Atmospheric scale height, H	km	1 000

by the Bayesian information criterion (Equation 3.6). This is consistent with theoretical expectations, as tidal circularisation of the orbit is expected to occur on $\lesssim 1\,\mathrm{Gyr}$ timescales due to the small distance between the planet and star (Section 1.3). An orbital eccentricity of zero is therefore adopted for this study. Other basic system parameters are summarised in Table 5.1.

Hartman et al. also found that the residuals of their RV fit were unusually high, with an amplitude of $80\,\mathrm{m\,s^{-1}}$. These departures from a simple Keplerian model raised the possibility of additional planets in the system, but the authors were unable to draw a firm conclusion based on the available data. Seeliger et al. (2014) have since presented a timing analysis of transits observed over a $\sim 6\,\mathrm{yr}$ baseline, and found no evidence for variations due to additional companions. Indeed, Hartman et al. preferred an explanation based on convective instabilities in the stellar atmosphere, which have been proposed as the primary source of RV jitter for F stars (Saar et al., 1998). Hartman et al. also noted that HAT-P-32 appears to be a photometrically quiet star, with no significant flux variations observed in the HATNet lightcurves.

Of particular relevance to the current observations is the fact that the atmospheric scale height H (Equation 1.8) of HAT-P-32b is among the largest of the known exoplanets ($H \sim 1000\,\mathrm{km}$; see Table 1.1), owing to its low density ($\rho_p \sim 0.2\,\mathrm{gm\,cm^{-3}}$) and high equilibrium temperature ($T_{\mathrm{eq}} \sim 1800\,\mathrm{K}$). The change in effective planetary radius corresponding to a single atmospheric scale height is equivalent to a change in the radius ratio of $\Delta R_p/R_\star \sim 0.0012$, which translates to a transit depth variation of $\sim 360\,\mathrm{ppm}$. Furthermore, the host star is moderately bright ($V = 11.4\,\mathrm{mag}$) and located in a field with nearby comparison stars of similar brightness. These properties make HAT-P-32b an especially favourable target for using MOS to measure the transmission spectrum.

5.3 NOT/ALFOSC observations

Two consecutive primary transits of HAT-P-32b were observed on the nights of 2013 September 6 and 8 using the Andalucia Faint Object Spectrograph (ALFOSC) mounted at the Cassegrain focus of the NOT. The instrument has a 2048×2048 pixel CCD with a 6.4×6.4 arcmin field of view and a plate scale of 0.19 arcsec pix^{-1}. Observations were made using grism #10, which has a wavelength coverage of ~350–1000 nm (Figure 5.1) and dispersion of 0.65 nm pix^{-1}. For a PSF FWHM of 1 arcsec and 2×2 pixel binning, which was used on both nights to reduce overheads, this translates to resolving powers of $R \sim 175$ at 400 nm and $R \sim 410$ at 700 nm. At the expense of fringing at wavelengths longer than ~ 730 nm, a second-order blocking filter was not used in order to preserve the full throughput for the shorter wavelengths. The fastest available read-out speed of 400 kpix sec^{-1} was used, resulting in a per-exposure overhead of 7 sec.

Pre-images of the field were obtained on 2013 July 22 and used to design a custom mask, with slitlets for HAT-P-32 and four nearby comparison stars (Figure 5.2). The mask was cut[1] from a 70×72 mm piece of black plastic and mounted in the ALFOSC aperture wheel, where the 2.05×4.10 mm slitlets translated to on-sky dimensions of 15×30 arcsec. During the observations, the telescope rotator was used to keep the stars aligned within the slitlets.

On both nights, a series of 10–20 dark frames were acquired with the same exposure time as the science frames. However, given the short exposure times and wavelength range of the science frames, the dark current was negligible. We measured the CCD bias level on the first and second nights to be 10^4 ADU counts and 1.6×10^4 ADU counts, respectively. The higher counts for the second night were caused by a bug in the ALFOSC system that occurs when the readout speed

[1]The mask was cut by Niels Michaelsen and Michael Anderson at the Niels Bohr Institute, University of Copenhagen, Denmark.

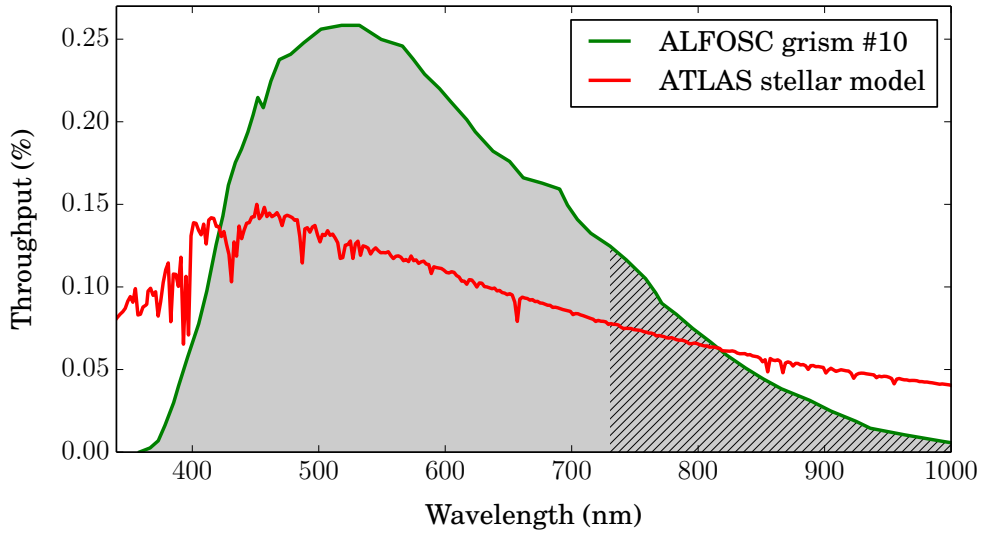

Figure 5.1: System throughput as a function of wavelength for the ALFOSC grism #10. The hatched region indicates wavelengths that are affected by second-order fringing effects and the red line shows a normalised ATLAS model spectrum (Kurucz, 1979, 1993) with parameters $T_{\mathrm{eff},\star} = 6250\,\mathrm{K}$, $\log g = 4.5$, and $[\mathrm{M/H}] = 0$, which are similar to those of HAT-P-32 (Table 5.1).

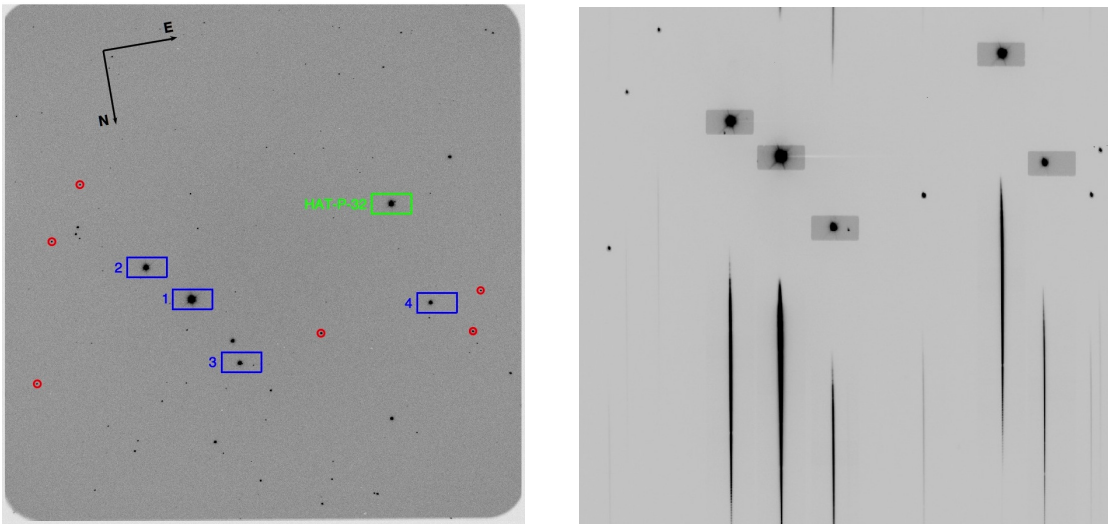

Figure 5.2: (Left) ALFOSC mask design overlayed on a pre-image of the HAT-P-32 field. The green box shows the 15×30 arcsec slitlet for the target, and the blue boxes show the slitlets for the comparison stars numbered in order of decreasing brightness. Red circles show the apertures that were used to accurately align the mask upon acquisition. The image field of view is 6.4×6.4 arcmin. (Right) An example science frame from the second night, with white to black indicating low to high counts. Dispersion is along the vertical axis. Zeroth-order spectra are projected within the slitlet outlines, while the first order spectra are visible as dark stripes directly below and above the corresponding slitlets.

is set after the pixel binning has been set. Apart from the increased bias level, the data appeared to be unaffected.

Flats were acquired during the afternoons by illuminating the science mask with a halogen lamp installed inside the ALFOSC instrument, with the grism in place but no filter. In addition, flats were obtained using a second calibration mask with the same slitlet layout but narrower 1 arcsec slitlets. The latter frames were intended to better simulate the instrument response for a point source under median seeing conditions. Unfortunately, these flats were compromised by a considerable misalignment between the calibration and science mask slitlets along the cross-dispersion axis due to an error in the mask cutting.

Wavelength arc frames were obtained immediately after the science observations on the first night, and immediately before and after the science observations on the second night. A helium lamp was used with the calibration mask to provide the reference spectrum. Despite the misalignment between the slitlets of the science and calibration masks along the cross-dispersion axis, the good alignment along the dispersion axis allowed an accurate wavelength calibration to be determined, as described in Section 5.4.3.4.

5.3.0.1 2013 September 6

Observations were made continuously for 384 min during dark time between UT 21:33–UT 03:57. This covered the entire 186 min transit, which occurred between UT 22:34–01:40, as well as a pre-transit baseline of 61 min and a post-transit baseline of 124 min. An exposure time of 30 sec was used to acquire 586 frames over the course of the night, corresponding to a duty cycle of 79%. Conditions were clear and there was no moon. The instrumental seeing was 2.8 arcsec for the entire pre-transit baseline, but decreased to 1 arcsec by the time of mid-transit, where it remained for the rest of the observations (Figure 5.3). At the start of the observations, the target was at an airmass of ~4, with the zeroth

order stellar spectra visibly elongated due to differential atmospheric refraction. By the time of ingress, the target had risen to an airmass of 2.6, and continued rising until the end of the observations, by which point the airmass was 1.1. The difference between the mask field angle and the parallactic angle varied from −36 deg to 72 deg over the course of the observations, as shown by the "ΔField" variable in Figure 5.4.

5.3.0.2 2013 September 8

Observations were made continuously for 328 min between UT 00:45–UT 06:13, starting in dark time and ending 27 min before sunrise. This included a pre-transit baseline of 85 min and a post-transit baseline of 57 min, with the transit occurring between UT 02:10–05:16. The median instrumental seeing was 1 arcsec throughout the night (Figure 5.3). The conditions were clear and there was no moon. An exposure time of 20 sec was used to allow for the possibility of the seeing improving and to increase the time sampling of the lightcurve. This allowed 723 frames to be obtained at a duty cycle of 73%, which was slightly lower than that achieved on the first night owing to the shorter exposure time. The target was at an airmass of 1.4 at the beginning of the pre-transit baseline, rising to minimum airmass of 1.1 at mid-transit, and then lowering again to an airmass of 1.2 by the end of the post-transit baseline (Figure 5.4). During this time, the difference between the mask field angle and the parallactic angle increased from −2 deg and to 147 deg.

5.4 Data reduction

5.4.1 Dark frames and static bad pixels

Master darks were generated for each night by median-combining the pixel values of the individual dark frames. These were subtracted from each science

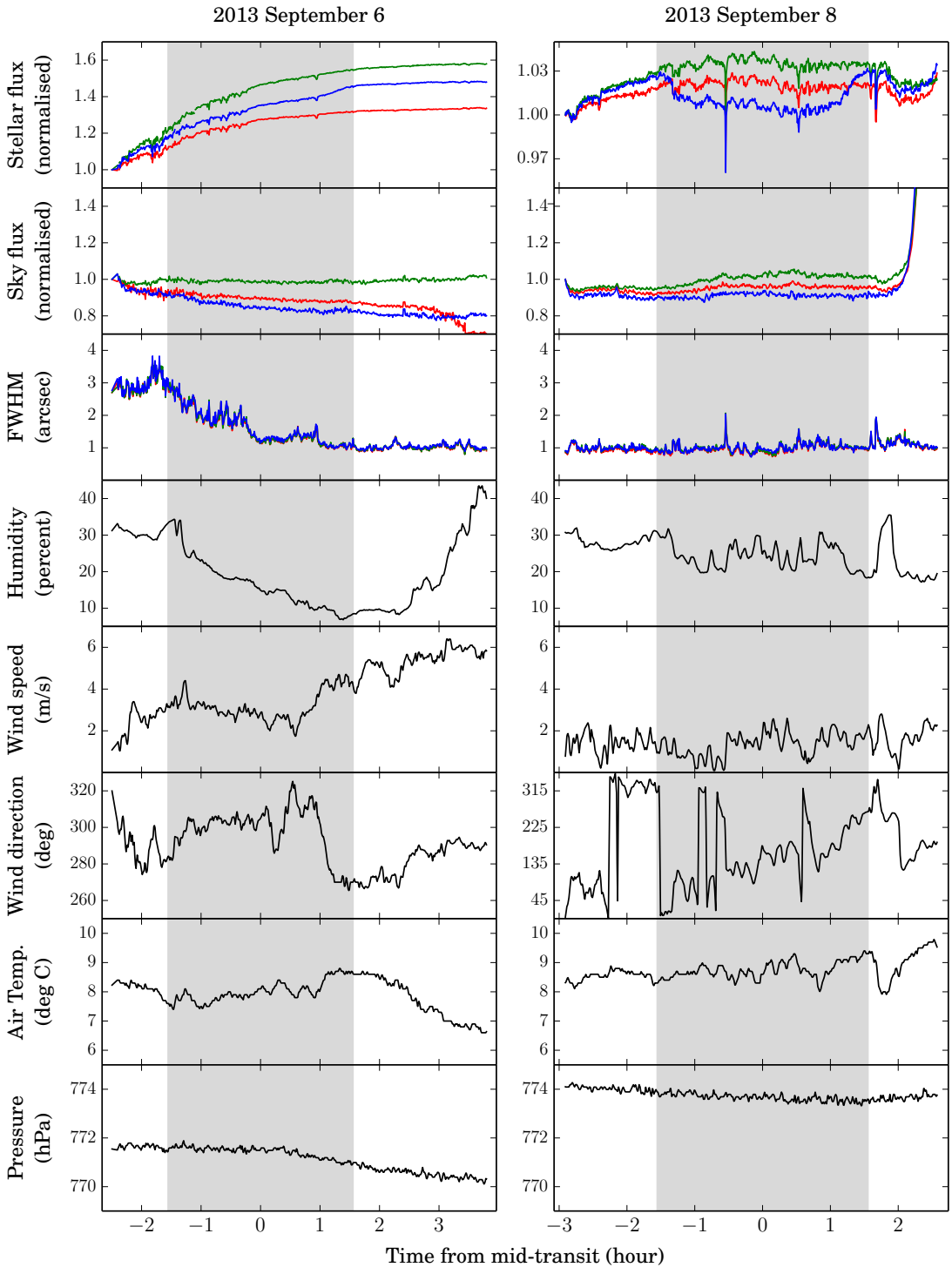

Figure 5.3: Observing conditions for the first (left) and second (right) nights. Blue lines show variables extracted for the target star; green and red lines show variables extracted for the two brightest comparison stars; and black lines show generic variables. Shaded grey regions show the times of transit and hatched regions show data that is discarded from the analysis (see Section 5.6.1). Rows from top to bottom are: (1) wavelength-integrated, median-normalised stellar flux; (2) wavelength-integrated, median-normalised sky flux; (3) FWHM of the spectra cross-dispersion profiles; (4) atmospheric humidity; (5) wind speed; (6) wind direction; (7) air temperature; (8) atmospheric pressure.

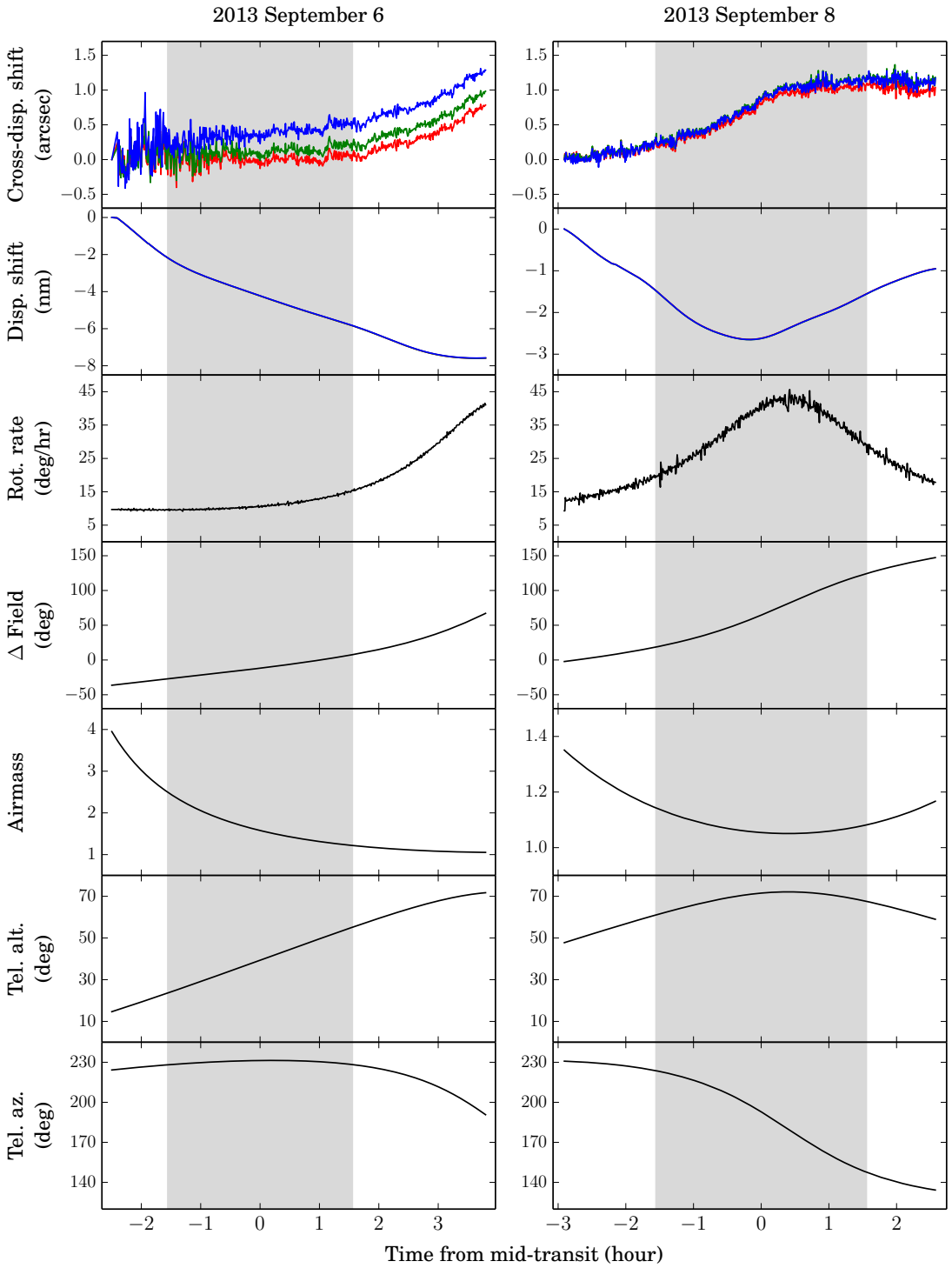

Figure 5.4: Variables pertaining to the instrument state for the first (left) and second (right) nights. Line colours and shading are the same as for Figure 5.3. Rows from top to bottom are: (1) spectral trace drift along dispersion axis (see Section 5.5); (2) spectral trace shift along cross-dispersion axis; (3) telescope rotator rate of change; (4) angle between the slitlet axes and the parallactic; (5) airmass; (6) target altitude; (7) target azimuth.

frame and used to generate a map of static bad pixels; namely, those with dark values that differed by more than 10σ from the median of all other pixels on the array. The bad pixels were masked in the subsequent analysis.

5.4.2 Flat fields

Flat fields were median-combined to produce master flats. To determine the pixel-to-pixel variations in the detector response, it was necessary to remove the contribution from the lamp spectrum. This was done by taking the median pixel values along the cross-dispersion axis within the spectral apertures of each slitlet, and fitting Gaussian processes (GP) to the resulting spectral profiles using squared exponential kernels (Equation 2.27). These spectral profiles were then subtracted from all columns along the cross-dispersion axis. Due to the cross-dispersion misalignment of the science and calibration masks (Section 5.3), only the flats taken with the science mask were used. For the lightcurve analysis, however, fully consistent results were obtained regardless of whether or not flatfielding was performed, owing to the relative nature of the flux measurement (i.e. the target flux is measured relative to the comparison fluxes). Below, results obtained from the analysis without flatfielding are presented, giving preference to the simpler reduction.

5.4.3 The `spectroscopy` pipeline

To extract the stellar spectra from the science frames following dark subtraction, the `spectroscopy` package was used. The software was written specifically for the current work, but it can be applied to any single slit or multiple slit spectroscopy dataset of a similar nature. It is written in the Python programming language and has been made publicly available online.[2]

[2] `http://www.github.com/tomevans`

5.4.3.1 Transient bad pixel flagging

The first step of the pipeline is to flag transient bad pixels that were not included in the static bad pixel maps generated from the master dark frames. This is done by computing the median values for each pixel from the 15 frames immediately preceding and the 15 frames immediately following each science frame. Any pixel with a value that differs from the median by $> 10\sigma$ is flagged as bad and masked in the subsequent analysis. Two iterations of this culling were performed. For the first night, 1456 bad pixels were flagged across the 586 frames. For the second night 2627 bad pixels were flagged across the 723 frames, with one frame in particular containing 1851 bad pixels.

5.4.3.2 Spectral trace fitting

The second step is to fit the spectral traces in each of the science frames. This is done for each star by first extracting the subarray containing the spectral trace, and then binning the flux along the dispersion axis. Bin widths of 30 pixels were used for the current study. To determine the location of the spectral trace center along the dispersion axis, the cross-dispersion profiles of each bin are fit with a 1D Gaussian plus linear slope. The means of the fitted Gaussians are then linearly interpolated or fit with a low-order polynomial to map the spectral trace coordinates to each pixel along the dispersion axis. For the current study, third order polynomial fits were performed. Meanwhile, the widths of the fitted Gaussian profiles provided the FWHM of the instrumental PSF (Figure 5.3).

5.4.3.3 Background subtraction and stellar flux extraction

The third step is to extract the stellar flux along the dispersion axis of each science frame. This is done for each pixel column along the dispersion axis by summing the pixels either side of the spectral trace center. Before summing

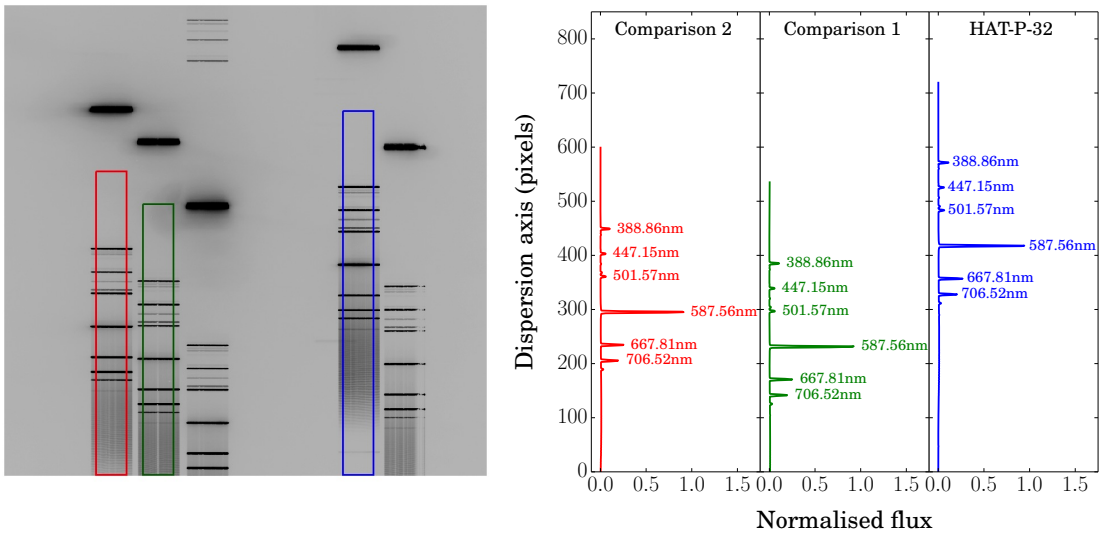

Figure 5.5: (Left) Wavelength arc frame obtained by illuminating the 1 arcsec slitlet calibration mask with a helium lamp. White indicates low counts, black indicates high counts. The blue box shows the region used to calibrate the wavelength scale for the target, while the green and red boxes show the regions used for the first and second comparisons, respectively. (Right) Extracted calibration spectra for the target and two comparison stars, with fiducial emission lines indicated.

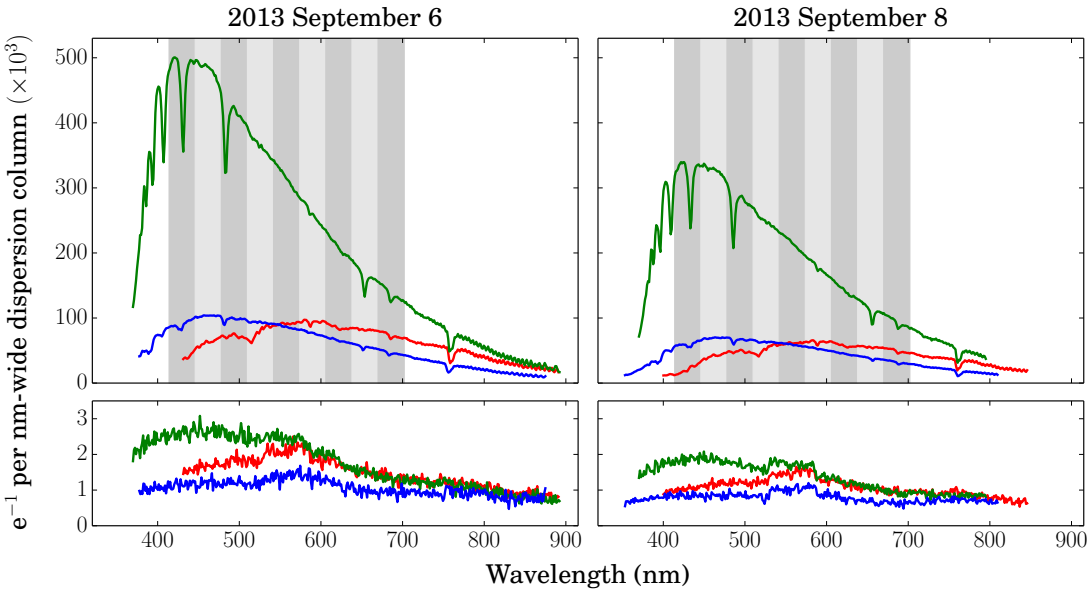

Figure 5.6: (Top) Example stellar spectra for the target (blue line) and two brightest comparisons (green and red lines) from each night. Grey shaded columns indicate the wavelength channels that were used to generate the spectroscopic lightcurves. (Bottom) Corresponding sky spectra extracted for each star. The fact that the sky spectra for each star are not identical might indicate contamination by the wings of the stellar PSFs; however, the sky spectra were extracted from regions 13.3 arcsec away from the center of the stellar cross-dispersion profiles, compared with the typical PSF widths of < 2 arcsec. Alternatively, it could be due to flatfielding effects, or diffuse stray light.

the flux within the aperture, however, the background must be estimated and removed. To do this, the median flux is calculated for a string of pixels either side of the trace, beyond the edges of the stellar PSF. These background values are linearly interpolated through the spectral aperture and removed. The stellar flux is then calculated by summing the background-subtracted pixel values within the aperture. Fractional fluxes are calculated for pixels that are only partially contained within the aperture at the edges.

For the current study, sky values were calculated from 10 pixel bands either side of the spectral trace at distances of 35 pixel (13.3 arcsec), and the stellar fluxes were extracted using aperture radii of 8 pixels (3.0 arcsec). This aperture radius encompasses a significant fraction of the PSF of a nearby ($\sim$ 2.8 arcsec) M dwarf companion to HAT-P-32 that was noted by Gibson et al. (2013b). Although this companion will dilute the flux extracted for HAT-P-32, Gibson et al. pointed out that the former is approximately 60 times fainter than the latter at wavelengths $\sim$ 850 nm and show that the effect on the inferred planet parameters will be negligible. This will also be true for the 414–702 nm wavelength range probed in the current study, where the flux from the M dwarf will be even smaller compared to that of HAT-P-32. For this reason, the contribution from the M dwarf companion to the extracted stellar flux is ignored.

Example spectra for the target and two brightest comparison stars are shown in Figure 5.6, along with the corresponding sky spectra. In theory, the sky spectra extracted for each star should be the same, but there are clear differences (compare the blue, green, and red lines in the bottom panels of Figure 5.6). It seems unlikely that this is due to contamination by the wings of the stellar PSFs, given that the FWHM was typically < 2 arcsec (Figure 5.3). However, the extracted sky spectra do appear to be correlated with the corresponding stellar spectra to some degree, suggesting that a diffuse contribution from the stellar

flux is non-negligible even at such a distance from the PSF center. An alternative explanation may be that the differences are caused by spatial variations in the detector sensitivity, given that flatfielding was not performed. A similar effect was observed, however, when the reduction was repeated using the flatfields described in Section 5.4.2, but this could also indicate that the flatfields were ineffective. This seems plausible, given that the flatfields were generated using the science mask with 15 arcsec slitlets, which is poorly matched to the seeing conditions.

5.4.3.4 Wavelength calibration

The fourth and final step of the reduction is to calibrate the wavelength scale of the spectral traces. To do this, calibration spectra were extracted from the master helium arcs in the regions containing the stellar spectra in the science frames, as illustrated in Figure 5.5. A number of fiducial lines were identified in the calibration spectra, and the corresponding pixel coordinates determined by fitting the line profiles with Gaussians. A third-order polynomial fit was then used to map the fiducial line centers from pixel coordinates to wavelengths. The maximum discrepancy between the polynomial fit and the measured line centers was < 0.05 nm for both nights.

5.5 Spectral trace drift

The first row of Figure 5.4 shows how the median cross-dispersion coordinate of the spectral traces vary over both nights. The cross-dispersion drifts are approximately correlated with the "Δ Field" angle shown in the fourth row of Figure 5.4; namely, the difference between the mask field angle and the parallactic angle, which is a proxy for the rotator angle. The amplitude of the cross-dispersion drift is similar for both nights, as is the range spanned

by the rotator angle. Taken together, this suggests that a mechanical effect could be responsible for the cross-dispersion drift of the trace, such as imperfect alignment of the stars within the slitlets or instrument flexure. Note that the cross-dispersion drift is automatically accounted for during the reduction process: the cross-dispersion profile of the spectral traces are fitted in each frame, before the stellar fluxes are extracted (Section 5.4.3.3).

The drift of the spectral traces along the dispersion axis can also be seen in the top panels of Figure 5.7, which shows the flux timeseries for the target star as a function of wavelength. In particular, the location of stellar absorption lines can be seen to vary with respect to the wavelength solution obtained from the calibration arcs (Section 5.4.3.4) during both nights. These drifts were traced by fitting Gaussian profiles to the stellar absorption lines at 434 nm, 486 nm, 656 nm, and 759 nm for each frame (Figure 5.6). The resulting time series of line center measurements were then smoothed by fitting them with squared exponential GPs (Equation 2.27), and interpolated to intermediate wavelengths. The bottom panels of Figure 5.7 show the calculated dispersion drifts as a function of wavelength and frame number.

Meanwhile, the second row of Figure 5.4 shows the dispersion drift of the median wavelength only, revealing an approximate inverse correlation with the telescope rotator rate, which is shown in the same figure. This suggests instrument flexure as a possible source of the spectral trace drift. However, a similar correlation is also observed for the airmass. Indeed, the amplitude of the dispersion drift was larger for the observations made on the first night, which encompassed a larger range in airmass. On the other hand, the rotator rate variation was similar on both nights, despite the variation in the spectral trace drift amplitude being more than two times greater on the first night. This argues in favour of atmospheric refraction being responsible for the observed dispersion

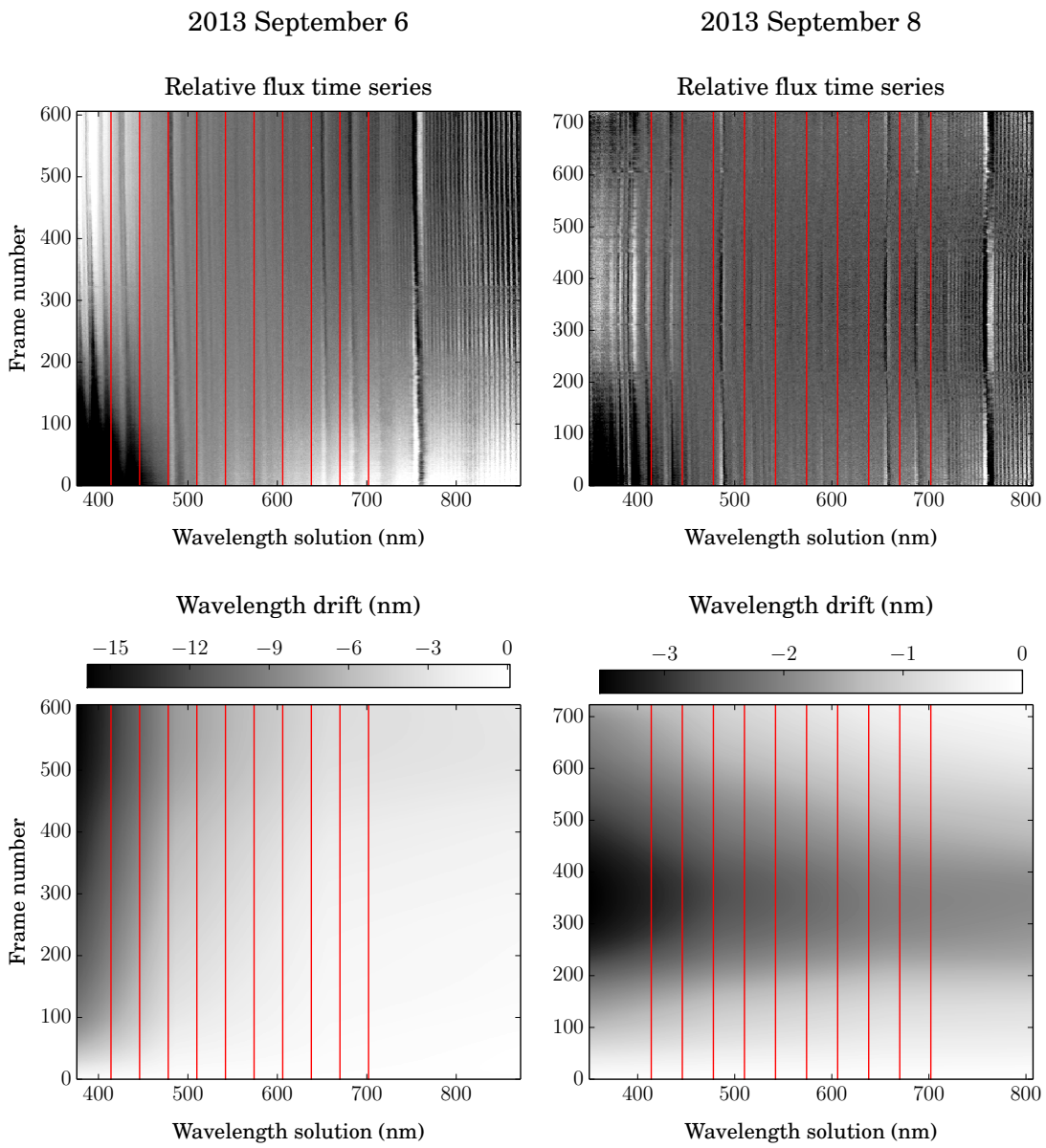

Figure 5.7: (Top panels) Measured flux for the target star as a function of wavelength along the horizontal axis and frame number along the vertical axis, for the first (left) and second (right) nights. Each column has been divided by the "white" lightcurve generated by integrating along the dispersion axis, and then normalised by the column mean. Black indicates low flux relative to the column mean, and white indicates relatively high flux. Vertical red lines show the edges of the wavelength channels used for the spectroscopic lightcurves, which are described in Section 5.6.2. Fringing effects are visible for wavelengths > 730 nm. (Bottom panels) Dispersion axis drifts as functions of wavelength and frame number, estimated by fitting Gaussian profiles to five stellar absorption lines at wavelengths 434 nm, 486 nm, 656 nm, and 759 nm for each frame. These line centers were interpolated to estimate the drift at other wavelengths along the dispersion axis. The colorbars give the amplitude of the wavelength drift for both nights.

drift. Further evidence is provided by the fact that the dispersion drift is largest at short wavelengths (Figure 5.7), as would be expected for refraction. Curiously, however, the dispersion drift minimum precedes the airmass minimum on the second night by ~30 min (Figure 5.4).

5.6 Lightcurve analysis

Raw flux timeseries were produced for the target and two brightest comparisons by integrating the spectra over the 414–702 nm wavelength range, which encompasses the bulk of the stellar flux (Figure 5.6). Broadband "white" lightcurves were generated for both nights by summing over the entire bandpass, and are shown in the top row of Figure 5.3. "Spectroscopic" lightcurves were also generated by summing within nine contiguous channels over the same wavelength range, each with a width of 32 nm (Figures 5.6 and 5.7). The target fluxes were then divided by the sum of the comparison fluxes, in order to remove time-dependent systematics that were common to all three stars. The resulting white lightcurves are shown in Figure 5.8, and the spectroscopic lightcurves are shown in Figure 5.9.

5.6.1 White lightcurves

The initial section of the white lightcurve from the first night is affected by high-amplitude noise. This is likely due to effects associated with observing at high airmass, such as scintillation, lower flux counts, and slit losses caused by an elongated PSF. For the second night, the last section of the white lightcurve exhibits an exponential rise, corresponding to a similar increase in the sky flux as the sun started to rise. However, the sky flux should be subtracted from the stellar flux during the data reduction (Section 5.4.3.3), so in theory the two should not be correlated. One possibility is that any relative errors in the background

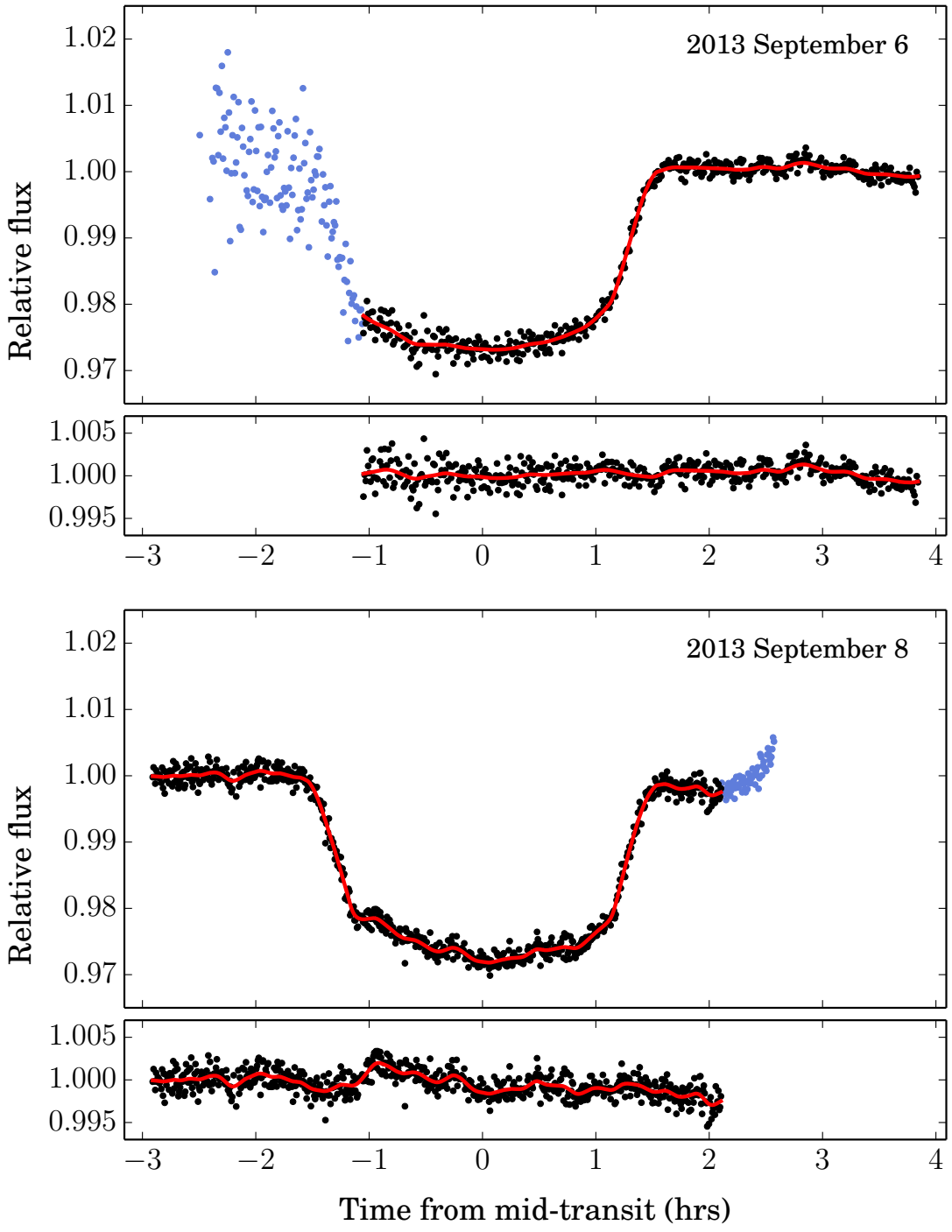

Figure 5.8: White lightcurves for 2013 September 6 and 8, which have been integrated over the 414–702 nm wavelength range. Black points show the data that is included in the lightcurve analyses, and blue points show the data that is discarded for the reasons explained in the main text. Red lines show the predictive means of the best-fit GP models. The axes immediately below each raw lightcurve show the residuals after dividing through by the best-fit transit signal.

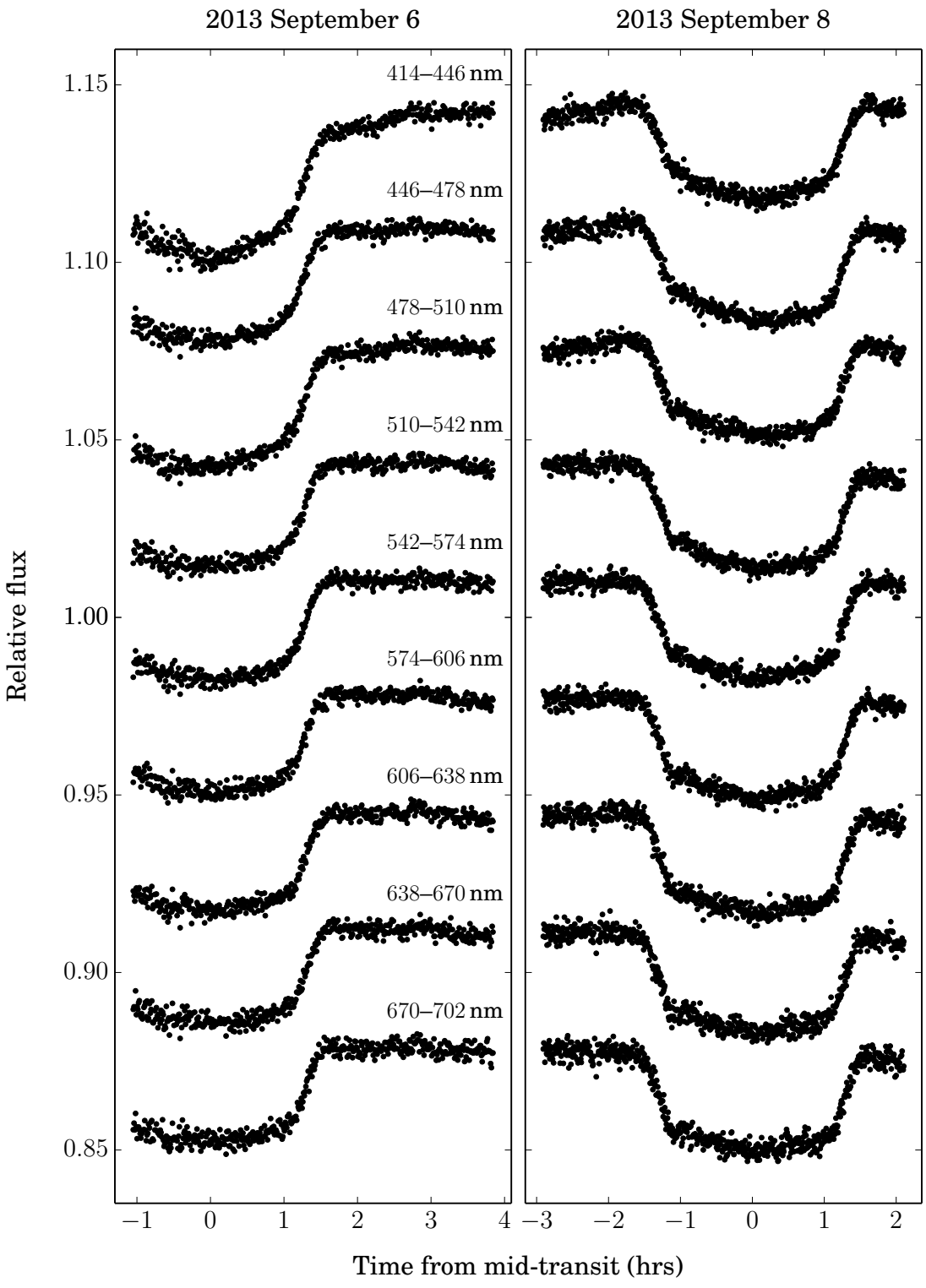

Figure 5.9: Raw spectroscopic lightcurves for 2013 September 6 and 8. Wavelength channels are labelled for 2013 September 6, and are identical for 2013 September 8. Vertical offsets have been added for clarity.

subtraction that were negligible for most of the observations were amplified as the sky flux increased exponentially at the end. To prevent these sections of the lightcurves biasing the final results, the initial 86 min of the first night and the final 28 min of the second night were discarded prior to fitting.

The remaining white lightcurves for both nights display low-amplitude correlated noise superimposed on the transit signals (Figure 5.8). To account for these systematics, a GP model was adopted for each lightcurve following the approach outlined in Section 2.3, and similar to that already described for the *Spitzer*/IRAC (Chapter 3) and *HST*/STIS (Chapter 4) analyses. The joint model likelihood therefore took the form:

$$p\left(\boldsymbol{d_1}, \boldsymbol{d_2} \,|\, \boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \boldsymbol{\gamma_1}, \boldsymbol{\gamma_2}\right) \;=\; \mathcal{N}\left(\boldsymbol{\mu_1}, \mathbf{K_1} + \boldsymbol{\Sigma_1}\right) \times \mathcal{N}\left(\boldsymbol{\mu_2}, \mathbf{K_2} + \boldsymbol{\Sigma_2}\right) \,, \quad (5.1)$$

where the $\boldsymbol{d_i}$ terms denote the datasets; the $\boldsymbol{\alpha_i}$ and $\boldsymbol{\gamma_i}$ terms denote the mean function and covariance parameters, respectively; and subscripts distinguish between the two nights, e.g. d_1 and d_2 denote the lightcurves from the first and second nights, respectively.

Again, Mandel & Agol (2002) transit functions were used for both GP means. The orbital period was fixed to $P = 2.1500080$ day, based on the ephemeris provided by Gibson et al. (2013b), and the eccentricity was fixed to zero. The radius ratio $R_p/R_\star$, normalised semimajor axis $a/R_\star$, and impact parameter b were shared between the two lightcurves and allowed to vary during the fitting. Separate transit mid-times, $T_{\mathrm{mid},1}$ and $T_{\mathrm{mid},2}$, were fit for each lightcurve. A quadratic law was used for the stellar limb darkening, with coefficients c_1 and c_2 shared between the two lightcurves. To constrain the latter within physically reasonable ranges, the quadratic coefficient was fixed to $c_2 = 0.35$, which was derived by fitting a quadratic law to the limb darkening profile of an ATLAS stellar model (Kurucz, 1979, 1993) with properties similar to HAT-P-32 integrated over

Table 5.2: White lightcurve fixed parameters and MCMC results. The orbital period P was fixed to the value of Gibson et al. (2013b). The quadratic limb darkening parameter c_2 was fixed to the value giving the best-fit to an ATLAS stellar model Kurucz (1979, 1993) with similar properties to HAT-P-32 integrated over the 414–702 nm bandpass. MLE values are quoted for the covariance parameters $\{A_1, L_{t,1}, \sigma_{w,1}, A_2, L_{t,2}, \sigma_{w,2}\}$, which were fixed for the MCMC analysis. MCMC chain medians are quoted for the remaining mean function parameters $\{R_p/R_\star, a/R_\star, b, c_1, T_{\mathrm{mid},1}, T_{\mathrm{mid},2}\}$, with uncertainties that correspond to the ranges either side of the medians that containing 34% of the chain samples. Orbital inclination i values are derived from the impact parameter $b = a\cos i/R_\star$ and system scale $a/R_\star$ values.

Parameter	Value	Unit
Global parameters		
P	2.1500080	day
$R_p/R_\star$	$0.1503^{+0.0009}_{-0.0009}$	–
$a/R_\star$	$6.09^{+0.05}_{-0.05}$	–
b	$0.118^{+0.046}_{-0.063}$	–
i	$88.89^{+0.60}_{-0.44}$	degree
c_1	$0.32^{+0.04}_{-0.04}$	–
c_2	0.35	–
Transit 1		
$T_{\mathrm{mid},1}$	$2456542.504601^{+0.000475}_{-0.000465}$	$\mathrm{HJD_{UTC}}$
A_1	543	ppm
$L_{t,1}$	9.2	minute
$\sigma_{w,1}$	1443	ppm
Transit 2		
$\mathrm{T_{mid,2}}$	$2456544.654790^{+0.000404}_{-0.000395}$	$\mathrm{HJD_{UTC}}$
A_2	969	ppm
$L_{t,2}$	10.3	minute
$\sigma_{w,2}$	1019	ppm

the 414–702 nm bandpass. The linear coefficient c_1, however, was allowed to vary.

The covariance for both nights was modelled using Matérn $\nu = 3/2$ kernels (Equation 2.29), with time as the only input variable. The Matérn kernel was chosen to handle non-smooth correlations in the time series (Figure 5.8), that would not be accounted for as well by a squared exponential kernel. Time was used as the only input, due to the lack of obvious correlations between the lightcurve systematics and the other auxiliary variables shown in Figures 5.3 and 5.4. The covariance parameters for the first night were thus a covariance amplitude A_1, a correlation length scale $L_{t,1}$, and a white noise term $\sigma_{w,1}$, with an analogous set of parameters for the second night.

The results of Gibson et al. (2013b) were used as priors for the mean function parameters $R_p/R_\star$, $a/R_\star$, and b. Namely:

$$p\left(R_p/R_\star\right) = \mathcal{N}\left(0.1515, 0.0012^2\right) , \tag{5.2}$$

$$p\left(a/R_\star\right) = \mathcal{N}\left(6.091, 0.005^2\right) , \tag{5.3}$$

$$p\left(b\right) = \mathcal{N}\left(0.093, 0.070^2\right) . \tag{5.4}$$

Uniform priors were assumed for the remaining parameters.

As with the lightcurve fits presented in Chapters 3 and 4 (in particular, see Section 3.7.4), the posterior distribution was first optimised by allowing the mean function and covariance parameters to vary. Explicitly, the free mean function parameters were:

$$\{\, \boldsymbol{\alpha_1} \,,\, \boldsymbol{\alpha_2} \,\} = \{\, R_p/R_\star \,,\, a/R_\star \,,\, b \,,\, c_1 \,,\, T_{\mathrm{mid},1} \,,\, T_{\mathrm{mid},2} \,\} , \tag{5.5}$$

and the free covariance parameters were:

$$\{\, \boldsymbol{\gamma_1} \,,\, \boldsymbol{\gamma_2} \,\} = \{\, A_1 \,,\, L_{t,1} \,,\, \sigma_{w,1} \,,\, A_2 \,,\, L_{t,2} \,,\, \sigma_{w,2} \,\} . \tag{5.6}$$

A number of optimisations were performed, each with different starting positions drawn randomly from the prior, in order to increase the chance of locating the global likelihood maximum.

The covariance parameters were then fixed to their MLE values, and marginalisation was performed for the mean function parameters using MCMC with Metropolis-Hastings sampling. Step sizes were pretuned to give an acceptance rate of 20–40 % using the same procedure described in Section 3.7.4 for the IRAC lightcurve analyses. Five independent chains of 10^5 steps were then run from well-dispersed starting locations. After conservatively discarding the first 5×10^4 steps as burn-in, the Gelman-Rubin statistics were found to be well within 1% for all parameters, as expected for chains that have reached stable states (Section 2.2.1).

Results for the MLE covariance parameters and marginalised mean function parameters are given in Table 5.2, and the best-fit model is shown in Figure 5.8. The derived values for $R_p/R_\star$, $a/R_\star$, and b are all consistent with those reported previously by Hartman et al. (2011) and Gibson et al. (2013b). A correlation length scale of approximately 10 min is inferred for both nights, as would be expected if the systematics are caused by the same process. On the other hand, a covariance amplitude of $A_1 \sim 500$ ppm is obtained for the first night, compared to $A_2 \sim 1000$ ppm for the second night. However, this is balanced against the higher white noise value of $\sigma_{w,1} \sim 1400$ ppm inferred for the first night, compared with the value of $\sigma_{w,2} \sim 1000$ ppm inferred for the second night, despite the shorter exposure times used for the latter (i.e. 20 sec versus 30 sec). This suggests that additional high-frequency noise is present in the lightcurve for the first night, separate to the correlations on ~ 10 min timescales, and which the GP model attempts to model as an additional source of white noise.

5.6.2 Spectroscopic lightcurves

A common-mode correction was applied to the raw spectroscopic lightcurves shown in Figure 5.9 using the residuals from the white lightcurve fits presented in the previous section. This was performed by dividing each spectroscopic lightcurve by the best-fit GP predictive mean of the corresponding white lightcurve, with the transit signal removed, i.e. the red lines plotted over the residuals in Figure 5.8.

The corrected spectroscopic lightcurves are shown in Figure 5.10. Most of the the low-amplitude correlated noise has been successfully removed; however, non-common-mode systematics are still visible for a number of lightcurves. In particular, smooth trends of increasing flux are evident for the three shortest wavelength channels, which may be the combination of two effects. Firstly, the drift along the dispersion axis noted in Section 5.5 is strongest at shorter wavelengths, and will cause the amount of flux falling within a given channel to vary with time. Secondly, the three shortest wavelength channels coincide with strong stellar absorption lines of Hγ at 434 nm and Hβ at 486 nm (Figure 5.6), which could exacerbate this effect. A possible solution would be to apply the estimated dispersion drifts shown in Figure 5.7 to the wavelength solutions on a frame-by-frame basis before extracting the spectroscopic lightcurves.

In the current analysis, however, the smooth systematics are accounted for by modelling the spectroscopic lightcurves individually as a transit signal multiplied by a quadratic polynomial in time. Using the same notation as above (e.g. Equation 5.1), the model likelihoods for each spectroscopic lightcurve are therefore given by:

$$p(\boldsymbol{d}_1, \boldsymbol{d}_2 | \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) = \mathcal{N}\left(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1\right) \times \mathcal{N}\left(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2\right) , \tag{5.7}$$

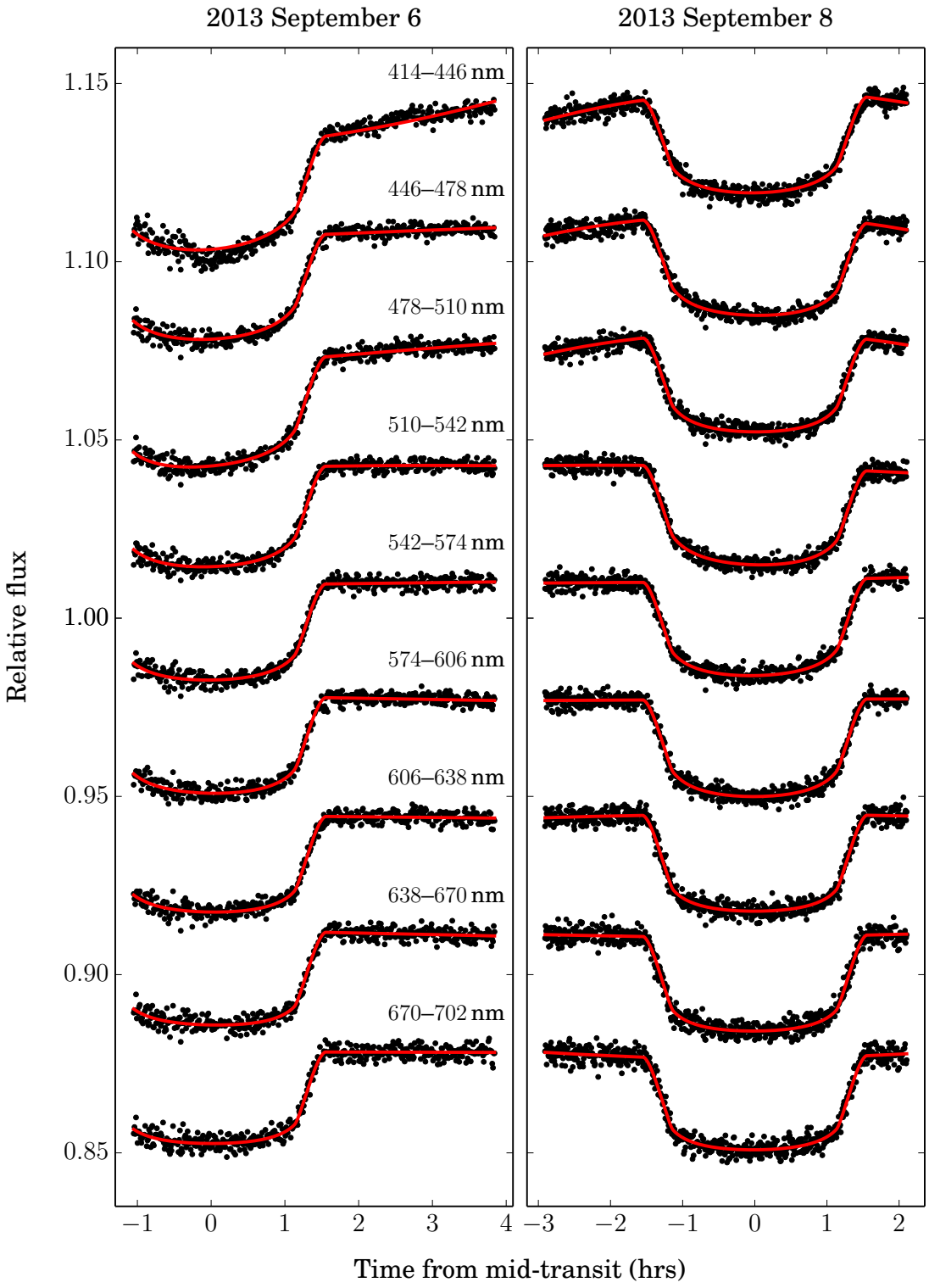

Figure 5.10: The same as Figure 5.9, but showing the spectroscopic lightcurves after dividing through by the common-mode correction, as described in Section 5.6.2. Red lines show the best-fit models, which consist of a transit signal multiplied by a quadratic polynomial in time.

with mean functions taking the form:

$$\mu(t) \quad = \quad E(t)\left(a_0 + a_1 t + a_2 t^2\right) \tag{5.8}$$

where E is a Mandel & Agol (2002) transit function. Note that the covariance matrices $\mathbf{\Sigma_1}$ and $\mathbf{\Sigma_2}$ in Equation 5.7 are diagonal; as was discussed in Section 2.3, this is equivalent to assuming that the noise superimposed on the mean function is uncorrelated. A full GP model with parameterised off-diagonal covariance entries would provide a more rigorous systematics treatment, similar to that performed for the white lightcurves in Section 5.6.1; however, such an analysis would be more computationally expensive, and is deferred to future work.

For the simpler analysis presented here, the only transit parameters allowed to vary during the fits were the radius ratio $R_p/R_\star$ and the linear limb darkening coefficient c_1. As with the white lightcurve analyses presented in Section 5.6.1, the quadratic limb darkening coefficients c_2 were fixed to values obtained by fitting to an ATLAS stellar model that had been integrated over the appropriate bandpasses. Again, the orbital period was fixed to $P = 2.1500080$ day (Gibson et al., 2013b), and the eccentricity set to zero. Remaining transit parameters ($a/R_\star$, b, $T_{\mathrm{mid},1}$, $T_{\mathrm{mid},2}$) were fixed to the MLE values determined from the white lightcurve analyses. For each of the nine spectroscopic lightcurves, the total set of free parameters thus consisted of:

$$\{\, \boldsymbol{\alpha_1}\,,\, \boldsymbol{\alpha_2}\,\} \quad = \quad \{\, R_p/R_\star\,,\, c_1\,,\, a_0\,,\, a_1\,,\, a_2,\, a_0'\,,\, a_1'\,,\, a_2'\,,\, \sigma_{w,1}\,,\, \sigma_{w,2}\,\}\,, \tag{5.9}$$

where the a_i and a_i' terms denote the coefficients of the quadratic polynomial baselines for the first and second nights, respectively. Uniform priors were adopted for all parameters.

Marginalisation of the model posterior distribution was performed using the same procedure described in Section 5.6.1 for the white lightcurves: first, the MLE solution was obtained, then five independent MCMC chains of 10^5 steps

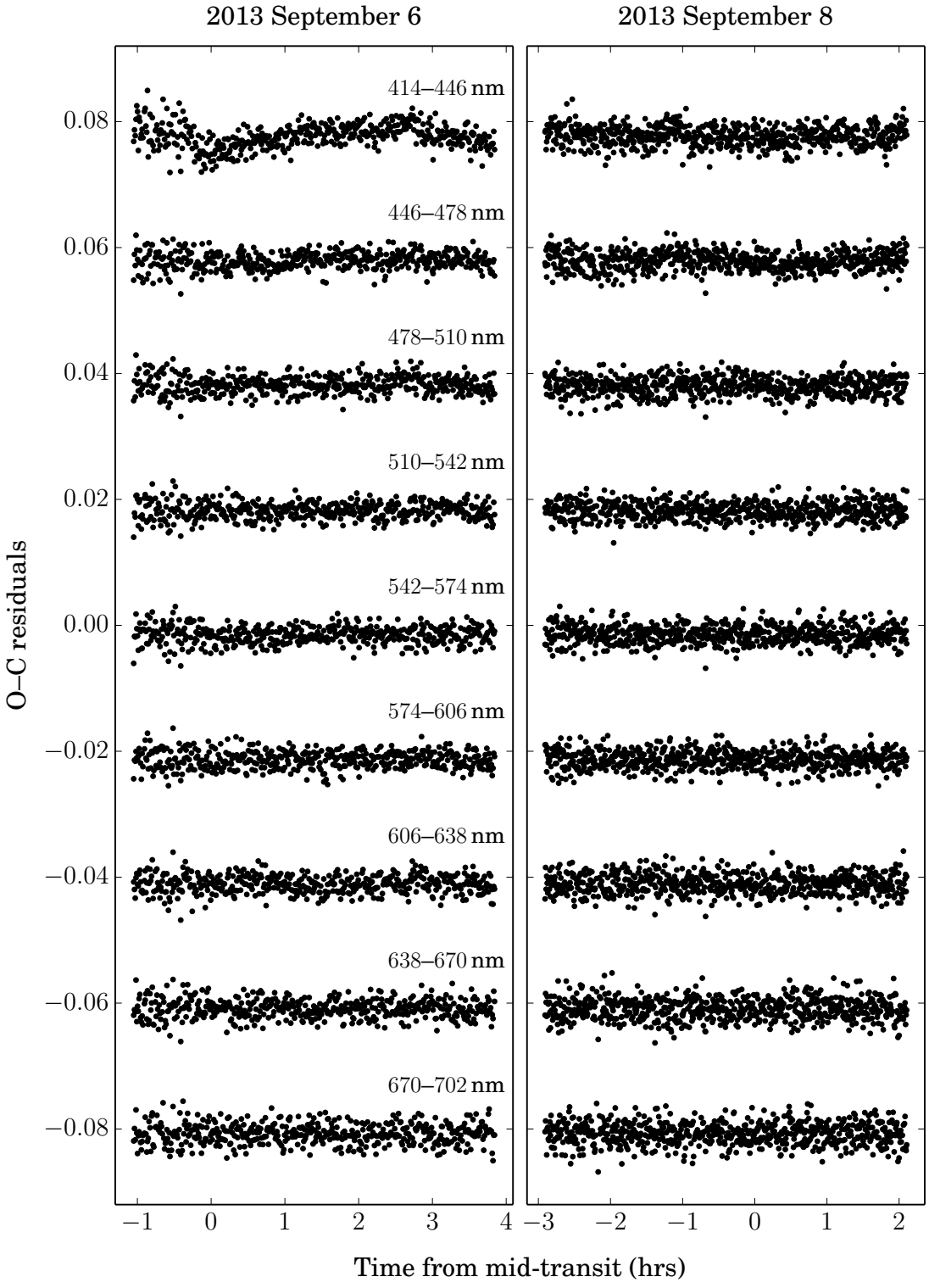

Figure 5.11: Observed-minus-calculated (O–C) residuals for the spectroscopic lightcurve fits shown in Figure 5.10.

Table 5.3: MCMC results for spectroscopic lightcurves. As for the white lightcurve analyses, the quadratic limb darkening parameter c_2 was fixed to the value giving the best-fit to an ATLAS stellar model Kurucz (1979, 1993) with similar properties to HAT-P-32 integrated over the appropriate bandpass. For the remaining parameters $\{R_p/R_\star\,,\,c_1\,,\,\sigma_{w,1}\,,\,\sigma_{w,2}\}$ the median of the MCMC chain is quoted, with uncertainties corresponding to the ranges either side of the medians that contain 34% of the chain samples.

Wavelength range (nm)	$R_p/R_\star$	c_1	c_2	$\sigma_{w,1}$ (ppm)	$\sigma_{w,2}$ (ppm)
414–446	$0.1519^{+0.0008}_{-0.0007}$	$0.47^{+0.01}_{-0.01}$	0.22	1844^{+65}_{-62}	1559^{+45}_{-44}
446–478	$0.1510^{+0.0006}_{-0.0006}$	$0.43^{+0.01}_{-0.01}$	0.29	1371^{+47}_{-44}	1449^{+41}_{-39}
478–510	$0.1510^{+0.0006}_{-0.0006}$	$0.40^{+0.01}_{-0.01}$	0.33	1404^{+48}_{-45}	1443^{+42}_{-39}
510–542	$0.1503^{+0.0006}_{-0.0006}$	$0.38^{+0.01}_{-0.01}$	0.33	1346^{+45}_{-43}	1343^{+38}_{-36}
542–574	$0.1491^{+0.0006}_{-0.0006}$	$0.33^{+0.01}_{-0.01}$	0.34	1392^{+47}_{-45}	1440^{+40}_{-39}
574–606	$0.1506^{+0.0006}_{-0.0006}$	$0.33^{+0.01}_{-0.01}$	0.35	1402^{+48}_{-45}	1414^{+40}_{-38}
606–638	$0.1522^{+0.0007}_{-0.0007}$	$0.27^{+0.01}_{-0.01}$	0.35	1445^{+50}_{-46}	1573^{+44}_{-42}
638–670	$0.1513^{+0.0007}_{-0.0007}$	$0.24^{+0.01}_{-0.01}$	0.36	1513^{+51}_{-49}	1719^{+49}_{-46}
670–702	$0.1491^{+0.0008}_{-0.0008}$	$0.24^{+0.02}_{-0.01}$	0.35	1639^{+54}_{-53}	1761^{+49}_{-47}

each were run. The results are reported in Table 5.3 and the best-fit models are shown in Figure 5.10. Overall, the fits do quite a good job of modelling the data. This can also be seen in Figure 5.11, which shows the fit residuals. The worst fit is obtained for the shortest wavelength channel (414–446 nm) on the first night, which is the most strongly affected by the high airmass systematics. The other lightcurves appear largely unaffected by visible residual correlations.

5.7 Discussion

The recovered transmission spectrum for HAT-P-32b is shown in Figure 5.12. A change in the effective radius equivalent to a single atmospheric scale height corresponds to a change in $R_p/R_\star$ of ~ 0.0012, which is approximately twice as large as the measured uncertainties (Table 5.3). In principle then, the transmission spectrum presented here can provide constraints on the wavelength-dependent

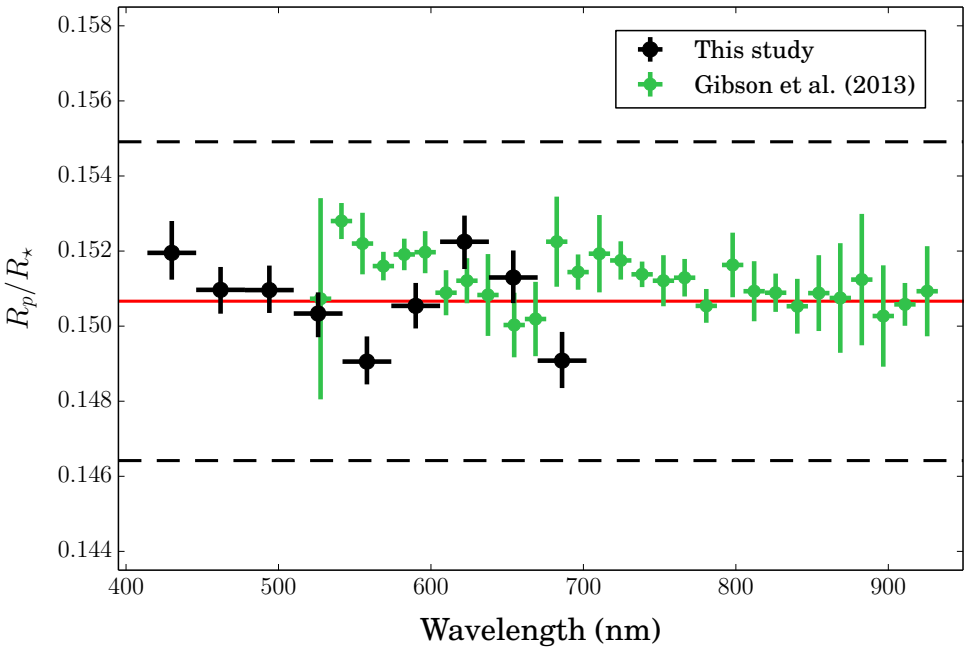

Figure 5.12: The combined transmission spectrum measured for HAT-P-32b over the 414–930 nm wavelength range. The horizontal red line shows the weighted mean of the measurements made in this study, $R_p/R_\star = 0.1507$. Horizontal dashed lines correspond to ±3 atmospheric scale heights about the weighted mean.

opacity of the atmosphere over the 32 nm-wide bandpasses of the spectroscopic lightcurves. In particular, the $R_p/R_\star$ values measured for the 542–574 nm and 670–702 nm wavelength channels are ~ 2–$3\,\sigma$ lower than the weighted mean for the full sample, $R_p/R_\star = 0.1507$. This creates the appearance of a broad opacity feature over the 550–700 nm wavelength range, where species such as Na and TiO/VO absorb (e.g. Fortney et al., 2010).

However, there are a number of caveats to be considered before invoking wavelength-dependent opacity to explain the variations in the transmission spectrum presented here. Firstly, a horizontal line fit to the data (i.e. a flat, featureless transmission spectrum) has a reduced χ^2 of 2; if the $R_p/R_\star$ uncertainties are uniformly increased by a factor of 1.5, the reduced χ^2 improves to 1. Indeed, it is quite possible that with correlated noise properly accounted for by a full GP analysis of the spectroscopic lightcurves, the uncertainties quoted in Table 5.3 will be revised upwards by such an amount. Secondly, there are a number of

discrepancies between the $R_p/R_\star$ results presented here and those previously published by Gibson et al. (2013b), which are also shown in Figure 5.12. The largest differences occur for the 542–574 nm and 670–702 nm wavelength channels. As was noted above, the $R_p/R_\star$ values obtained for these channels in the current study show the most significant departure from a flat transmission spectrum. The apparent disagreement with the corresponding values of Gibson et al. casts doubt on the reliability of this result.

However, as has been emphasised above, the transmission spectrum presented here should be considered a preliminary analysis. Further tests must be done before the results should be considered robust, and an attempt made to provide a physical interpretation. These include:

- varying basic data reduction parameters, such as the spectral aperture size, to determine how these influence the final results;
- independent analyses of the first and second night lightcurves, to check for consistency between the two;
- using full GP models for the spectroscopic lightcurve fits;
- repeating the lightcurve MCMC analyses with marginalisation over the covariance parameters as well as the mean function parameters;
- using different wavelength binnings for the spectroscopic lightcurves, including an investigation of narrower bins.

Also, as has been noted above in Section 5.4.3.3, the sky spectra extracted for each star show clear differences, despite the expectation that the background be constant across the FOV. It is desirable to understand what is causing these differences, and whether or not they are introducing biases to the sky subtraction that propagate through to the final estimates of the planet parameters.

For now, the lightcurve analyses reported in this chapter offer encouraging signs. The derived uncertainties for the spectroscopic $R_p/R_\star$ values are 600–800 ppm per 30 nm wavelength bin, compared with the typical uncertainties of 300–800 ppm per 15 nm wavelength bin obtained by Gibson et al. (2013b). This is quite impressive, given that Gibson et al. used the 8 m *Gemini-North* telescope to make their measurements as opposed to the 2.5 m *NOT* telescope used for the current study. It will be interesting to see how the uncertainties increase when the analysis is repeated with 15 nm wavelength channels. A naive estimation based on Poisson statistics would suggest that by reducing the number of photons per wavelength channel by half, the uncertainty on the $R_p/R_\star$ value would increase by a factor of $\sqrt{2} = 1.4$. However, given that $R_p/R_\star$ is constrained by the overall shape of the lightcurve, not just the depth of the transit, the increase in uncertainty may turn out to be smaller than this.

If it can be verified that the ALFOSC lightcurves presented here provide useful constraints for the transmission spectrum of HAT-P-32b, it will represent an important development in exoplanet atmospheric characterisation. Until now, successful ground-based observations of this type have predominantly been performed with heavily subscribed 6–8 m telescopes, such as *VLT* (e.g. Bean et al., 2010), *Gemini* (e.g. Gibson et al., 2013a,b; Stevenson et al., 2014), and *Magellan* (e.g. Bean et al., 2013). By demonstrating that useful observations can also be performed using modest-aperture telescopes such as the NOT, the field would be opened up to many more facilities. Indeed, a complementary approach may prove fuitful: a single transit observation performed with a 6–8 m telescope could be used to accurately constrain bulk system parameters, such as $a/R_\star$ and b, which could then be fixed during the analysis of multiple spectroscopic lightcurves acquired with smaller aperture telescopes.

Chapter 6

Summary and conclusions

This thesis has presented analyses of transit and eclipse lightcurves obtained with *Spitzer*/IRAC, *HST*/STIS, and *NOT*/ALFOSC for the hot Jupiters HD 189733b, HD 209458b, and HAT-P-32b. The aim was to use these measurements to infer information about the properties of the planetary atmospheres, including the re-evaluation of previous claims that have been made in the literature. Gaussian process (GP) models were used for these analyses, in order to provide a robust treatment of instrumental systematics within the Bayesian framework. The important results are summarised in Sections 6.1–6.3 below. A final comment on the application of GPs to lightcurve modelling is given in Section 6.4.

6.1 Chapter 3 – *Spitzer*/IRAC

Chapter 3 presented analyses of transit and eclipse measurements that have been made for HD 189733b and HD 209458b using *Spitzer*/IRAC. The main contribution of this study was the application of GPs to the lightcurve modelling, which give a more robust treatment of instrumental systematics than widely-used parametric methods, such as linear decorrelation. I obtain an overall improvement in consistency for the normalised semimajor axes $a/R_\star$ and orbital inclinations i across different epochs and wavelength channels, compared to a number of results that were previously published for these datasets. This provides

evidence that the GP models are more effectively accounting for the systematics, and less prone to underestimating uncertainties, compared with other lightcurve fitting approaches used in the literature.

For the HD 189733b emission data, a clear atmosphere model provides a much better fit to my revised IRAC data than an isothermal blackbody model. In particular, the blackbody model significantly under-predicts the 3.6μm emission and over-predicts the 4.5μm emission. The inability of an isothermal model to explain the emission data suggests that atmosphere layers of different temperatures are being probed by the four IRAC bandpasses, which in turn implies a wavelength-dependent atmosphere opacity. The clear atmosphere model fit demonstrates that one possibility is absorption by molecules such as H_2O and CO. An alternative possibility proposed by Pont et al. (2013) is that a thick layer of cloud/haze could be responsible for the discrepancy between the measured 3.6μm emission and the simple isothermal blackbody prediction, by producing a thermal inversion in the upper atmosphere. However, some doubt has been cast on this hypothesis by the inference of significant Na absorption in the reflection signal (see Section 6.2). Another notable result of my emission analysis is the downward revision of the 3.6 μm 2005 Nov 24 eclipse depth by 7.2σ from the value obtained by Charbonneau et al. (2008), bringing it much closer to the 3.6μm eclipse depths obtained for the 2010 Dec 27–30 phase curve.

For the HD 189733b transmission data, I do not verify the claimed detection of H_2O absorption made by Tinetti et al. (2007) based on the lightcurve analyses of Knutson et al. (2007a) and Beaulieu et al. (2008). My results are consistent with those of Désert et al. (2009, 2011b), who reported a tentative detection of CO absorption based on the deeper transit measured in the 4.5μm channel. However, when I apply the star spot corrections of Pont et al. (2013), my $R_p/R_\star$ values shift and the uncertainties increase, producing a transmission spectrum

that is consistent with being flat. I therefore conclude that the IRAC transmission data are unable to confidently identify absorption features in the atmosphere of HD 189733b.

For the HD 209458b emission data, I obtain results that are consistent with those published in the literature for all four channels, although my uncertainties are larger by factors of 1.2–1.8. The only exception is the 4.5μm 2005 Nov 28 dataset, for which I obtain an eclipse depth that is 2.7σ lower than the value reported by Knutson et al. (2008). My revised value is consistent with the depths obtained for the two other 4.5μm eclipses measured as part of the 2011 Jan 12–16 phase curve. Furthermore, it was the relatively deep eclipses measured in the 4.5μm and 5.8μm channels that were originally interpreted as evidence for a thermal inversion in the atmosphere of HD 209458b. The lower value I obtain for the 4.5μm eclipse and the larger uncertainty I obtain for the 5.8μm eclipse weaken this interpretation substantially. Indeed, simple model fits suggest that a non-inverted atmosphere provides a much better match to the revised emission data. Therefore, in contrast to earlier studies (e.g. Burrows et al., 2007), I conclude that the emission data does not imply a thermal inversion for the atmosphere of HD 209458b.

For the HD 209458b transmission data, I do not confirm the claim of H_2O absorption made by Beaulieu et al. (2010), which was based on deeper transit depths measured for the 5.8μm and 8.0μm channels relative to the 3.6μm and 4.5μm channels. Instead, I obtain transit depths across all four wavelength channels that are fit reasonably well by a flat spectrum. This could provide evidence for an opacity source that extends across the IRAC bandpasses, suppressing molecular absorption features. An opacity source along these lines has also been suggested by Deming et al. (2013), based on the smaller-than-expected amplitude of the H_2O feature measured over the 1–1.7μm wavelength range

with *HST*/WFC3. A clear atmosphere model was fit to the muted H_2O feature measured by Deming et al., but overpredicted the CO absorption in the 4.5μm channel relative to the transit depth measured for the 2010 Jan 19 lightcurve. This could indicate that the putative opacity source responsible for muting the H_2O feature over the 1–1.7μm wavelength range has an even stronger effect in the IRAC bandpasses.

Taken together, the revised results presented in this thesis illustrate how sensitive IRAC lightcurve analyses are to the systematics treatment. GP analyses have been shown to produce results that are generally more stable, and with uncertainties that are relatively conservative, compared to those obtained using other approaches. Combined with the lack of spectral resolution afforded by the broad bandpasses, this suggests that statements made previously in the literature about exoplanet atmospheres relying heavily on the interpretation of IRAC data should be regarded with caution.

6.2 Chapter 4 – *HST*/STIS

Chapter 4 presented the first ever multi-wavelength observation of an exoplanet secondary eclipse in reflected light. The observation was made for HD 189733b using *HST*/STIS over the 290–570 nm wavelength range. Geometric albedos of $A_g = 0.37^{+0.12}_{-0.13}$ and $A_g = 0.13^{+0.13}_{-0.12}$ were measured across the 290–450 nm and 450–570 nm wavelength ranges, respectively. I argued that stellar variability is unlikely to be responsible for the measured signal, but noted that it cannot be definitively ruled out. Assuming, however, that the planetary eclipse has been detected, I propose that the nonzero albedo measured for the 290–450 nm channel is due to incident stellar radiation that is scattered back to space by cloud/haze or H_2 molecules, while the lower albedo measured for the 450–570 nm channel is due to absorption by the broad wings of the Na 589 nm doublet. The detection of

broad Na absorption disfavours the hypothesis put forward by Pont et al. (2013), in which the dayside hemisphere of HD 189733b is dominated by a high-altitude layer of optically thick reflective cloud. However, the low precision and spectral resolution of the STIS data prevent a more detailed statement to be made about the atmosphere properties, as demonstrated by the retrieval study of Barstow et al. (2014). Nonetheless, the derived albedo constraints show that HD 189733b occupies a region of the colour-colour plane that is distinct from other solar system bodies with thick atmospheres, and that it would probably appear a deep blue colour to the human eye.

6.3 Chapter 5 – *NOT*/ALFOSC

Chapter 5 presented two transit observations for HAT-P-32b made in multi-object spectroscopy (MOS) mode with *NOT*/ALFOSC over the 414–702 nm wavelength range. Both white lightcurves were jointly analysed using GP models, giving planet parameters that are consistent with those reported previously by Hartman et al. (2011) and Gibson et al. (2013b). Spectroscopic lightcurves were also produced by binning the time series spectra into 32 nm-wide wavelength channels. Preliminary fits were made to the spectroscopic lightcurves, using quadratic trends in time to model the smooth systematics. The uncertainties on the radius ratios $R_p/R_\star$ inferred from these analyses were equivalent to approximately half an atmospheric scale height. The recovered transmission spectrum shows evidence for variations in the atmospheric opacity, possibly due to a species such as TiO or VO, which models predict may be present in the gas state for hot atmospheres such as HAT-P-32b (e.g. Fortney et al., 2010). However, this result is currently very tentative, and based entirely on the low radius ratios inferred for two of the nine wavelength channels (i.e. 542–574 nm and 670–702 nm; Table 5.3 and Figure 5.12). If the uncertainties have been underestimated by a factor

of 1.5, then the transmission spectrum is perfectly consistent with being flat. Previous experience (e.g. the IRAC analyses presented in Chapter 3) has shown that uncertainties can be underestimated by such an amount when lightcurve systematics are not rigorously accounted for. The spectroscopic lightcurves will therefore be re-analysed using a more rigorous GP treatment in the near future. On a more general note, the good quality of the lightcurves obtained in this study offer encouraging signs for attempting future exoplanet atmosphere characterisations using small, ground-based telescopes with MOS.

6.4 Gaussian process models

The lightcurve analyses presented in this work were performed using GPs to simultaneously model the planet signal and instrumental systematics. Under this approach, the model likelihood is defined as a multivariate normal distribution, with off-diagonal entries of the covariance matrix encoding the poorly understood systematics contribution. This offers a number of advantages compared to other approaches adopted in the literature. In particular, the use of covariance kernels such as the squared exponential (Equation 2.27) and Matérn (Equation 2.28) allow rich regions of function space to be marginalised with a small number of free parameters. Alternatively, if an explicit functional form is adopted for the systematics, it must either include a large number of parameters to account for the same range of behaviours, making full marginalisation intractable, or else it must be a simple function, such as a low order polynomial decorrelation. By restricting the range of function space that is marginalised over, the latter approach runs the risk of underestimating the uncertainties on model parameters. For instance, this might explain why the uncertainties that I derived using GPs are typically 1–5 times larger than those published previously for a number of the IRAC lightcurves.

The main disadvantage of the GP approach is that it is computationally expensive, with likelihood evaluation times scaling as $\mathcal{O}\left(N^3\right)$, where N is the number of data points. This restricts the use of GPs to lightcurves with $N \lesssim 1000$ data points. Methods exist, however, for approximating certain aspects of GP models, in order to increase the speed of likelihood evaluations (e.g. Quiñonero-Candela & Rasmussen, 2005). The application of fast GP methods will be investigated in future work.

Another improvement to the analyses presented here would be to use a method such as nested sampling (Section 2.2.2) to compute the model evidences. For instance, model evidences could be evaluated for a number of different GP covariance kernels, in order to make a principled decision on the most appropriate one to adopt for a given application.

In conclusion, GPs allow transit lightcurve models to be elegantly formulated within a rigorous Bayesian framework. They provide a natural mechanism for handling potentially complex correlations in the data that are unrelated to the astrophysical signal of interest. Uncertainty is rigorously propagated through all levels of the model in a clear and transparent manner, and Occam's razor is automatically implemented, mitigating against overfitting. GPs thus provide a flexible and robust means of handling uncertainty in transit lightcurves, making them well-suited to the task of atmosphere characterisation.

Appendix A

Spitzer/IRAC PSF and centroiding

This appendix describes the *Spitzer*/IRAC point source function (PSF) and three different centroiding algorithms. Tests are performed to compare the performance of the Gaussian profile fitting and flux-weighted mean centroiding algorithms, and the results presented.

A.1 IRAC point source function

The top row of Figure A.1 shows the IRAC point response function (PRF), as measured by the *Spitzer* Science Team in each of the four wavelength channels. The middle row shows the same, but plotted on a log colour scale to exaggerate the high contrast structure away from the core, including the diffraction spikes and Airy rings. The PRF is the convolution of the instrument point spread function (PSF) with the response function of a single pixel. To appreciate this more clearly, consider a star centered at detector coordinates (x_c, y_c) with total flux $F_\star$. The value z recorded at a pixel with center coordinates (x_i, y_j) following background subtraction is equal to:

$$z(x_i, y_j) = F_\star \int_{x_i-0.5}^{x_i+0.5} \int_{y_j-0.5}^{y_j+0.5} s(x'-x_i, y'-y_j)\, \Pi_S\left(x'-x_c, y-y_c\right)\, dy'\, dx' \quad \text{(A.1)}$$

where s is the response function of the pixel and Π_S is the instrumental PSF. Since $s = 0$ for $|x - x_i| > 0.5$ and $|y - y_j| > 0.5$, the above equation is equivalent

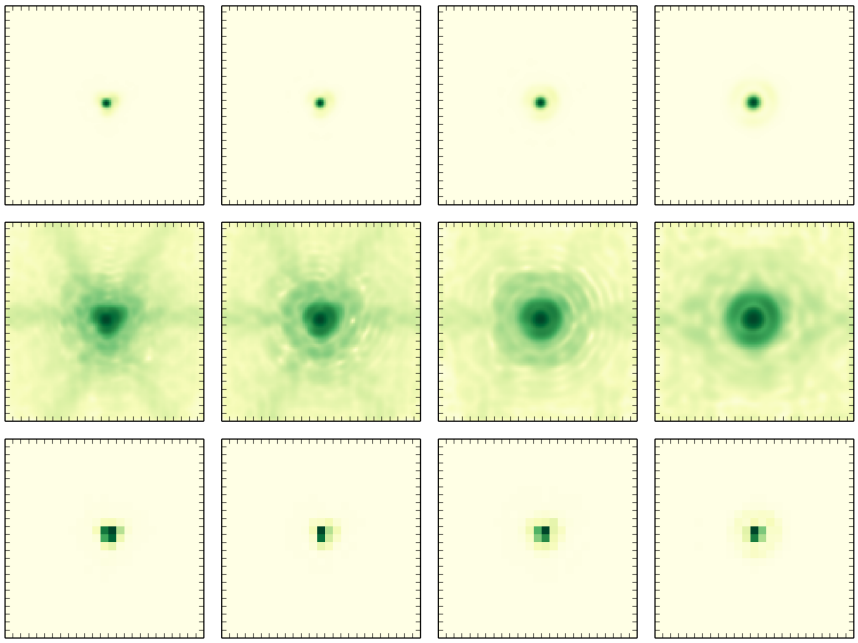

Figure A.1: (Top row) IRAC PRF for each wavelength channel with linear colour scale, at a resolution 5× the native pixel resolution. (Middle row) The same as above, but with log colour scale to emphasise fine-scale structure. (Bottom row) IRAC PSF for each wavelength channel with linear colour scale, obtained by sampling the super-resolution PRFs at the native pixel resolution. For all axes, tick marks indicate pixel boundaries.

to:

$$z(\Delta x, \Delta y) = F_\star \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} s(x, y)\, \Pi_{\rm S}(\Delta x - x, \Delta y - y)\, dy\, dx$$

$$= F_\star\, \Pi_{\rm R}(\Delta x, \Delta y)\,, \tag{A.2}$$

where we have made the substitutions $x = x_i - x'$, $y = y_j - y'$, $\Delta x = x_i - x_c$ and $\Delta y = y_j - y_c$, and $\Pi_{\rm R}$ denotes the PRF. Thus, we see that the PRF gives the fraction of stellar light integrated across a full pixel as a function of the pixel center coordinates relative to the stellar PSF centroid. In practice, the PRF is obtained by measuring how pixel values for one or more point sources vary as a function of a sub-pixel dither pattern. For further details, see Anderson & King (2000).

The bottom row of Figure A.1 shows the PSF, obtained by sampling the PRF to

the native detector resolution in accordance with Equation A.2. From the latter, it is evident that for each wavelength channel, most of the flux is concentrated in only ~3–4 pixels. Indeed, the measured full-width at half maximum (FWHM) is 1.6″, 1.6″, 1.8″, and 1.9″ for the 3.6μm, 4.5μm, 5.8μm, and 8.0μm channels, respectively. Given the pixel scale of $\sim 1.2''\,\mathrm{pixel}^{-1}$, these FWHMs translate to 1.3 pixels, 1.3 pixels, 1.5 pixels, and 1.6 pixels, respectively. The IRAC PSF is therefore undersampled in all four channels, posing a significant challenge to accurate centroiding and photometry.

A.2 Centroiding algorithms

A.2.1 Gaussian profile fitting

The idealised PSF of a circular aperture is a squared Bessel function, recognisable as the well-known Airy disk pattern. The core of such a PSF can be well-approximated by a 2D Gaussian function in xy. It follows that the marginal distributions along the x and y axes are 1D Gaussians with means at x_c and y_c, respectively. This property holds regardless of the xy covariance in the 2D PSF distribution. The centroid coordinates (x_c, y_c) can therefore be determined by fitting 1D Gaussian profiles to the marginal pixel distributions using least squares.

A.2.2 Flux-weighted mean

Analogous to the center-of-mass for a system, it is possible to determine the center-of-light for an array of pixels, which can be referred to as the flux-weighted mean. If z_i denotes the value of the pixel with coordinates (x_i, y_i), the centroid coordinates are calculated according to:

$$x_c = \frac{\sum_i z_i\, x_i}{\sum_i z_i} \quad , \quad y_c = \frac{\sum_i z_i\, y_i}{\sum_i z_i} \, . \tag{A.3}$$

Before applying Equation A.3, it is important to subtract the background, otherwise the background pixels will bias the result.

A.2.3 PSF fitting

If an accurate, high-resolution PSF model is available, it can be shifted in x and y until the closest match to the measured pixel values is identified. In addition to the centroid coordinates(x_c, y_c), a scaling constant for the PSF amplitude and an additive offset for the background are allowed to vary. At each step in the optimisation, the high-resolution PSF must be digitized to the native resolution of the detector before computing the residuals. In theory, such an approach is possible for centroiding IRAC datasets, as super-resolution PRFs have been made available by the *Spitzer* Science Team (Section A.1). However, Ballard et al. (2010) concluded that the performance of PSF fitting was relatively inferior compared to simpler methods, such as flux-weighted mean, when applied to IRAC data. For this reason, I do not consider PSF fitting any further.

A.3 Comparison of centroiding algorithms

Most of the published analyses of IRAC datasets in the exoplanet field have employed standard Interactive Data Language (IDL) routines for computing stellar centroids. In particular, the `gcntrd.pro` routine of the IDL Astronomy User's Library[1] is used to implement Gaussian profile fitting (Section A.2.1), and the `box_centroider.pro` routine made available online by the *Spitzer* Science Team[2] implements the flux-weighted mean (Section A.2.2). In this section, I compare the performance of these widely-used IDL routines with the equivalent `pyirac` versions, to verify the latter.

[1] `http://idlastro.gsfc.nasa.gov/ftp/pro/idlphot/gcntrd.pro`
[2] `http://irsa.ipac.caltech.edu/data/SPITZER/docs/irac`

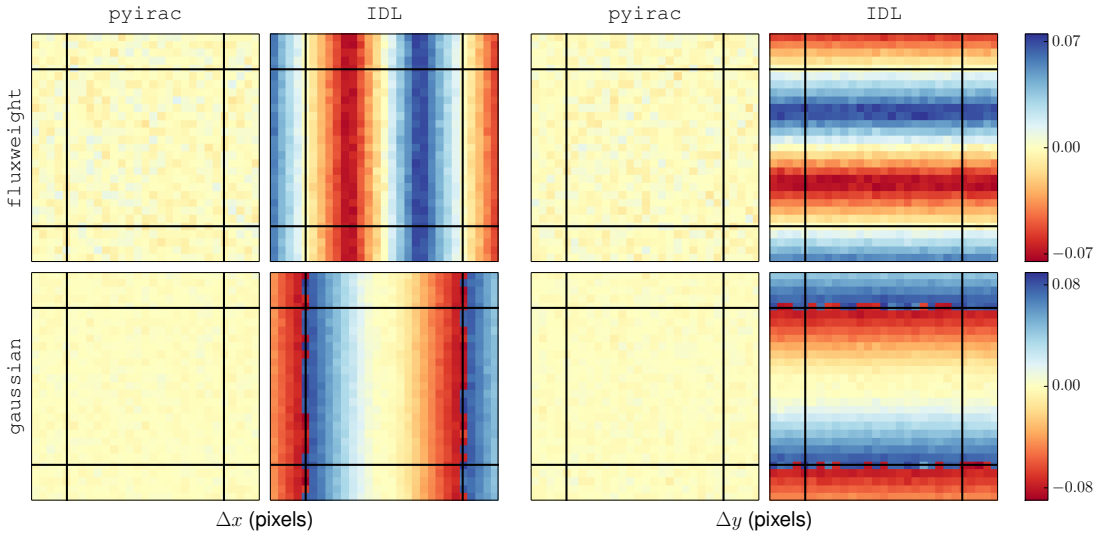

Figure A.2: Errors in centroid estimates as a function of the true centroid location for the case of an idealised Gaussian PSF. Black lines show pixel boundaries. The colour scale indicates the centroid error in units of pixels. First and second columns give x-coordinate errors for the `pyirac` and IDL routines, respectively, and the third and fourth columns give the y-coordinate errors.

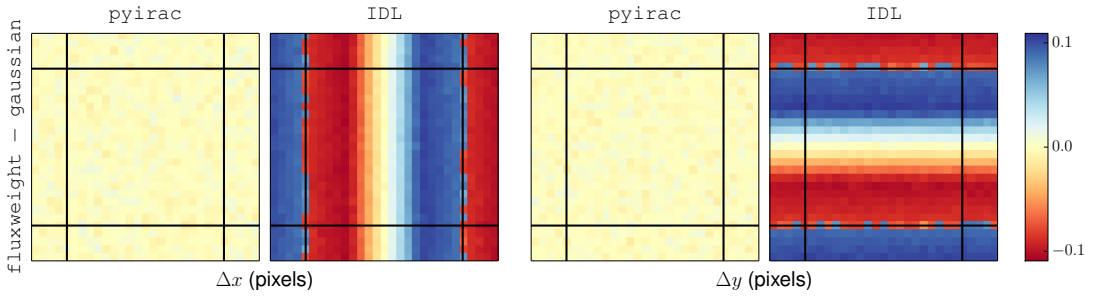

Figure A.3: Similar to Figure A.2, but showing discrepancies between centroid coordinates estimated using the Gaussian and flux-weighted centroid methods, as applied to the idealised Gaussian PSF. Notably, the discrepancies between the IDL implementations of the two methods are significantly larger than those for the `pyirac` implementations.

To start with, I considered an idealised case in which the PSF was given by a 2D axisymmetric Gaussian with a FWHM of 1.4 pixels, close to the true value of the IRAC PSF. The center of the idealised Gaussian PSF was shifted to each point on a 20×20 rectangular grid spanning a single pixel, and binned down to the native pixel resolution on an array of 10×10 pixels. A constant background level corresponding to 0.03% of the PSF peak was added to all pixels. The resulting 400 datasets were then duplicated 100 times, with random errors added to each pixel in accordance with the photon noise properties expected for a bright stellar target. This gave 4×10^4 simulated datasets in total, and the `pyirac` and `IDL` centroiding algorithms were run on each of them.

Figure A.2 shows the median errors in the estimated centroids as a function of the true centroid location for each algorithm. The `pyirac` routines are found to perform substantially better than the equivalent `IDL` routines, with the latter exhibiting errors of up to ~ 0.08 pixels in both the x and y centroid coordinates. The superior performance of the `pyirac` routines are also demonstrated in Figure A.3, which shows the difference between the centroid coordinates calculated using Gaussian profile fitting and the flux-weighted mean. There is good agreement between the two methods for the `pyirac` implementations (within 0.01 pixels), whereas the `IDL` routines systematically disagree by up to ~ 0.1 pixels.

The same analysis was repeated using the IRAC PRFs shown in Figure A.1 interpolated to a grid of 100×100 points per pixel. The true centroid was estimated as the flux-weighted mean of the super-resolution PRF according to Equation A.3, and rounded to the nearest 0.01 pixel.[3] The PRF was then shifted to each point on a 20×20 grid spanning a single pixel, and resampled to the native resolution on an array of 10×10 pixels. Background offsets of 0.01%, 0.02%, 0.03%, and 0.21% the peak stellar flux were added for the 3.6μm, 4.5μm,

[3] Although not strictly equal to the centroid of the PSF, it is a good approximation and at worst will introduce a constant offset to the centroid errors estimated subsequently.

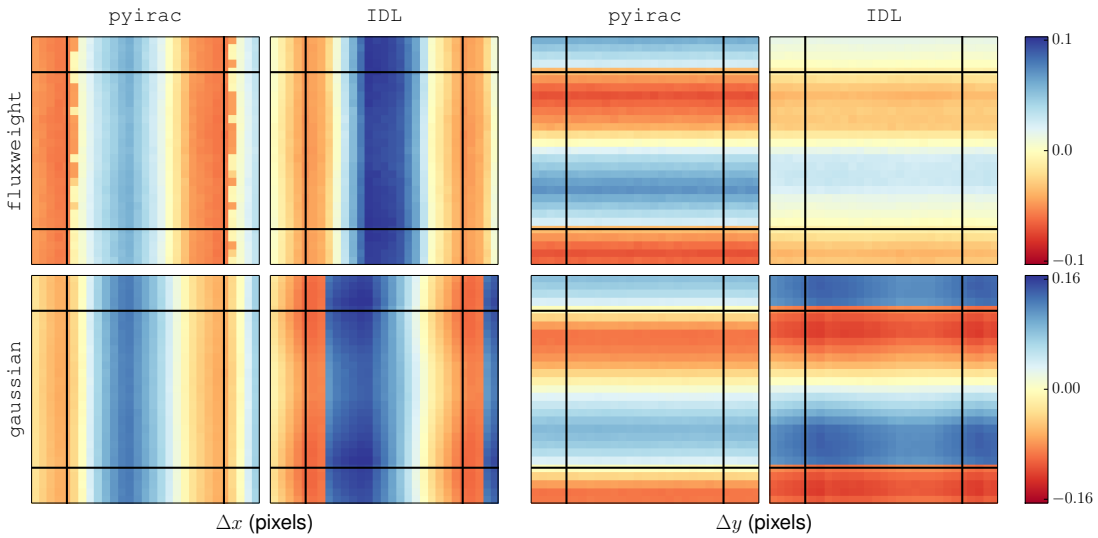

Figure A.4: Same as Figure A.2, except for the IRAC 3.6μm PSF instead of the idealised Gaussian PSF. The asymetry of the IRAC PSF results in larger centroid errors compared with those obtained for the idealised Gaussian PSF.

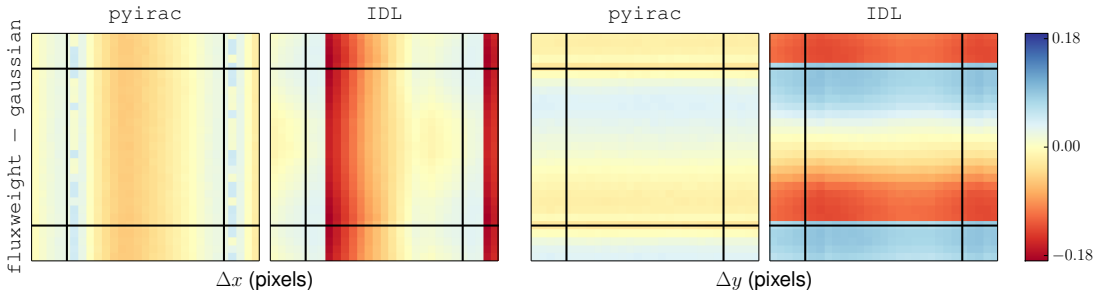

Figure A.5: Same as Figure A.3, except for the IRAC 3.6μm PSF. Again, the discrepancies in centroid estimates are larger overall for the IDL methods compared to the `pyirac` methods, but this difference is less pronounced than it is for the case of the idealised Gaussian PSF.

5.8μm, and 8.0μm wavelength channels, respectively. These values are typical of those encountered in the archival datasets described in Section 3.3. As with the idealised Gaussian PSF, the resulting 400 datasets for each wavelength channel were duplicated 100 times with random photon noise errors added to each pixel, and the centroiding algorithms were applied.

Figure A.4 shows the median errors in the estimated centroids for the 3.6μm channel PSF, and the differences between the two centroiding methods are plotted in Figure A.5. Similar results were obtained for the other three wavelength channels, but are not plotted. In all channels, I find that the `pyirac` and `IDL` routines give comparable results, and the flux-weighted mean gives lower errors than the Gaussian profile fits overall. For instance, the errors in the 3.6μm channel are up to ~ 0.10 pixels for the flux-weighted mean calculations, and up to ~ 0.15 pixels for the Gaussian profile fits.

I have therefore verified that the `pyirac` implementations of the Gaussian profile fitting and flux-weighted mean centroiding methods perform at a level comparable to, or better than, the equivalent `IDL` routines that are used widely in the community. My tests also suggest that the flux-weighted mean method is generally better-suited than Gaussian profile fitting for handling the undersampled IRAC PSF. I note that both methods can give systematic errors of 0.1 pixels, depending on where the PSF is centered within a pixel. However, these errors tend to be reasonably constant over significant fractions of a pixel surface. Therefore, if the pointing drift is only ~0.1 pixels over the course of an observation, the *relative* centroid errors will likely be smaller (~0.01 pixels).

Appendix B

Spitzer/IRAC lightcurve covariance parameters

This appendix contains the maximum likelihood estimates (MLEs) for the GP covariance parameters of the *Spitzer*/IRAC lightcurve models described in Chapter 3. Table B.1 gives values for the 3.6μm and 4.5μm channels, and Table B.2 gives values for the for the 5.8μm and 8.0μm channels. These are the values that the covariance parameters were fixed to for the MCMC analysis described in Section 3.7.4.

Table B.1: Covariance parameter MLEs for the 3.6μm and 4.5μm lightcurves.

Channel (μm)	Target	Date	Signal	A_{xy} (ppm)	L_x (pix)	L_y (pix)	A_t (ppm)	L_t (min)	β
3.6	HD 189733	2005 Nov 24	Eclipse	17293	0.423	0.319	582	110.4	1.124
		2006 Oct 31	Transit	9436	0.247	0.155	0	$> 10^5$	1.112
		2007 Nov 25	Transit	2948	0.121	0.053	100	9.4	1.336
		2010 Dec 27	Eclipse	23138	0.223	0.213	344	22.6	1.595
		2010 Dec 29	Transit	22724	0.305	0.242	108	7.0	1.298
		2010 Dec 30	Eclipse	8943	0.129	0.102	250	3.6	1.483
	HD 209548	2005 Nov 28	Transit	24842	0.721	0.483	708	435.2	1.254
		2007 Dec 31	Transit	3657	0.125	0.146	412	$> 10^5$	1.158
		2008 Jul 19	Transit	2026	0.095	0.083	15	$> 10^5$	1.137
		2011 Jan 12	Eclipse	20688	0.242	0.226	263	4.4	1.384
		2011 Jan 14	Transit	9663	0.153	0.191	1372	37.5	4.370
		2011 Jan 16	Eclipse	18813	0.234	0.221	352	8.2	1.466
4.5	HD 189733	2005 Nov 24	Eclipse	22604	0.635	0.486	1227	278.2	1.150
		2007 Nov 23	Transit	316	0.045	6.895	413	31.6	1.045
		2009 Dec 22	Eclipse	5729	0.173	0.183	86	4.2	1.089
		2009 Dec 23	Transit	10432	0.279	0.278	79	10.0	1.145
		2009 Dec 24	Eclipse	10936	0.545	0.252	180	15.7	1.502
	HD 209548	2005 Nov 28	Eclipse	4477	0.147	0.177	3909	2222.0	1.317
		2008 Jul 22	Transit	2059	82.343	0.123	1161	679.8	1.256
		2010 Jan 18	Eclipse	7707	0.232	0.195	223	13.3	1.134
		2010 Jan 19	Transit	11247	0.313	0.239	213	$> 10^5$	1.274
		2010 Jan 21	Eclipse	5849	0.172	0.197	194	16.8	1.251

Table B.2: Covariance parameter MLEs for the 5.8μm and 8.0μm lightcurves.

Channel (μm)	Target	Date	Signal	A_τ (ppm)	L_τ	c_0	c_1 (min)	A_t (ppm)	L_t (min)	β
5.8	HD 189733	2005 Nov 24	Eclipse	6510	302.4	−0.155	292.2	0	$> 10^5$	1.502
		2006 Oct 31	Transit	7004	341.0	0.323	1349.9	0	$> 10^5$	1.123
	HD 209458	2005 Nov 28	Eclipse	554	1.5	709.878	58024.5	711	87.4	1.781
		2007 Dec 31	Transit	2030	1212.9	2.938	33590.2	0	$> 10^5$	1.070
		2008 Jul 22	Transit	1837	3555.6	5.957	2117.1	13	$> 10^5$	1.125
8.0	HD 189733	2005 Nov 24	Eclipse	5403	1292.8	1.779	1454.5	1482	155.3	1.176
		2006 Oct 29	Transit	6494	1512.4	1.332	7611.6	99	25.5	1.127
		2006 Oct 30	Eclipse	83	0.7	0.045	21024.9	117	240.6	1.065
		2007 Nov 23	Transit	8618	934.0	−1.574	818.1	101	$> 10^5$	1.201
	HD 209458	2005 Nov 28	Eclipse	1799	85.1	−0.153	4648.1	26	$> 10^5$	1.674
		2007 Dec 24	Transit	3158	1787.8	−2.524	15225.7	198	10.1	1.188
		2007 Dec 25	Eclipse	131	0.4	0.009	21753.8	124	3.9	1.232
		2008 Jul 22	Transit	2787	167.9	0.424	15242.5	0	$> 10^5$	1.157

Appendix C

HST/STIS linear basis model tests

Before performing the GP modelling of the HD 189733b STIS lightcurve presented in Chapter 4, a number of tests were performed using linear basis models to identify the relevant input variables and verify the robustness of the signal. For these tests, the lightcurve models took the form:

$$\mu = c_0 + \sum_{i=1}^{n_\phi} a_i \phi^i + b_1 t + b_2 \psi + b_3 x + b_4 y + f_0 B \,, \tag{C.1}$$

where the c_0, a_i, b_i, and f_0 terms are the unknown coefficients of the linear model; ϕ is the *HST* orbital phase; t is time; ψ is the trace slope (Equation 4.1); x and y are the trace shifts along the dispersion and cross-dispersion axes, respectively; and B is the boxcar function for the eclipse (Equation 4.3).

Tests were performed by solving for the unknown coefficients using linear least squares. Values of $n_\phi = 3, 4, 5$ were tested for the *HST* orbital phase polynomial, with different combinations of linear terms for the remaining inputs. Table C.1 reports the combinations that gave the lowest Bayesian information criterion (BIC) (Equation 3.6) values for the two-channel lightcurves (i.e. the 290–450 nm and 450–570 nm channels), along with the fitted eclipse depths. Uncertainties for the latter were estimated by randomly permuting the model residuals 1000 times with replacement, repeating the fit for each of these synthetic datasets, and taking the standard deviation of the resulting eclipse depths. The results are

Table C.1: Linear decorrelations that give the lowest BICs for the two-channel binning. Polynomial orders used for the decorrelations are listed in the columns labelled n_ϕ, n_ψ, n_t, n_x, and n_y, with the subscript indicating the corresponding auxiliary variable. Results are divided between those obtained without a linear trend in time (i.e. $n_t = 0$) and those with a linear trend in time (i.e. $n_t = 1$). Eclipse depth uncertainties were estimated using residual permutation.

290–450 nm channel							450–570 nm channel						
n_ϕ	n_t	n_ψ	n_x	n_y	δ_s (ppm)	BIC	n_ϕ	n_t	n_ψ	n_x	n_y	δ_s (ppm)	BIC
4	0	1	0	0	121 ± 32	195.92	4	0	1	0	0	-21 ± 20	234.52
4	0	1	0	1	123 ± 31	197.01	4	0	1	1	0	10 ± 29	234.75
5	0	1	0	0	121 ± 31	200.07	5	0	1	1	0	14 ± 27	235.26
4	0	1	1	0	116 ± 46	200.47	5	0	1	0	0	-20 ± 20	236.06
4	0	1	1	1	123 ± 30	201.23	4	0	1	0	1	-21 ± 20	239.07
4	1	1	0	0	107 ± 36	199.51	4	1	1	0	0	-63 ± 23	213.05
4	1	1	0	1	107 ± 36	200.39	5	1	1	0	0	-63 ± 21	213.79
5	1	1	0	0	72 ± 60	203.32	4	1	1	1	0	-80 ± 36	217.04
4	1	1	1	0	107 ± 35	203.60	4	1	1	0	1	-63 ± 22	217.64
5	1	1	0	1	73 ± 61	204.21	5	1	1	1	0	-76 ± 38	218.09

divided between those that included a linear t term and those that did not. In all cases, $n_\phi = 4$ is favoured by the BIC. For the 290–450 nm channel, the lowest BIC overall includes an additional linear ψ term, and gives $\delta_s = 121 \pm 32$ ppm. The lowest BIC for the 450–570 nm channel is obtained with additional linear terms in both ψ and t, and gives $\delta_s = -60 \pm 23$ ppm. This corresponds to a negative eclipse depth at the 2.6σ level, which is unphysical. It would seem that the inclusion of a linear t dependence biases the inferred eclipse depths to lower values, and for this reason, I favour the results that do not include a t dependence. Of the latter, the lowest BIC for the 450–570 nm channel is obtained with a linear ψ term only, and gives $\delta_s = -20 \pm 21$ ppm. It is also reassuring that for both channels, all the results quoted in Table C.1 (i.e. corresponding to the lowest BICs) for the case of no t dependence are consistent with each other.

Before proceeding to the GP analysis, a number of other tests were performed using the linear decorrelation models. First, the data set was divided into two halves: the first consisting of every odd-numbered data point ($i = 1, 3, 5, \ldots$) and

Table C.2: Results of linear decorrelation tests. Second column ('Decorr') gives the *HST* orbits that were used to compute the linear decorrelation. Third column ('In') gives the *HST* orbit that was assumed to be in-eclipse. Fourth column ('Out') gives the *HST* orbits that were assumed to be out-of-eclipse when estimating the eclipse depth. Fifth and sixth columns give the inferred eclipse depths for the 290–450 nm and 450–570 nm wavelength channels, respectively.

	HST orbits			Inferred δ_s (ppm)	
Test #	Decorr.	In	Out	290–450 nm	450–570 nm
1	1, 2, 3	1	3	-20 ± 36	-13 ± 21
2	1, 3	1	3	2 ± 36	-2 ± 21
3	1, 2, 3	2	1, 3	120 ± 31	-20 ± 19
4	1, 2, 3	2	1	130 ± 35	-14 ± 26
5	1, 2, 3	2	3	107 ± 36	-27 ± 23
6	1, 2, 3	3	1, 2	-45 ± 32	20 ± 20
7	1, 2, 3	1	2, 3	-75 ± 34	1 ± 20
8	1, 3	1	2, 3	-59 ± 31	8 ± 20
9	1, 3	3	1, 2	-63 ± 34	11 ± 20

the other consisting of every even-numbered data point ($i = 2,4,6,\ldots$). These subsets were then fit separately, and gave eclipse depths that were consistent with those listed in Table C.1, but with larger uncertainties commensurate with the fact that they comprised half the number of data points in each lightcurve. Second, random subsets of 60 datapoints were sampled and fits were made to these samples. This process was repeated multiple times, in all cases giving inferred eclipse depths that were consistent with those listed in Table C.1. These tests provide evidence that the inferred eclipse depths are not disproportionately influenced by a small subset of anomalous data points.

Tests were also performed to determine how the inferred eclipse depth varied according to which data points were assumed to be in-eclipse and out-of-eclipse. For instance, we assumed that *HST* orbits 1 and 2 were out-of-eclipse and orbit 3 was in-eclipse, and adjusted the definition for the B function given by Equation 4.3 accordingly. The results of these tests are reported in Table C.2. In summary:

- Tests 1–2 show that if we assume orbit 1 is in-eclipse and orbit 3 is out-of-

eclipse, then the inferred eclipse depth is consistent with zero, regardless of whether or not orbit 2 is included in the systematics decorrelation.

- Tests 3–5 show that consistent eclipse depths are inferred when we using orbits 1, 2, and 3 for the systematics decorrelation and assume orbit 2 is in-eclipse, regardless of whether we use orbit 1 or orbit 3 or both for the out-of-eclipse.

- Tests 6–9 show that if we include orbit 2 as out-of-eclipse and use either orbit 1 or 3 as in-eclipse, the inferred eclipse depth in the 290–450 nm channel is always negative while in the 450–570 nm channel it remains consistent with zero. These results are obtained regardless of whether or not orbit 2 is included in the systematics decorrelation, and are what would be expected if a genuine eclipse with $\delta_s > 0$ occurs during orbit 2 for the 290–450 nm channel. This is because the flux reduction during orbit 2 caused by the eclipse would decrease the estimated baseline flux level, and cause a negative eclipse depth to be inferred for orbit 1 or 3.

The ensemble of tests presented in this appendix suggest that the STIS lightcurve is amenable to robust decorrelation. This gives confidence in the eclipse measurements and albedo constraints that are reported in Chapter 4.

References

Agol, E., Cowan, N. B., Knutson, H. A., Deming, D., Steffen, J. H., Henry, G. W., & Charbonneau, D. 2010, ApJ, 721, 1861

Albrecht, S., Winn, J. N., Johnson, J. A., Howard, A. W., Marcy, G. W., Butler, R. P., Arriagada, P., Crane, J. D., Shectman, S. A., Thompson, I. B., Hirano, T., Bakos, G., & Hartman, J. D. 2012, ApJ, 757, 18

Allison, R. & Dunkley, J. 2014, MNRAS, 437, 3918

Alonso, R., Brown, T. M., Torres, G., Latham, D. W., Sozzetti, A., Mandushev, G., Belmonte, J. A., Charbonneau, D., Deeg, H. J., Dunham, E. W., O'Donovan, F. T., & Stefanik, R. P. 2004, ApJ, 613, L153

Alonso, R., Guillot, T., Mazeh, T., Aigrain, S., Alapini, A., Barge, P., Hatzes, A., & Pont, F. 2009, A&A, 501, L23

Anderson, D. R., Gillon, M., Maxted, P. F. L., Barman, T. S., Collier Cameron, A., Hellier, C., Queloz, D., Smalley, B., & Triaud, A. H. M. J. 2010, A&A, 513, L3

Anderson, J. & King, I. R. 2000, PASP, 112, 1360

Auvergne, M., Bodin, P., Boisnard, L., Buey, J.-T., Chaintreuil, S., Epstein, G., Jouret, M., Lam-Trong, T., Levacher, P., Magnan, A., Perez, R., Plasson, P., Plesseria, J., Peter, G., Steller, M., Tiphène, D., Baglin, A., Agogué, P., Appourchaux, T., Barbet, D., Beaufort, T., et al. 2009, A&A, 506, 411

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

Ballard, S., Charbonneau, D., Deming, D., Knutson, H. A., Christiansen, J. L., Holman, M. J., Fabrycky, D., Seager, S., & A'Hearn, M. F. 2010, PASP, 122, 1341

Barman, T. 2007, ApJ, 661, L191

Barman, T. S., Hauschildt, P. H., & Allard, F. 2001, ApJ, 556, 885

Barstow, J. K., Aigrain, S., Irwin, P. G. J., Fletcher, L. N., & Lee, J.-M. 2013, MNRAS, 434, 2616

Barstow, J. K., Aigrain, S., Irwin, P. G. J., Hackler, T., Fletcher, L. N., Lee, J. M., & Gibson, N. P. 2014, ApJ, 786, 154

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

Bean, J. L., Désert, J.-M., Kabath, P., Stalder, B., Seager, S., Miller-Ricci Kempton, E., Berta, Z. K., Homeier, D., Walsh, S., & Seifahrt, A. 2011, ApJ, 743, 92

Bean, J. L., Désert, J.-M., Seifahrt, A., Madhusudhan, N., Chilingarian, I., Homeier, D., & Szentgyorgyi, A. 2013, ApJ, 771, 108

Bean, J. L., Miller-Ricci Kempton, E., & Homeier, D. 2010, Nature, 468, 669

Beaulieu, J. P., Carey, S., Ribas, I., & Tinetti, G. 2008, ApJ, 677, 1343

Beaulieu, J. P., Kipping, D. M., Batista, V., Tinetti, G., Ribas, I., Carey, S., Noriega-Crespo, J. A., Griffith, C. A., Campanella, G., Dong, S., Tennyson, J., Barber, R. J., Deroo, P., Fossey, S. J., Liang, D., Swain, M. R., Yung, Y., & Allard, N. 2010, MNRAS, 409, 963

Benneke, B. & Seager, S. 2012, ApJ, 753, 100

Berdyugina, S. V., Berdyugin, A. V., Fluri, D. M., & Piirola, V. 2008, ApJ, 673, L83

—. 2011, ApJ, 728, L6

Berta, Z. K., Charbonneau, D., Désert, J.-M., Miller-Ricci Kempton, E., McCullough, P. R., Burke, C. J., Fortney, J. J., Irwin, J., Nutzman, P., & Homeier, D. 2012, ApJ, 747, 35

Bessel, F. W. 1844, MNRAS, 6, 136

Birkby, J. L., de Kok, R. J., Brogi, M., de Mooij, E. J. W., Schwarz, H., Albrecht, S., & Snellen, I. A. G. 2013, MNRAS, 436, L35

Bishop, C. M. 2006, Pattern Recognition and Machine Learning (Secaucus, NJ, USA: Springer-Verlag New York, Inc.)

Blecic, J., Harrington, J., Madhusudhan, N., Stevenson, K. B., Hardy, R. A., Cubillos, P. E., Hardin, M., Bowman, O., Nymeyer, S., Anderson, D. R., Hellier, C., Smith, A. M. S., & Collier Cameron, A. 2014, ApJ, 781, 116

Blecic, J., Harrington, J., Madhusudhan, N., Stevenson, K. B., Hardy, R. A., Cubillos, P. E., Hardin, M., Campo, C. J., Bowman, W. C., Nymeyer, S., Loredo, T. J., Anderson, D. R., & Maxted, P. F. L. 2013, ApJ, 779, 5

Bond, I. A., Udalski, A., Jaroszyński, M., Rattenbury, N. J., Paczyński, B., Soszyński, I., Wyrzykowski, L., Szymański, M. K., Kubiak, M., Szewczyk, O., Żebruń, K., Pietrzyński, G., Abe, F., Bennett, D. P., Eguchi, S., Furuta, Y., Hearnshaw, J. B., Kamiya, K., Kilmartin, P. M., Kurata, Y., Masuda, K., Matsubara, Y., Muraki, Y., Noda, S., Okajima, K., Sako, T., Sekiguchi, T., Sullivan, D. J., Sumi, T., Tristram, P. J., et al. 2004, ApJ, 606, L155

Borucki, W. J., Koch, D., Basri, G., Batalha, N., Brown, T., Caldwell, D., Caldwell, J., Christensen-Dalsgaard, J., Cochran, W. D., DeVore, E., Dunham, E. W., Dupree, A. K., Gautier, T. N., Geary, J. C., Gilliland, R., Gould, A., Howell, S. B., Jenkins, J. M., Kondo, Y., Latham, D. W., Marcy, G. W., Meibom, S., Kjeldsen, H., Lissauer, J. J., Monet, D. G., Morrison, D., Sasselov, D., Tarter, J., Boss, A., Brownlee, D., et al. 2010, Science, 327, 977

Borucki, W. J., Koch, D., Jenkins, J., Sasselov, D., Gilliland, R., Batalha, N., Latham, D. W., Caldwell, D., Basri, G., Brown, T., Christensen-Dalsgaard, J., Cochran, W. D., DeVore, E., Dunham, E., Dupree, A. K., Gautier, T., Geary, J., Gould, A., Howell, S., Kjeldsen, H., Lissauer, J., Marcy, G., Meibom, S., Morrison, D., & Tarter, J. 2009, Science, 325, 709

Boss, A. P. 1995, Science, 267, 360

Bouchy, F., Queloz, D., Deleuil, M., Loeillet, B., Hatzes, A. P., Aigrain, S., Alonso, R., Auvergne, M., Baglin, A., Barge, P., Benz, W., Bordé, P., Deeg, H. J., de La Reza, R., Dvorak, R., Erikson, A., Fridlund, M., Gondoin, P., Guillot, T., Hébrard, G., Jorda, L., Lammer, H., Léger, A., Llebaria, A., Magain, P., Mayor, M., Moutou, C., Ollivier, M., Pätzold, M., Pepe, F., et al. 2008, A&A, 482, L25

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

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

—. 2003, ApJ, 593, L125

Brown, T. M. & Charbonneau, D. 1999, in Bulletin of the American Astronomical Society, Vol. 31, American Astronomical Society Meeting Abstracts, 109.07

Brown, T. M., Charbonneau, D., Gilliland, R. L., Noyes, R. W., & Burrows, A. 2001, ApJ, 552, 699

Burke, C. J., Bryson, S. T., Mullally, F., Rowe, J. F., Christiansen, J. L., Thompson, S. E., Coughlin, J. L., Haas, M. R., Batalha, N. M., Caldwell, D. A., Jenkins, J. M., Still, M., Barclay, T., Borucki, W. J., Chaplin, W. J., Ciardi, D. R., Clarke, B. D., Cochran, W. D., Demory, B.-O., Esquerdo, G. A., Gautier, III, T. N., Gilliland, R. L., Girouard, F. R., Havel, M., Henze, C. E., Howell, S. B., Huber, D., Latham, D. W., Li, J., Morehead, R. C., et al. 2014, ApJS, 210, 19

Burrows, A., Hubeny, I., Budaj, J., Knutson, H. A., & Charbonneau, D. 2007, ApJ, 668, L171

Burrows, A., Ibgui, L., & Hubeny, I. 2008, ApJ, 682, 1277

Butler, R. P. & Marcy, G. W. 1996, ApJ, 464, L153

Butler, R. P., Marcy, G. W., Vogt, S. S., & Apps, K. 1998, PASP, 110, 1389

Butler, R. P., Marcy, G. W., Williams, E., Hauser, H., & Shirts, P. 1997, ApJ, 474, L115

Butler, R. P., Wright, J. T., Marcy, G. W., Fischer, D. A., Vogt, S. S., Tinney, C. G., Jones, H. R. A., Carter, B. D., Johnson, J. A., McCarthy, C., & Penny, A. J. 2006, ApJ, 646, 505

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

Charbonneau, D., Allen, L. E., Megeath, S. T., Torres, G., Alonso, R., Brown, T. M., Gilliland, R. L., Latham, D. W., Mandushev, G., O'Donovan, F. T., & Sozzetti, A. 2005, ApJ, 626, 523

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., Jha, S., & Noyes, R. W. 1998, ApJ, 507, L153

Charbonneau, D., Knutson, H. A., Barman, T., Allen, L. E., Mayor, M., Megeath, S. T., Queloz, D., & Udry, S. 2008, ApJ, 686, 1341

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

Chatterjee, S., Ford, E. B., Matsumura, S., & Rasio, F. A. 2008, ApJ, 686, 580

Chen, G., van Boekel, R., Madhusudhan, N., Wang, H., Nikolov, N., Seemann, U., & Henning, T. 2014, A&A, 564, A6

Choi, J., McCarthy, C., Marcy, G. W., Howard, A. W., Fischer, D. A., Johnson, J. A., Isaacson, H., & Wright, J. T. 2013, ApJ, 764, 131

Christiansen, J. L., Ballard, S., Charbonneau, D., Madhusudhan, N., Seager, S., Holman, M. J., Wellnitz, D. D., Deming, D., A'Hearn, M. F., & EPOXI Team. 2010, ApJ, 710, 97

Claret, A. 2004, A&A, 428, 1001

Cochran, W. D. & Hatzes, A. P. 1994, Ap&SS, 212, 281

Collier-Cameron, A., Horne, K., James, D., Penny, A., & Semel, M. 2004, in IAU Symposium, Vol. 202, Planetary Systems in the Universe, ed. A. Penny, 75

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

Correia, A. C. M. & Laskar, J. 2011, Tidal Evolution of Exoplanets, ed. S. Seager (Tucson, Arizona, USA: University of Arizona Press), 239–266

Coughlin, J. L. & López-Morales, M. 2012, AJ, 143, 39

Courteau, S., Widrow, L. M., McDonald, M., Guhathakurta, P., Gilbert, K. M., Zhu, Y., Beaton, R. L., & Majewski, S. R. 2011, ApJ, 739, 20

Cowan, N. B. & Agol, E. 2008, ApJ, 678, L129

Cowan, N. B., Machalek, P., Croll, B., Shekhtman, L. M., Burrows, A., Deming, D., Greene, T., & Hora, J. L. 2012, ApJ, 747, 82

Cox, R. T. 1946, AmJPh, 14, 1

Crossfield, I. J. M., Hansen, B. M. S., & Barman, T. 2012a, ApJ, 746, 46

Crossfield, I. J. M., Knutson, H., Fortney, J., Showman, A. P., Cowan, N. B., & Deming, D. 2012b, ApJ, 752, 81

Cubillos, P., Harrington, J., Madhusudhan, N., Stevenson, K. B., Hardy, R. A., Blecic, J., Anderson, D. R., Hardin, M., & Campo, C. J. 2013, ApJ, 768, 42

de Kok, R. J., Brogi, M., Snellen, I. A. G., Birkby, J., Albrecht, S., & de Mooij, E. J. W. 2013, A&A, 554, A82

de Mooij, E. J. W., Brogi, M., de Kok, R. J., Snellen, I. A. G., Kenworthy, M. A., & Karjalainen, R. 2013, A&A, 550, A54

de Mooij, E. J. W., de Kok, R. J., Nefs, S. V., & Snellen, I. A. G. 2011, A&A, 528, A49

de Mooij, E. J. W. & Snellen, I. A. G. 2009, A&A, 493, L35

Deeg, H. J., Moutou, C., Erikson, A., Csizmadia, S., Tingley, B., Barge, P., Bruntt, H., Havel, M., Aigrain, S., Almenara, J. M., Alonso, R., Auvergne, M., Baglin, A., Barbieri, M., Benz, W., Bonomo, A. S., Bordé, P., Bouchy, F., Cabrera, J., Carone, L., Carpano, S., Ciardi, D., Deleuil, M., Dvorak, R., Ferraz-Mello, S., Fridlund, M., Gandolfi, D., Gazzano, J.-C., Gillon, M., Gondoin, P., et al. 2010, Nature, 464, 384

Deich, A. N. & Orlova, O. N. 1977a, Soviet Ast., 21, 182

—. 1977b, AZh, 54, 327

Deming, D., Harrington, J., Laughlin, G., Seager, S., Navarro, S. B., Bowman, W. C., & Horning, K. 2007, ApJ, 667, L199

Deming, D., Harrington, J., Seager, S., & Richardson, L. J. 2006, ApJ, 644, 560

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

Deming, D., Wilkins, A., McCullough, P., Burrows, A., Fortney, J. J., Agol, E., Dobbs-Dixon, I., Madhusudhan, N., Crouzet, N., Desert, J.-M., Gilliland, R. L., Haynes, K., Knutson, H. A., Line, M., Magic, Z., Mandell, A. M., Ranjan, S., Charbonneau, D., Clampin, M., Seager, S., & Showman, A. P. 2013, ApJ, 774, 95

Demory, B.-O., de Wit, J., Lewis, N., Fortney, J., Zsom, A., Seager, S., Knutson, H., Heng, K., Madhusudhan, N., Gillon, M., Barclay, T., Desert, J.-M., Parmentier, V., & Cowan, N. B. 2013, ApJ, 776, L25

Demory, B.-O., Gillon, M., Deming, D., Valencia, D., Seager, S., Benneke, B., Lovis, C., Cubillos, P., Harrington, J., Stevenson, K. B., Mayor, M., Pepe, F., Queloz, D., Ségransan, D., & Udry, S. 2011a, A&A, 533, A114

Demory, B.-O., Seager, S., Madhusudhan, N., Kjeldsen, H., Christensen-Dalsgaard, J., Gillon, M., Rowe, J. F., Welsh, W. F., Adams, E. R., Dupree, A., McCarthy, D., Kulesa, C., Borucki, W. J., & Koch, D. G. 2011b, ApJ, 735, L12

Désert, J.-M., Charbonneau, D., Fortney, J. J., Madhusudhan, N., Knutson, H. A., Fressin, F., Deming, D., Borucki, W. J., Brown, T. M., Caldwell, D., Ford, E. B., Gilliland, R. L., Latham, D. W., Marcy, G. W., & Seager, S. 2011a, ApJS, 197, 11

Désert, J.-M., Lecavelier des Etangs, A., Hébrard, G., Sing, D. K., Ehrenreich, D., Ferlet, R., & Vidal-Madjar, A. 2009, ApJ, 699, 478

Désert, J.-M., Sing, D., Vidal-Madjar, A., Hébrard, G., Ehrenreich, D., Lecavelier Des Etangs, A., Parmentier, V., Ferlet, R., & Henry, G. W. 2011b, A&A, 526, A12

Eastman, J., Gaudi, B. S., & Agol, E. 2013, PASP, 125, 83

Esteves, L. J., De Mooij, E. J. W., & Jayawardhana, R. 2013, ApJ, 772, 51

Evans, T. M., Pont, F., Sing, D. K., Aigrain, S., Barstow, J. K., Désert, J.-M., Gibson, N., Heng, K., Knutson, H. A., & Lecavelier des Etangs, A. 2013, ApJ, 772, L16

Fabrycky, D. & Tremaine, S. 2007, ApJ, 669, 1298

Fazio, G. G. 2004, ApJS, 154, 10

Fischer, D. A., Marcy, G. W., Butler, R. P., Vogt, S. S., & Apps, K. 1999, PASP, 111, 50

Fischer, D. A. & Valenti, J. 2005, ApJ, 622, 1102

Ford, E. B. 2005, AJ, 129, 1706

—. 2006, ApJ, 642, 505

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Fortney, J. J., Demory, B.-O., Désert, J.-M., Rowe, J., Marcy, G. W., Isaacson, H., Buchhave, L. A., Ciardi, D., Gautier, T. N., Batalha, N. M., Caldwell, D. A., Bryson, S. T., Nutzman, P., Jenkins, J. M., Howard, A., Charbonneau, D., Knutson, H. A., Howell, S. B., Everett, M., Fressin, F., Deming, D., Borucki, W. J., Brown, T. M., Ford, E. B., Gilliland, R. L., Latham, D. W., Miller, N., Seager, S., Fischer, D. A., Koch, D., et al. 2011, ApJS, 197, 9

Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, ApJ, 678, 1419

Fortney, J. J. & Nettelmann, N. 2010, Space Sci. Rev., 152, 423

Fortney, J. J., Shabram, M., Showman, A. P., Lian, Y., Freedman, R. S., Marley, M. S., & Lewis, N. K. 2010, ApJ, 709, 1396

Fressin, F., Torres, G., Charbonneau, D., Bryson, S. T., Christiansen, J., Dressing, C. D., Jenkins, J. M., Walkowicz, L. M., & Batalha, N. M. 2013, ApJ, 766, 81

Gatewood, G. 1974, AJ, 79, 52

Gatewood, G. & Eichhorn, H. 1973, AJ, 78, 769

Gaudi, B. S. 2012, ARA&A, 50, 411

Gaudi, B. S., Bennett, D. P., Udalski, A., Gould, A., Christie, G. W., Maoz, D., Dong, S., McCormick, J., Szymański, M. K., Tristram, P. J., Nikolaev, S., Paczyński, B., Kubiak, M., Pietrzyński, G., Soszyński, I., Szewczyk, O., Ulaczyk, K., Wyrzykowski, Ł., OGLE Collaboration, DePoy, D. L., Han, C., Kaspi, S., Lee, C.-U., Mallia, F., Natusch, T., Pogge, R. W., Park, B.-G., μ-Fun Collabortion, Abe, F., Bond, I. A., Botzler, C. S., Fukui, A., et al. 2008, Science, 319, 927

Gelman, A., Roberts, G., & Gilks, W. 1996, in Bayesian Statistics, ed. J. M. Bernado et al., Vol. 5 (OUP), 599

Gelman, A. & Rubin, D. B. 1992, Stat. Sci., 7, 457

Gibson, N. P., Aigrain, S., Barstow, J. K., Evans, T. M., Fletcher, L. N., & Irwin, P. G. J. 2013a, MNRAS, 428, 3680

—. 2013b, MNRAS, 436, 2974

Gibson, N. P., Aigrain, S., Pollacco, D. L., Barros, S. C. C., Hebb, L., Hrudková, M., Simpson, E. K., Skillen, I., & West, R. 2010a, MNRAS, 404, L114

Gibson, N. P., Aigrain, S., Pont, F., Sing, D. K., Désert, J.-M., Evans, T. M., Henry, G., Husnoo, N., & Knutson, H. 2012a, MNRAS, 422, 753

Gibson, N. P., Aigrain, S., Roberts, S., Evans, T. M., Osborne, M., & Pont, F. 2012b, MNRAS, 419, 2683

Gibson, N. P., Pollacco, D. L., Barros, S., Benn, C., Christian, D., Hrudková, M., Joshi, Y. C., Keenan, F. P., Simpson, E. K., Skillen, I., Steele, I. A., & Todd, I. 2010b, MNRAS, 401, 1917

Gibson, N. P., Pont, F., & Aigrain, S. 2011, MNRAS, 411, 2199

Gillon, M., Demory, B.-O., Barman, T., Bonfils, X., Mazeh, T., Pont, F., Udry, S., Mayor, M., & Queloz, D. 2007, A&A, 471, L51

Gould, A., Dorsher, S., Gaudi, B. S., & Udalski, A. 2006, Acta Astron., 56, 1

Grillmair, C. J., Burrows, A., Charbonneau, D., Armus, L., Stauffer, J., Meadows, V., van Cleve, J., von Braun, K., & Levine, D. 2008, Nature, 456, 767

Grillmair, C. J., Charbonneau, D., Burrows, A., Armus, L., Stauffer, J., Meadows, V., Van Cleve, J., & Levine, D. 2007, ApJ, 658, L115

Harrington, R. S., Kallarakal, V. V., & Dahn, C. C. 1983, AJ, 88, 1038

Hartman, J. D., Bakos, G. Á., Torres, G., Latham, D. W., Kovács, G., Béky, B., Quinn, S. N., Mazeh, T., Shporer, A., Marcy, G. W., Howard, A. W., Fischer, D. A., Johnson, J. A., Esquerdo, G. A., Noyes, R. W., Sasselov, D. D., Stefanik, R. P., Fernandez, J. M., Szklenár, T., Lázár, J., Papp, I., & Sári, P. 2011, ApJ, 742, 59

Hastings, W. K. 1970, Biometrika, 57, 97

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

Hébrard, G., Evans, T. M., Alonso, R., Fridlund, M., Ofir, A., Aigrain, S., Guillot, T., Almenara, J. M., Auvergne, M., Baglin, A., Barge, P., Bonomo, A. S., Bordé, P., Bouchy, F., Cabrera, J., Carone, L., Carpano, S., Cavarroc, C., Csizmadia, S., Deeg, H. J., Deleuil, M., Díaz, R. F., Dvorak, R., Erikson, A., Ferraz-Mello, S., Gandolfi, D., Gibson, N., Gillon, M., Guenther, E., Hatzes, A., et al. 2011, A&A, 533, A130

Heintz, W. D. 1988, JRASC, 82, 140

Heng, K. & Demory, B.-O. 2013, ApJ, 777, 100

Henry, G. W. 1999, PASP, 111, 845

Henry, G. W., Baliunas, S. L., Donahue, R. A., Soon, W. H., & Saar, S. H. 1997, ApJ, 474, 503

Henry, G. W., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2000, ApJ, 529, L41

Holman, M. J. & Murray, N. W. 2005, Science, 307, 1288

Horne, K. 2003, in Astronomical Society of the Pacific Conference Series, Vol. 294, Scientific Frontiers in Research on Extrasolar Planets, ed. D. Deming & S. Seager, 361–370

Howard, A. W., Marcy, G. W., Bryson, S. T., Jenkins, J. M., Rowe, J. F., Batalha, N. M., Borucki, W. J., Koch, D. G., Dunham, E. W., Gautier, III, T. N., Van Cleve, J., Cochran, W. D., Latham, D. W., Lissauer, J. J., Torres, G., Brown, T. M., Gilliland, R. L., Buchhave, L. A., Caldwell, D. A., Christensen-Dalsgaard, J., Ciardi, D., Fressin, F., Haas, M. R., Howell, S. B., Kjeldsen, H., Seager, S., Rogers, L., Sasselov, D. D., Steffen, J. H., Basri, G. S., et al. 2012, ApJS, 201, 15

Howell, S. B., Sobeck, C., Haas, M., Still, M., Barclay, T., Mullally, F., Troeltzsch, J., Aigrain, S., Bryson, S. T., Caldwell, D., Chaplin, W. J., Cochran, W. D., Huber, D., Marcy, G. W., Miglio, A., Najita, J. R., Smith, M., Twicken, J. D., & Fortney, J. J. 2014, PASP, 126, 398

Huitson, C. M., Sing, D. K., Pont, F., Fortney, J. J., Burrows, A. S., Wilson, P. A., Ballester, G. E., Nikolov, N., Gibson, N. P., Deming, D., Aigrain, S., Evans, T. M., Henry, G. W., Lecavelier des Etangs, A., Showman, A. P., Vidal-Madjar, A., & Zahnle, K. 2013, MNRAS, 434, 3252

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

Irwin, P. G. J., Teanby, N. A., de Kok, R., Fletcher, L. N., Howett, C. J. A., Tsang, C. C. C., Wilson, C. F., Calcutt, S. B., Nixon, C. A., & Parrish, P. D. 2008, J. Quant. Spec. Radiat. Transf., 109, 1136

Jackson, B. K., Lewis, N. K., Barnes, J. W., Drake Deming, L., Showman, A. P., & Fortney, J. J. 2012, ApJ, 751, 112

Jacob, W. S. 1855, MNRAS, 15, 228

Jensen, A. G., Redfield, S., Endl, M., Cochran, W. D., Koesterke, L., & Barman, T. S. 2011, ApJ, 743, 203

Karkoschka, E. 1994, Icarus, 111, 174

Kipping, D. & Bakos, G. 2011, ApJ, 730, 50

Kipping, D. M. & Spiegel, D. S. 2011, MNRAS, 417, L88

Knutson, H. A., Charbonneau, D., Allen, L. E., Burrows, A., & Megeath, S. T. 2008, ApJ, 673, 526

Knutson, H. A., Charbonneau, D., Allen, L. E., Fortney, J. J., Agol, E., Cowan, N. B., Showman, A. P., Cooper, C. S., & Megeath, S. T. 2007a, Nature, 447, 183

Knutson, H. A., Charbonneau, D., Cowan, N. B., Fortney, J. J., Showman, A. P., Agol, E., & Henry, G. W. 2009a, ApJ, 703, 769

Knutson, H. A., Charbonneau, D., Cowan, N. B., Fortney, J. J., Showman, A. P., Agol, E., Henry, G. W., Everett, M. E., & Allen, L. E. 2009b, ApJ, 690, 822

Knutson, H. A., Charbonneau, D., Noyes, R. W., Brown, T. M., & Gilliland, R. L. 2007b, ApJ, 655, 564

Knutson, H. A., Howard, A. W., & Isaacson, H. 2010, ApJ, 720, 1569

Knutson, H. A., Lewis, N., Fortney, J. J., Burrows, A., Showman, A. P., Cowan, N. B., Agol, E., Aigrain, S., Charbonneau, D., Deming, D., Désert, J.-M., Henry, G. W., Langton, J., & Laughlin, G. 2012, ApJ, 754, 22

Konacki, M., Torres, G., Jha, S., & Sasselov, D. D. 2003, Nature, 421, 507

Kurucz, R. 1993, ATLAS9 Stellar Atmosphere Programs and 2 km/s grid. Kurucz CD-ROM No. 13. Cambridge, Mass.: Smithsonian Astrophysical Observatory, 1993., 13

Kurucz, R. L. 1979, ApJS, 40, 1

Latham, D. W., Stefanik, R. P., Mazeh, T., Mayor, M., & Burki, G. 1989, Nature, 339, 38

Lecavelier Des Etangs, A., Pont, F., Vidal-Madjar, A., & Sing, D. 2008, A&A, 481, L83

Lee, J.-M., Fletcher, L. N., & Irwin, P. G. J. 2012, MNRAS, 420, 170

Leigh, C., Collier Cameron, A., Horne, K., Penny, A., & James, D. 2003, MNRAS, 344, 1271

Levenberg, K. 1944, Q Appl Math, 2, 164

Lewis, N. K., Knutson, H. A., Showman, A. P., Cowan, N. B., Laughlin, G., Burrows, A., Deming, D., Crepp, J. R., Mighell, K. J., Agol, E., Bakos, G. Á., Charbonneau, D., Désert, J.-M., Fischer, D. A., Fortney, J. J., Hartman, J. D., Hinkley, S., Howard, A. W., Johnson, J. A., Kao, M., Langton, J., & Marcy, G. W. 2013, ApJ, 766, 95

Lin, D. N. C., Bodenheimer, P., & Richardson, D. C. 1996, Nature, 380, 606

Line, M. R. & Yung, Y. L. 2013, ApJ, 779, 3

Lissauer, J. J. 1993, ARA&A, 31, 129

Lovis, C. & Fischer, D. 2011, Exoplanets, ed. S. Seager (Tucson, Arizona, USA: University of Arizona Press), 27–53

Lucas, P. W., Hough, J. H., Bailey, J. A., Tamura, M., Hirst, E., & Harrison, D. 2009, MNRAS, 393, 229

Machalek, P., Greene, T., McCullough, P. R., Burrows, A., Burke, C. J., Hora, J. L., Johns-Krull, C. M., & Deming, D. L. 2010, ApJ, 711, 111

Madhusudhan, N. & Burrows, A. 2012, ApJ, 747, 25

Madhusudhan, N., Harrington, J., Stevenson, K. B., Nymeyer, S., Campo, C. J., Wheatley, P. J., Deming, D., Blecic, J., Hardy, R. A., Lust, N. B., Anderson, D. R., Collier-Cameron, A., Britt, C. B. T., Bowman, W. C., Hebb, L., Hellier, C., Maxted, P. F. L., Pollacco, D., & West, R. G. 2011, Nature, 469, 64

Madhusudhan, N. & Seager, S. 2009, ApJ, 707, 24

—. 2010, ApJ, 725, 261

Mädler, J. H. 1842, Astronomische Nachrichten, 19, 201

Mandel, K. & Agol, E. 2002, ApJ, 580, L171

Mandell, A. M., Drake Deming, L., Blake, G. A., Knutson, H. A., Mumma, M. J., Villanueva, G. L., & Salyk, C. 2011, ApJ, 728, 18

Marcy, G. W. & Butler, R. P. 1992, PASP, 104, 270

—. 1996, ApJ, 464, L147

Marcy, G. W., Isaacson, H., Howard, A. W., Rowe, J. F., Jenkins, J. M., Bryson, S. T., Latham, D. W., Howell, S. B., Gautier, III, T. N., Batalha, N. M., Rogers, L., Ciardi, D., Fischer, D. A., Gilliland, R. L., Kjeldsen, H., Christensen-Dalsgaard, J., Huber, D., Chaplin, W. J., Basu, S., Buchhave, L. A., Quinn, S. N., Borucki, W. J., Koch, D. G., Hunter, R., Caldwell, D. A., Van Cleve, J., Kolbl, R., Weiss, L. M., Petigura, E., Seager, S., et al. 2014, ApJS, 210, 20

Marley, M. S., Gelino, C., Stephens, D., Lunine, J. I., & Freedman, R. 1999, ApJ, 513, 879

Marois, C., Macintosh, B., Barman, T., Zuckerman, B., Song, I., Patience, J., Lafrenière, D., & Doyon, R. 2008, Science, 322, 1348

Maxted, P. F. L., Anderson, D. R., Doyle, A. P., Gillon, M., Harrington, J., Iro, N., Jehin, E., Lafrenière, D., Smalley, B., & Southworth, J. 2013, MNRAS, 428, 2645

Mayor, M. & Queloz, D. 1995, Nature, 378, 355

McCullough, P. R., Stys, J. E., Valenti, J. A., Fleming, S. W., Janes, K. A., & Heasley, J. N. 2005, PASP, 117, 783

McQuillan, A., Aigrain, S., & Roberts, S. 2012, A&A, 539, A137

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., & Teller, A. H. 1953, J. Chem. Phys., 21, 1087

Morales-Calderón, M., Stauffer, J. R., Kirkpatrick, J. D., Carey, S., Gelino, C. R., Barrado y Navascués, D., Rebull, L., Lowrance, P., Marley, M. S., Charbonneau, D., Patten, B. M., Megeath, S. T., & Buzasi, D. 2006, ApJ, 653, 1454

Mordasini, C., Alibert, Y., Benz, W., & Naef, D. 2009, A&A, 501, 1161

Morris, B. M., Mandell, A. M., & Deming, D. 2013, ApJ, 764, L22

Moutou, C., Deleuil, M., Guillot, T., Baglin, A., Bordé, P., Bouchy, F., Cabrera, J., Csizmadia, S., & Deeg, H. J. 2013, Icarus, 226, 1625

Murgas, F., Pallé, E., Zapatero Osorio, M. R., Nortmann, L., Hoyer, S., & Cabrera-Lavers, A. 2014, A&A, 563, A41

Murray, C. D. & Correia, A. C. M. 2010, Keplerian Orbits and Dynamics of Exoplanets, ed. S. Seager (Tucson, Arizona, USA: University of Arizona Press), 15–23

Muterspaugh, M. W., Lane, B. F., Kulkarni, S. R., Konacki, M., Burke, B. F., Colavita, M. M., Shao, M., Hartkopf, W. I., Boss, A. P., & Williamson, M. 2010, AJ, 140, 1657

Nelder, J. A. & Mead, R. 1965, Computer Journal, 7, 308

Nikolov, N., Sing, D. K., Pont, F., Burrows, A. S., Fortney, J. J., Ballester, G. E., Evans, T. M., Huitson, C. M., Wakeford, H. R., Wilson, P. A., Aigrain, S., Deming, D., Gibson, N. P., Henry, G. W., Knutson, H., Lecavelier des Etangs, A., Showman, A. P., Vidal-Madjar, A., & Zahnle, K. 2014, MNRAS, 437, 46

Nutzman, P. & Charbonneau, D. 2008, PASP, 120, 317

Oppenheimer, B. R. & Hinkley, S. 2009, ARA&A, 47, 253

Papaloizou, J. C. B. & Terquem, C. 2006, Reports on Progress in Physics, 69, 119

Parmentier, V., Showman, A. P., & Lian, Y. 2013, A&A, 558, A91

Parviainen, H., Deeg, H. J., & Belmonte, J. A. 2013, A&A, 550, A67

Perna, R., Heng, K., & Pont, F. 2012, ApJ, 751, 59

Plavchan, P. & Bilinski, C. 2013, ApJ, 769, 86

Pollacco, D. L., Skillen, I., Collier Cameron, A., Christian, D. J., Hellier, C., Irwin, J., Lister, T. A., Street, R. A., West, R. G., Anderson, D. R., Clarkson, W. I., Deeg, H., Enoch, B., Evans, A., Fitzsimmons, A., Haswell, C. A., Hodgkin, S., Horne, K., Kane, S. R., Keenan, F. P., Maxted, P. F. L., Norton, A. J., Osborne, J., Parley, N. R., Ryans, R. S. I., Smalley, B., Wheatley, P. J., & Wilson, D. M. 2006, PASP, 118, 1407

Pont, F., Husnoo, N., Mazeh, T., & Fabrycky, D. 2011, MNRAS, 414, 1278

Pont, F., Knutson, H., Gilliland, R. L., Moutou, C., & Charbonneau, D. 2008a, MNRAS, 385, 109

Pont, F., Sing, D. K., Gibson, N. P., Aigrain, S., Henry, G., & Husnoo, N. 2013, MNRAS, 432, 2917

Pont, F., Tamuz, O., Udalski, A., Mazeh, T., Bouchy, F., Melo, C., Naef, D., Santos, N. C., Moutou, C., Diaz, R. F., Gieren, W., Gillon, M., Hoyer, S., Kubiak, M., Mayor, M., Minniti, D., Pietrzynski, G., Queloz, D., Ramirez, S., Ruiz, M. T., Shporer, A., Soszyński, I., Szewczyk, O., Szymański, M. K., Udry, S., Ulaczyk, K., Wyrzykowski, Ł., & Zoccali, M. 2008b, A&A, 487, 749

Queloz, D., Mayor, M., Weber, L., Blécha, A., Burnet, M., Confino, B., Naef, D., Pepe, F., Santos, N., & Udry, S. 2000, A&A, 354, 99

Quiñonero-Candela, J. & Rasmussen, C. E. 2005, JMLR, 6, 1939

Quintana, E. V., Rowe, J. F., Barclay, T., Howell, S. B., Ciardi, D. R., Demory, B.-O., Caldwell, D. A., Borucki, W. J., Christiansen, J. L., Jenkins, J. M., Klaus, T. C., Fulton, B. J., Morris, R. L., Sanderfer, D. T., Shporer, A., Smith, J. C., Still, M., & Thompson, S. E. 2013, ApJ, 767, 137

Ranjan, S., Charbonneau, D., Désert, J.-M., Madhusudhan, N., Deming, D., Wilkins, A., & Mandell, A. M. 2014, ApJ, 785, 148

Rasio, F. A., Tout, C. A., Lubow, S. H., & Livio, M. 1996, ApJ, 470, 1187

Rasmussen, C. E. & Williams, C. 2006, Gaussian Processes for Machine Learning (Cambridge, Massachusetts, USA: MIT Press)

Rauscher, E. & Menou, K. 2012, ApJ, 745, 78

Rauscher, E., Menou, K., Seager, S., Deming, D., Cho, J. Y.-K., & Hansen, B. M. S. 2007, ApJ, 664, 1199

Reach, W. T., Megeath, S. T., Cohen, M., Hora, J., Carey, S., Surace, J., Willner, S. P., Barmby, P., Wilson, G., Glaccum, W., Lowrance, P., Marengo, M., & Fazio, G. G. 2005, PASP, 117, 978

Redfield, S., Endl, M., Cochran, W. D., & Koesterke, L. 2008, ApJ, 673, L87

Reuyl, D. & Holmberg, E. 1943, ApJ, 97, 41

Richardson, L. J., Deming, D., Horning, K., Seager, S., & Harrington, J. 2007, Nature, 445, 892

Richardson, L. J., Harrington, J., Seager, S., & Deming, D. 2006, ApJ, 649, 1043

Roberts, G. O., Gelman, A., & Gilks, W. R. 1997, The Annals of Applied Probability, 7, 110

Rodler, F., Kürster, M., & Barnes, J. R. 2013, MNRAS, 432, 1980

Rowe, J. F., Bryson, S. T., Marcy, G. W., Lissauer, J. J., Jontof-Hutter, D., Mullally, F., Gilliland, R. L., Issacson, H., Ford, E., Howell, S. B., Borucki, W. J., Haas, M., Huber, D., Steffen, J. H., Thompson, S. E., Quintana, E., Barclay, T., Still, M., Fortney, J., Gautier, III, T. N., Hunter, R., Caldwell, D. A., Ciardi, D. R., Devore, E., Cochran, W., Jenkins, J., Agol, E., Carter, J. A., & Geary, J. 2014, ApJ, 784, 45

Rowe, J. F., Matthews, J. M., Seager, S., Miller-Ricci, E., Sasselov, D., Kuschnig, R., Guenther, D. B., Moffat, A. F. J., Rucinski, S. M., Walker, G. A. H., & Weiss, W. W. 2008, ApJ, 689, 1345

Russell, H. N. 1916, ApJ, 43, 173

Saar, S. H., Butler, R. P., & Marcy, G. W. 1998, ApJ, 498, L153

Santerne, A., Bonomo, A. S., Hébrard, G., Deleuil, M., Moutou, C., Almenara, J.-M., Bouchy, F., & Díaz, R. F. 2011, A&A, 536, A70

Schlaufman, K. C. 2010, ApJ, 719, 602

Schwarz, G. E. 1978, Ann. Stat., 6, 461

Seager, S. 2010, Exoplanet Atmospheres: Physical Processes (Princeton: Princeton Univ. Press)

Seager, S. & Deming, D. 2010, ARA&A, 48, 631

Seager, S. & Sasselov, D. D. 1998, ApJ, 502, L157

—. 2000, ApJ, 537, 916

Seager, S., Whitney, B. A., & Sasselov, D. D. 2000, ApJ, 540, 504

See, T. J. J. 1896, AJ, 16, 17

Seeliger, M., Dimitrov, D., Kjurkchieva, D., Mallonn, M., Fernandez, M., Kitze, M., Casanova, V., Maciejewski, G., Ohlert, J. M., Schmidt, J. G., Pannicke, A., Puchalski, D., Göğüş, E., Güver, T., Bilir, S., Ak, T., Hohle, M. M., Schmidt, T. O. B., Errmann, R., Jensen, E., Cohen, D., Marschall, L., Saral, G., Bernt, I., Derman, E., Gałan, C., & Neuhäuser, R. 2014, MNRAS, 441, 304

Showman, A. P., Cho, J. Y.-K., & Menou, K. 2011, Atmospheric Circulation of Exoplanets, ed. S. Seager (Tucson, Arizona, USA: University of Arizona Press), 471–516

Showman, A. P., Fortney, J. J., Lian, Y., Marley, M. S., Freedman, R. S., Knutson, H. A., & Charbonneau, D. 2009, ApJ, 699, 564

Showman, A. P. & Guillot, T. 2002, A&A, 385, 166

Sing, D. K. 2010, A&A, 510, A21

Sing, D. K., Désert, J.-M., Fortney, J. J., Lecavelier Des Etangs, A., Ballester, G. E., Cepa, J., Ehrenreich, D., López-Morales, M., Pont, F., Shabram, M., & Vidal-Madjar, A. 2011a, A&A, 527, A73

Sing, D. K., Désert, J.-M., Lecavelier Des Etangs, A., Ballester, G. E., Vidal-Madjar, A., Parmentier, V., Hebrard, G., & Henry, G. W. 2009, A&A, 505, 891

Sing, D. K., Huitson, C. M., Lopez-Morales, M., Pont, F., Désert, J.-M., Ehrenreich, D., Wilson, P. A., Ballester, G. E., Fortney, J. J., Lecavelier des Etangs, A., & Vidal-Madjar, A. 2012, MNRAS, 426, 1663

Sing, D. K., Lecavelier des Etangs, A., Fortney, J. J., Burrows, A. S., Pont, F., Wakeford, H. R., Ballester, G. E., Nikolov, N., Henry, G. W., Aigrain, S., Deming, D., Evans, T. M., Gibson, N. P., Huitson, C. M., Knutson, H., Showman, A. P., Vidal-Madjar, A., Wilson, P. A., Williamson, M. H., & Zahnle, K. 2013, MNRAS, 436, 2956

Sing, D. K., Pont, F., Aigrain, S., Charbonneau, D., Désert, J.-M., Gibson, N., Gilliland, R., Hayek, W., Henry, G., Knutson, H., Lecavelier Des Etangs, A., Mazeh, T., & Shporer, A. 2011b, MNRAS, 416, 1443

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

Skilling, J. 2004, AIP Conf. Proc., 735, 395

Skrutskie, M. F., Forrest, W. J., & Shure, M. A. 1987, ApJ, 312, L55

Snellen, I. A. G., Albrecht, S., de Mooij, E. J. W., & Le Poole, R. S. 2008, A&A, 487, 357

Snellen, I. A. G. & Covino, E. 2007, MNRAS, 375, 307

Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010a, Nature, 465, 1049

Snellen, I. A. G., de Mooij, E. J. W., & Albrecht, S. 2009, Nature, 459, 543

Snellen, I. A. G., de Mooij, E. J. W., & Burrows, A. 2010b, A&A, 513, A76

Sozzetti, A., Giacobbe, P., Lattanzi, M. G., Micela, G., Morbidelli, R., & Tinetti, G. 2014, MNRAS, 437, 497

Stevenson, K. B., Bean, J. L., Seifahrt, A., Désert, J.-M., Madhusudhan, N., Bergmann, M., Kreidberg, L., & Homeier, D. 2014, AJ, 147, 161

Stevenson, K. B., Harrington, J., Fortney, J. J., Loredo, T. J., Hardy, R. A., Nymeyer, S., Bowman, W. C., Cubillos, P., Bowman, M. O., & Hardin, M. 2012a, ApJ, 754, 136

Stevenson, K. B., Harrington, J., Lust, N. B., Lewis, N. K., Montagnier, G., Moses, J. I., Visscher, C., Blecic, J., Hardy, R. A., Cubillos, P., & Campo, C. J. 2012b, ApJ, 755, 9

Strand, K. A. 1937, Annalen van de Sterrewacht te Leiden, 18, B1

—. 1943, PASP, 55, 29

—. 1952, AJ, 57, 97

Struve, O. 1952, The Observatory, 72, 199

Sudarsky, D., Burrows, A., Hubeny, I., & Li, A. 2005, ApJ, 627, 520

Sudarsky, D., Burrows, A., & Pinto, P. 2000, ApJ, 538, 885

Swain, M. R., Bouwman, J., Akeson, R. L., Lawler, S., & Beichman, C. A. 2008a, ApJ, 674, 482

Swain, M. R., Deroo, P., Griffith, C. A., Tinetti, G., Thatte, A., Vasisht, G., Chen, P., Bouwman, J., Crossfield, I. J., Angerhausen, D., Afonso, C., & Henning, T. 2010, Nature, 463, 637

Swain, M. R., Tinetti, G., Vasisht, G., Deroo, P., Griffith, C., Bouwman, J., Chen, P., Yung, Y., Burrows, A., Brown, L. R., Matthews, J., Rowe, J. F., Kuschnig, R., & Angerhausen, D. 2009a, ApJ, 704, 1616

Swain, M. R., Vasisht, G., & Tinetti, G. 2008b, Nature, 452, 329

Swain, M. R., Vasisht, G., Tinetti, G., Bouwman, J., Chen, P., Yung, Y., Deming, D., & Deroo, P. 2009b, ApJ, 690, L114

Tegmark, M., Strauss, M. A., Blanton, M. R., Abazajian, K., Dodelson, S., Sandvik, H., Wang, X., Weinberg, D. H., Zehavi, I., Bahcall, N. A., Hoyle, F., Schlegel, D., Scoccimarro, R., Vogeley, M. S., Berlind, A., Budavari, T., Connolly, A., Eisenstein, D. J., Finkbeiner, D., Frieman, J. A., Gunn, J. E., Hui, L., Jain, B., Johnston, D., Kent, S., Lin, H., Nakajima, R., Nichol, R. C., Ostriker, J. P., Pope, A., et al. 2004, Phys. Rev. D, 69, 103501

Tinetti, G., Deroo, P., Swain, M. R., Griffith, C. A., Vasisht, G., Brown, L. R., Burke, C., & McCullough, P. 2010, ApJ, 712, L139

Tinetti, G., Vidal-Madjar, A., Liang, M.-C., Beaulieu, J.-P., Yung, Y., Carey, S., Barber, R. J., Tennyson, J., Ribas, I., Allard, N., Ballester, G. E., Sing, D. K., & Selsis, F. 2007, Nature, 448, 169

Titov, D. V., Bullock, M. A., Crisp, D., Renno, N. O., Taylor, F. W., & Zasova, L. V. 2007, Washington DC American Geophysical Union Geophysical Monograph Series, 176, 121

Todorov, K., Deming, D., Harrington, J., Stevenson, K. B., Bowman, W. C., Nymeyer, S., Fortney, J. J., & Bakos, G. A. 2010, ApJ, 708, 498

Tody, D. 1986, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 627, Instrumentation in astronomy VI, ed. D. L. Crawford, 733

Torres, G., Winn, J. N., & Holman, M. J. 2008, ApJ, 677, 1324

Udalski, A., Paczynski, B., Zebrun, K., Szymanski, M., Kubiak, M., Soszynski, I., Szewczyk, O., Wyrzykowski, L., & Pietrzynski, G. 2002a, Acta Astron., 52, 1

Udalski, A., Szewczyk, O., Zebrun, K., Pietrzynski, G., Szymanski, M., Kubiak, M., Soszynski, I., & Wyrzykowski, L. 2002b, Acta Astron., 52, 317

Udry, S. & Santos, N. C. 2007, ARA&A, 45, 397

van de Kamp, P. 1944, Proc. Am. Phil. Soc., 88, 372

—. 1969, AJ, 74, 757

Waldmann, I. P., Tinetti, G., Drossart, P., Swain, M. R., Deroo, P., & Griffith, C. A. 2012, ApJ, 744, 35

Walker, G. A. H., Bohlender, D. A., Walker, A. R., Irwin, A. W., Yang, S. L. S., & Larson, A. 1992, ApJ, 396, L91

Walker, M. G., Mateo, M., Olszewski, E. W., Peñarrubia, J., Wyn Evans, N., & Gilmore, G. 2009, ApJ, 704, 1274

Weiss, L. M., Marcy, G. W., Rowe, J. F., Howard, A. W., Isaacson, H., Fortney, J. J., Miller, N., Demory, B.-O., Fischer, D. A., Adams, E. R., Dupree, A. K., Howell, S. B., Kolbl, R., Johnson, J. A., Horch, E. P., Everett, M. E., Fabrycky, D. C., & Seager, S. 2013, ApJ, 768, 14

Welsh, W. F., Orosz, J. A., Seager, S., Fortney, J. J., Jenkins, J., Rowe, J. F., Koch, D., & Borucki, W. J. 2010, ApJ, 713, L145

Wiktorowicz, S. J. 2009, ApJ, 696, 1116

Wilkins, A. N., Deming, D., Madhusudhan, N., Burrows, A., Knutson, H., McCullough, P., & Ranjan, S. 2014, ApJ, 783, 113

Winn, J. N., Fabrycky, D., Albrecht, S., & Johnson, J. A. 2010, ApJ, 718, L145

Winn, J. N., Holman, M. J., Shporer, A., Fernández, J., Mazeh, T., Latham, D. W., Charbonneau, D., & Everett, M. E. 2008, AJ, 136, 267

Wolszczan, A. 1994, Science, 264, 538

Wolszczan, A. & Frail, D. A. 1992, Nature, 355, 145

Wolszczan, A. & Kuchner, M. 2011, Exoplanets, ed. S. Seager (Tucson, Arizona, USA: University of Arizona Press), 175–190

Woodgate, B. E., Kimble, R. A., Bowers, C. W., Kraemer, S., Kaiser, M. E., Danks, A. C., Grady, J. F., Loiacono, J. J., Brumfield, M., Feinberg, L., Gull, T. R., Heap, S. R., Maran, S. P., Lindler, D., Hood, D., Meyer, W., Vanhouten, C., Argabright, V., Franka, S., Bybee, R., Dorn, D., Bottema, M., Woodruff, R., Michika, D., Sullivan, J., Hetlinger, J., Ludtke, C., Stocker, R., Delamere, A., Rose, D., et al. 1998, PASP, 110, 1183

Wright, J. T., Marcy, G. W., Howard, A. W., Johnson, J. A., Morton, T. D., & Fischer, D. A. 2012, ApJ, 753, 160

Wright, J. T., Upadhyay, S., Marcy, G. W., Fischer, D. A., Ford, E. B., & Johnson, J. A. 2009, ApJ, 693, 1084

Zellem, R. T., Lewis, N. K., Knutson, H. A., Griffith, C. A., Showman, A. P., Fortney, J. J., Cowan, N. B., Agol, E., Burrows, A., Charbonneau, D., Deming, D., Laughlin, G., & Langton, J. 2014, ApJ, 790, 53

Zhao, M., Milburn, J., Barman, T., Hinkley, S., Swain, M. R., Wright, J., & Monnier, J. D. 2012, ApJ, 748, L8