ABSTRACT

Title of Dissertation:	REVEALING UNIQUE EXOPLANET ATMOSPHERES WITH MULTI-INSTRUMENT SPACE TELESCOPE TRANSIT AND ECLIPSE SPECTROSCOPY
	Kyle Sheppard Doctor of Philosophy, 2021
Dissertation Directed by:	Professor Drake Deming Department of Astronomy

Atmospheres act as windows into their host planets, containing measurable information on their planets' chemistry, climate, and atmospheric physics. The bulk properties of planets outside of the Solar System (exoplanets) prove to be much more varied than the Solar System, allowing the ability to test atmospheric models over a range of temperatures, radii, and host star properties. Modeling and observing exoplanet atmospheres provides a better understanding of both atmospheric processes and planetary diversity, and it places the Solar System in a greater context to understand how unique it is, if at all. I take a broad approach, analyzing both transit and emission spectroscopy of 5 exoplanets populating the edges of parameter space, ranging from cool, Earth-sized planets (T$\sim$500K, R=0.8$R_\oplus$) up to massive, ultra-hot Jupiters (T$\sim$2500K, M=10M_{Jup}). I use my publicly available, open source Python 3 analysis pipeline `DEFLATE` to process telescope data and produce verifiable spectra. I then retrieve atmospheric properties using a forward model + Bayesian sampler

retrieval tool, exploring how both inter- and intra- modeling assumptions impact results. I retrieve unexpected atmospheres, including: evidence of stellar activity mimicking water vapor features in two terrestrial planets in the multi-planet L9859 system; evidence of a clear atmosphere and a superstellar atmospheric metallicity and water abundance (5σ detection) in the hot Jupiter HAT-P-41b ($R=1.65R_{\mathrm{Jup}}$, $T_{\mathrm{eq}}=1950$ K); a potentially non-TiO driven thermal inversion and a photometric CO detection (6σ) in the ultrahot Jupiter WASP-18b; and a water absorption feature (2.8σ) and non-inverted T-P profile in the water-dissociation-vulnerable hot Jupiter WASP-19b ($R=1.4R_{\mathrm{Jup}}$, $T_{\mathrm{eq}}=2120$ K). Overall, these results expand already extensive diversity of exoplanet atmospheres.

Revealing Unique Exoplanet Atmospheres with Multi-instrument Space Telescope Transit and Eclipse Spectroscopy

by

Kyle Sheppard

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:
Professor Drake Deming, Chair/Advisor
Professor Kayo Ide, Dean's Representative
Professor Eliza Kempton
Dr. Avi Mandell
Professor Derek Richardson

Preface

Disclaimer: I originally published the majority of Chapter 3 in the Astronomical Journal (Sheppard et al., 2021). I published the majority of the WASP-18b-related analyses in Chapter 4 in Astrophysical Journal Letters (Sheppard et al., 2017). Even though I was first author and in charge of writing and organizing each paper, I also worked with collaborators, who contributed a significant amount. There are sections of the dissertation where a collaborator performed a retrieval or data analysis and my contribution was limited to discussing, interpreting, and writing. These sections are necessary to include for a complete understanding of each chapter. To clearly designate sections where a collaborator contributed significantly to the analysis, I use “we”, “a collaborator”, or “they”.

Acknowledgments

No one achieves anything alone. I have been surrounded by supportive people which was not only helpful, but also necessary for me to attain a PhD. First and foremost, I want to thank my parents, Amy Feinberg Sheppard and Paul Sheppard, for always encouraging, supporting, and believing in me — both in my education and in life in general. The same goes for my older siblings Mike and Brianna, who helped guide me through the early years. I owe a lot to my girlfriend Megan, who was beside me the entire way. She made the lows of my time in graduate school tolerable, and the highs immeasurably better.

I'd like to thank my extended family and Megan's family who have been on my side every step of the way. I'd also like to thank my friends, who entertained and distracted me enough to keep me sane during the more stressful moments. I want to acknowledge my graduate classmates, who were always welcoming and made everything from classes to quals easier.

I owe a lot to my advisors, Drake Deming and Avi Mandell, both of who consistently advocated for me. They helped guide my research, coauthored my papers, and connected me to other researchers outside of UMD. In particular, Avi spent countless hours reading over my drafts, video chatting with me, helping me develop code, and mentoring me in general, and for that I'm grateful.

Thanks to my thesis committee for helping me refine my research and for their instrumental feedback on my dissertation. I want to acknowledge my collaborators — especially Nikolay, Madhu, Luis, and Sid — who were great to work with (both as scientists and people). Special thanks to Cole Miller for long discussions on Bayesian statistics that were indispensable to my research.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Astronomical Symbols and Units:

AU	Astronomical Unit (1.496×10^{11} m)
$M_\oplus$	Earth mass (5.972×10^{24} kg)
M_{Jupiter}	Jupiter mass (1.899×10^{27} kg)
$M_\odot$	Solar mass (1.988×10^{30} kg)
$R_\oplus$	Earth radius (6.378×10^{6} m)
R_{Jupiter}	Jupiter radius (7.149×10^{7} m)
$R_\odot$	Solar radius (6.957×10^{8} m)

Chemical Symbols and Formulae:

AlO	Aluminum Oxide
Ca	Calcium
C/H	Carbon-to-Hydrogen ratio
CH_4	Methane
CO	Carbon Monoxide
CO_2	Carbon Dioxide
C/O	Carbon-to-Oxygen ratio
Fe	Iron
FeH	Iron Hydride
H^-	Negative Hydrogen Ion
H_2	Molecular Hydrogen
H_2O	Water
He	Helium
K	Potassium
$MgSiO_3$	Pervoskite
Na	Sodium
Ni	Nickel

O/H	Oxygen-to-Hydrogen ratio
SiO	Silicon Monoxide
Ti	Titanium (chemical symbol)
TiO	Titanium Oxide
VO	Vanadium Oxide

Statistical Symbols:

$\mathcal{U}$(min, max)	Uniform distribution
$\mathcal{N}$(mean, width)	Normal distribution
$\mathcal{LU}$(min, max)	Log-uniform distribution
$\mathcal{Z}$	Bayesian Evidence
$\mathcal{O}$	Odds Ratio
BF	Bayes Factor

Acronyms:

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
ESPs	Earth-sized Planets
GO	General Observer
HAT	Hungarian-made Automated Telescope
HST	Hubble Space Telescope
HZ	Habitable Zone
IR	Infrared
IRAC	Infrared Array Camera
JWST	James Webb Space Telescope
KELT	Kilodegree Extremely Little Telescope
LD	Limb-Darkening
LTE	Local Thermodynamic Equilibrium
MAST	Mikulski Archive for Space Science
MCMC	Markov Chain Monte Carlo

NASA	National Aeronautics and Space Administration
NICMOS	Near Infrared Camera and Multi-Object Spectrometer
NIR	Near Infrared
NIRISS	Near InfraRed Imager and Slitless Spectrograph
NIRSpec	Near InfraRed Spectrograph
NUV	Near Ultraviolet
OOT	Out-Of-Transit
PanCET	Panchromatic Comparative Exoplanet Treasury
PASP	Publications of the Astronomical Society of the Pacific
PI	Principal Investigator
PLATON	PLanetary Atmospheric Tool for Observer Noobs
RMS	Root-mean-square, or standard deviation
S/N	Signal-to-Noise
SNR	Signal-to-Noise Ratio
SOSS	Single Object Slitless Spectroscopy
STIS	Space Telescope Imaging Spectrograph
STScI	Space Telescope Science Institute
TESS	Transiting Exoplanet Survey Satellite
TIC	TESS Input Catalog
TICv8	TESS Input Catalog version 8
T-P	Temperature-Pressure
TRAPPIST	TRAnsiting Planets and PlanetesImals Small Telescope
TRES	Tillinghast Reflector Echelle Spectrograph
TSM	Transmission Spectroscopy Metric
UMD	University of Maryland
UV	ultraviolet
VLT	Very Large Telescope
WASP	Wide-Angle Search for Planets
WFC3	Wide Field Camera 3
UVIS	WFC3/Ultraviolet and Visible Light
Z	Metallicity

Chapter 1: Introduction

Atmospheres are windows into their host planets' characteristics. They contain measurable information on their planets' chemistry and climate, which directly informs atmospheric processes and can even provide insight into the formation and evolution of a planet. Understanding these physics is an interesting endeavor, since planets are more diverse and not as "neat" as stars. We do not have to look further than the Solar System to see that there is no obvious main sequence for planetary atmospheres. Despite similar radii, similar orbits, and the exact same formation conditions (in terms of host star and protoplanetary disk), Earth (N_2 and O_2) and Venus (CO_2) evolved to have two wildly different atmospheres. The rest of the Solar System also exhibits astonishing diversity, with cold ice giants, hydrogen dominated gas giants, and atmosphere-free terrestrial planets. Fortunately, thanks to observing missions such as Kepler and TESS, we have access to thousands of test cases in the form of planets outside of the Solar System (extrasolar planets, or exoplanets). The bulk properties of exoplanets prove to be much more varied than the Solar System, allowing the ability to test atmospheric models over a range of temperatures, radii, and host star properties. By modeling and observing exoplanet atmospheres, we can better understand both atmospheric processes and planetary diversity, and place the

Solar System in a greater context to understand how unique it is, if at all.

1.1 Background

The first step in characterizing the atmosphere of a planet is detecting the planet. Figure 1.1 gives context for the extreme increase in the detections in the last few years (left panel), as well as a sample of roughly characterizable planets (right panel). These planets are all in the Milky way, typically on the order of 100 parsecs away. This makes them close enough to observe, but too far away to spatially resolve from their host star. Atmospheres are observable, but only indirectly, primarily utilizing two detection methods. In transit detections (used by Kepler and the **T**ransiting **E**xoplanet **S**urvey **S**atellite, TESS), the light from stars in a wide angle of sky is monitored for a long time, and periodic dips indicate that a planet is periodically orbiting in front of a star and blocking a small fraction of light. The size of the dips reveal, among other properties, the radius of the planet relative to the star. Larger planets with smaller orbits and smaller host stars are most easily observed in transit. Radial velocity (RV) detections (e.g, **H**igh **A**ccuracy **R**adial velocity **P**lanet **S**earcher HARPS) monitor wide angles of the sky for star "wiggles" (i.e, doppler-shifted spectra), which are indicative of the star orbiting a joint planet-star center of mass. RV measurements provide the mass of the planet. High mass planets with low mass host stars are more easily observed in RV. The typically path to atmospheric characterization of a planet is: detected by transit (radius), followed up with RV measurements (mass), which combined provide density, gravity and

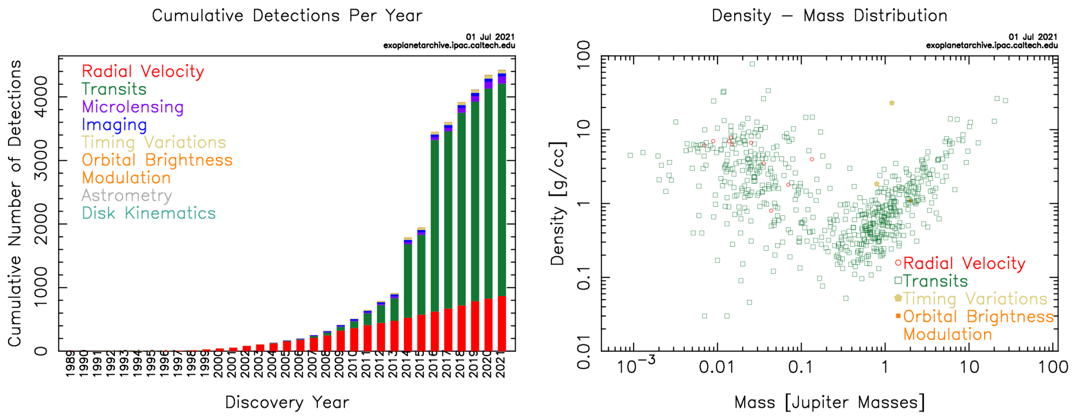


Figure 1.1: Exoplanet detection information from the NASA Exoplanet Archive. **Left:** Cumulative number of confirmed planets each year. **Right:** Sample of planets with both mass and radius measured.[a]

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

amenability to atmospheric characterization.

We then follow up promising planets with transit or eclipse spectroscopy. The infographic in Figure 1.2, taken from Kreidberg (2018), helps demonstrate the geometry of this process. Given the distances involved, we observe the combined planet-star light as a single point. The idea behind transit and eclipse spectroscopy is to use relative, wavelength-dependent changes in the combined planet-star light to indirectly infer the spectrum of the planet. Primary transit occurs when a planet passes in front of its host star. If a solid, atmosphere-free planet passes in front of a star, it will be equally opaque at every wavelength (since solids block all light in the NUV-IR range typical of observations) and so the planet will look the same size regardless of wavelength. However, if it has a gaseous atmosphere, that gas will interact with light in different ways depending on its composition. Therefore, the atmosphere can then leave its imprint on light that passes through it. Depending on its composition and physical properties, different amount of lights will be blocked

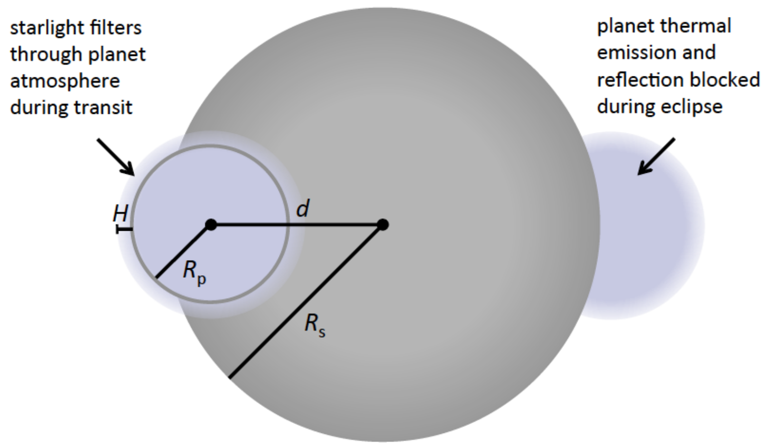


Figure 1.2: Transit and Eclipse Geometry, from Kreidberg (2018)

at varying heights for different wavelengths. Since atomic and molecular gaseous species have different spectral signatures, we can relate the changes in depth with specific atoms/molecules. An atmosphere of all water will look bigger at 1.4μm, where water is opaque, than at 0.5μm, where it has a relatively low opacity. The amount of light blocked is known as a transit depth, $(R_p/R_s)^2$. We can match observed transit depth variation with wavelength with spectral signatures to infer composition and physics in the atmosphere. Transit spectroscopy measures the limb of the atmosphere, otherwise known as the day-night terminator, and typically probes high in the atmosphere at pressures around 1 mbar.

In eclipse spectroscopy, we observe the system immediately before the planet goes behind the star, and thus get a total flux of the system. When the planet passes behind the star, we measure only stellar flux. We compare this flux ratio to total flux to infer emergent flux from the planet itself. The change in flux corresponds to the composition of the atmosphere and, unlike transits, is also highly sensitive to the thermal structure of the atmosphere. Eclipses probe the dayside of the atmosphere (i.e, the half of the planet facing the host star, which is often permanent since — due to circularization — many planets at orbits this close are inferred to be tidally locked), at pressures deeper in the atmosphere around 100 mbar. Phase curves, where entire orbits of an exoplanet are observed to better inform climate and atmospheric dynamics, are another popular characterization method.

Space-based instruments are, with good reason, the most common way to perform transit and eclipse spectroscopy. Though ground-spaced spectroscopy is useful, particularly for high-resolution cross-correlated spectra (Birkby, 2018), it faces an extra challenge of correcting for Earth's time-variable atmosphere. In addition to avoiding this complication, space-based instruments are able to observe in wavelengths at which Earth's atmosphere is opaque (e.g, water bands are too difficult too observe from the ground), and they have much cooler thermal background noise, which is especially important for IR observations. Consequently, space-based instruments HST/WFC3 (1.1–1.7μm; near-IR), HST/STIS (0.3–0.9μm; near-UV-optical), and *Spitzer* IRAC (3.6–8μm; mid-IR) have been the most productive instruments for characterizing exoplanet atmospheres (e.g, Benneke & Seager, 2013; Deming et al., 2013; Mandell et al., 2013; Kreidberg et al., 2014b; Stevenson et al., 2014, among countless others).

These instruments cover separate wavelength regimes which are able to probe different pressures, physics, and spectroscopically active chemical species. Combining observations from multiple instruments maximizes the spectral baseline and allows for the most complete characterization of exoplanet atmospheres (Nikolov et al., 2014; Sing et al., 2016; Beatty et al., 2017; Mansfield et al., 2018; Chachan et al., 2019a). Further, combining optical and IR spectra can break potential degeneracies between molecular abundances and planetary properties (Griffith, 2014; Line & Parmentier, 2016; Welbanks & Madhusudhan, 2019). One must be careful combining data from different instruments since the absolute depth is typically less well constrained then the *relative* change in depth (i.e, the shape of an instruments spectrum). Systematic errors could potentially bias the depths of an instrument relative to other instruments, and that should be considered in analysis (Garhart et al., 2020).

1.2 Atmospheric Retrievals

To derive atmospheric properties of an exoplanet from transmission or emission spectroscopy, we must solve the inverse problem of "what chemical composition and physics produce the observed spectrum?" Atmospheric retrieval is the process of *retrieving* chemical and physical information about an atmosphere based on an observed spectrum (Irwin et al., 2008; Madhusudhan & Seager, 2009; Lee et al., 2012; Line et al., 2013; Benneke & Seager, 2013; Waldmann et al., 2015b). The two primary components of atmospheric retrieval are forward models (a spectrum generated given exactly known atmospheric properties) and a parameter estimation method (for a given model, which properties lead to the best fit to the data?). In total, the observed spectrum is compared to many simulated spectra (from forward models) in order to constrain the possible values of parameters of interest (such as planet radius, temperature, or water abundance). This is not a novel process: Rodgers (2000) used atmospheric retrieval to analyze remote sensing data of Earth, and Irwin et al. (2008) similarly used atmospheric retrieval on Solar System planets. An infographic summary from a recent review paper (Madhusudhan, 2018) is shown in Figure 1.3.

First, I summarize the forward modeling of exoplanet transit and eclipse spectra. Through both quantum mechanics and laboratory experiments, we understand how light interacts with many different molecules and atoms (Barber et al., 2006; Rothman et al., 2009, 2010; Tennyson & Yurchenko, 2012; Allard et al., 2019). This is important, since knowing the spectral signatures of species can allow us to learn about the composition of an atmosphere from an observed spectrum.

Atoms generally interact via spectral lines with shapes dictated by energy level transitions (and various sources of broadening). Molecules typically form "bands", which are combinations of millions of electronic, vibrational, and rotational transitions (Tennyson & Sutcliffe, 1982). The shapes and magnitudes of these molecular bands or atomic lines constitute the spectral feature for a given species. The shape of these spectral features, for a given temperature and pressure, are constantly being updated in libraries of line lists such as EXOMOL (Tennyson & Yurchenko, 2012; Tennyson et al., 2016; Tennyson & Yurchenko, 2018; Tennyson et al., 2020), HITEMP (Rothman et al., 2009), and HITRAN (Rothman et al., 2010). These libraries are particularly relevant to exoplanet atmospheres since they include spectral features of species at higher temperatures.

Included species are informed by chemistry (e.g, the carbon and hydrogen are common elements, so H_2, CH_4, and C_2H_2 should be considered), as well as observations of Solar System planet atmospheres (e.g, Hubbard et al., 2002), cool stellar atmospheres (e.g, Allard et al., 1996; Tsuji et al., 1996), and, importantly, similar temperature brown dwarf atmospheres (e.g, Burrows & Sharp, 1999). Though hundreds of species are chemically expected, many are at such low abundances in chemical equilibrium, or have such low opacity in telescope wavelength bands, that they can be neglected in analyses (Kempton et al., 2017). This generally results in about 30 relevant species. Of special importance are water, CO, CO_2, CH_4, Na, K, TiO, and VO. I emphasize that these are all gaseous species. In an exoplanet atmospheric context, water always means water vapor. If the temperature was low enough (at a given pressure) for a species to condense into a cloud, it would interact with light in a much different, grayer (constant with wavelength) manner.

Forward models utilize line lists for these species in order to derive the opacity (effective area of an interaction per unit mass of material) for each molecule and atom. Continuum opacity sources such as collision induced H_2-He absorption, H- bound-bound and bound-free interactions, and Rayleigh scattering are also generally included as important opacity sources. Finally, opacity due to aerosols is considered. This is broken down into two groups: clouds (similar to rain clouds on Earth; condensates which form due to thermochemical equilibrium) and hazes (small, complex particles likely created by the combination of photochemical byproducts of molecules). Cloud and haze opacities are often treated as a flat line (at a cloud top pressure) and a scaled version of Rayleigh scattering, respectively, since a full microphysical treatment is computationally expensive.

With opacity sources well understood, it is possible to model a transit (or emission spectrum). In modeling exoplanet atmospheres, a useful philosophy to have is summarized by a quote famous statistician George Box: "All models are wrong, but some are useful." A completely self-consistent 3D model that simultaneously accounts for disequilibrium photochemistry would be a relatively accurate atmospheric model, but it is prohibitively computationally expensive to be useful for retrievals, which often require hundreds of thousands of model evaluations. Accordingly, approximations, such as 1D atmospheres (the temperature and abundances at each longitude and latitude are uniform, only pressure (height) matters) are necessary. Fortunately, we can check approximations both against full treatments and better quality Solar System data. For example, the atmospheric physics used for exoplanets are derived from stellar atmospheres, and have been validated on stars (Allard et al., 1996; Tsuji et al., 1996), brown dwarfs (Burrows & Sharp, 1999) and Solar System planets (Hubbard et al., 2002). Further, the exoplanet atmosphere modeling community is good at self-regulating. Madhusudhan & Seager (2009) validated that their temperature-pressure (T-P) profile parameterization is able to fit the T-P profile of each Solar System planet. Both Burrows et al. (2010) and Fortney et al. (2010) compared 1D models to 3D general circulation models (GCMs) and found tolerable agreement. Blecic et al. (2017) similarly benchmarked the performance of 1D models to more updated 3D GCM models for several data resolutions. Line et al. (2013) found consistent results between their five-parameter T-P profile to the level-by-level approach used in Earth analyses. As a final example, Kempton et al. (2017) showed that a transit spectrum generated from an isothermal temperature profile sufficiently matched one generated from a self-consistent radiative-convective T-P profile.

In summary, the bulk of the literature finds the error caused by computationally useful approximations to be generally negligible for transit and eclipse analyses as compared to current observational precision. Additionally, the physics used to predict model atmospheres successfully describes higher-quality observations from stars, brown dwarfs, and Solar System planet atmospheres.

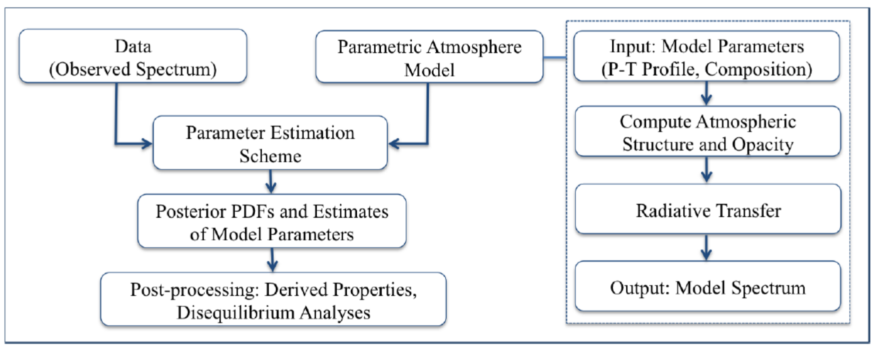


Figure 1.3: Summary of the Atmospheric Retrieval Process, from Madhusudhan (2018).

Forward models as used in retrievals are summarized in the rightmost panel of Figure 1.3. They generally work as follows: assume a 1-D, plane-parallel atmosphere; set a T-P profile; set abundances at each T-P point for each important species (potentially using chemical equilibrium); calculate opacity of each species for a given temperature and pressure for each wavelength; determine physical height associated with each pressure by solving hydrostatic equilibrium (Eq 1.1; μ is the mean molecular weight of the atmosphere); perform an abundance-weighted sum and combine with Rayleigh scattering to get total opacity at each atmospheric layer.

$$\frac{dP}{dr} = \frac{-GM}{r^2}\frac{\mu m_{amu} P}{kT} \tag{1.1}$$

The final step is a radiative transfer calculation, where an incident ray from the star is traced through atmospheric layers and towards the observer at Earth. The optical depth experienced by a ray of a given impact parameter is determined by the "width" and total opacity of each layer it passes through. The amount of light blocked by the atmosphere at each wavelength is determined by the optical depth experienced by rays at different impact parameters. The apparent size of the planet due to this blocked light gives the transit depth at that wavelength. This is formalized in Equation 1.2. D_λ is the transit depth at wavelength λ, τ_λ is the optical depth at wavelength λ. The transit depth at different wavelengths gives the modeled transit spectrum.

$$D_\lambda = (R_{\rm bottom}/R_s)^2 + 2\int_{R_{\rm bottom}}^{\infty} \frac{r}{R_s^2}(1 - e^{-\tau_\lambda})\, dr \tag{1.2}$$

A similar process is used to model eclipses, but instead of light blocked by the atmosphere now flux emitted by the planet at each wavelength (and stellar flux) is important. Additionally, the T-P profile is more important, as emitted light is more sensitive to temperature gradients (Kempton et al., 2018). Self-consistent models are computationally expensive, so profiles are often parameterized such that different profile slopes can be captured (Madhusudhan & Seager, 2009; Guillot, 2010; Line et al., 2012). For a given set of T-P parameters, a T-P profile is generated. Again, abundances are set at each pressure layer, and a height is assigned to each pressure via hydrostatic equilibrium. Total opacity is similarly determined for each wavelength, and the emergent planetary flux is determined by assuming the source function is the Planck function (i.e, neglect scattering, which is too computationally expensive to model) and summing contributions over each pressure layer for entire planetary disk. Stellar flux is often determined from a stellar spectral library. The exact eclipse depth (D_λ) is given in Equation 1.4 and depends on the planetary flux ($F_{p,\lambda}$) calculated in Equation 1.3 (Zhang et al., 2020a). $B_\lambda(\tau_\lambda)$ is the Planck function (B) at optical depth τ, and it is integrated over both optical depth and the viewing angle (μ).

$$D_\lambda = \left(\frac{R_{p,\lambda}}{R_s}\right)^2 \frac{F_{p,\lambda}}{F_{s,\lambda}} \tag{1.3}$$

$$F_{p,\lambda} = \int_0^\infty \int_0^1 B_\lambda(\tau_\lambda) e^{\frac{-\tau_\lambda}{\mu}} \, d\mu \, d\tau_\lambda \tag{1.4}$$

The parameter estimation component is best accomplished with a Bayesian sampler. As temperature, composition, and other physical parameters such as planetary mass vary, so will the output spectrum, but in a non-linear way; temperatures impact opacities, as well as abundances (e.g, condensation). Gaseous species abundances impact both optical depth and scale height. Further, there are many degeneracies (Seager & Sasselov, 2000; Madhusudhan & Seager, 2009; Line & Parmentier, 2016; Welbanks & Madhusudhan, 2019). Accordingly, Bayesian inference is a necessary tool to properly retrieve values and uncertainties of important parameters, such as temperature and water abundance. This involves setting reasonable priors on parameters of interest, and using an intelligent sampler, such as Markov Chain Monte Carlo (MCMC Goodman & Weare, 2010) or nested sampling (Skilling, 2004), to fully explore prior parameter space and finely sample the posterior near its peak values. Bayesian inference can provide both credible intervals for parameters and the Bayesian evidence, which is useful for model comparison. As an example, the ratio of Bayesian evidences of a model with water opacity to one without water opacity gives the strength of a water detection.

1.3 Hot Jupiter Climates

Due to their high temperatures and large radii, planets in the hot Jupiter archetype are the most amenable to characterization via emission spectroscopy (eclipse signal $\propto R_p^2 \times T_p$). The first thermal emission from an exoplanet was detected by Deming et al. (2005) and Charbonneau et al. (2005), and since then WFC3 and *Spitzer* data have been jointly used to constrain the dayside atmosphere of tens of hot Jupiters. Early theories on hot Jupiter atmospheres were generally driven by two connections: extrapolating models of brown dwarf atmospheres to lower temperatures and adding stellar irradiation, or assuming they are higher temperature analogues of Solar System gas giant atmospheres. These early predictions were iteratively tested against observations, and more complicated physics would be invoked if the models could not describe the data. This process contextualizes the current state of the field.

Hot Jupiters typically lie on orbits which are 20% the orbital distance of Mercury, making them both tidally-locked and subject to intense irradiation. This strongly influences their thermal structure, which is also sensitive to many other factors (such as opacity and advection) and are complicated to model (Fortney et al., 2021). There is an additional sub-branch of hot Jupiters known as ultrahot Jupiters (typically T_{eq} >2300 K). For context on important gaseous species, I show emission spectrum *cross sections* at typical hot Jupiter temperature 2000 K in Figure 1.4. A cross-section is a temperature, pressure, and wavelength-dependent "effective area" of a particular photon-particle interaction. It is a measure of the likelihood of a photon interacting with a single particle via that interaction. The cross sections are derived from the model presented in Gandhi & Madhusudhan (2018), though the figure is from private communications.

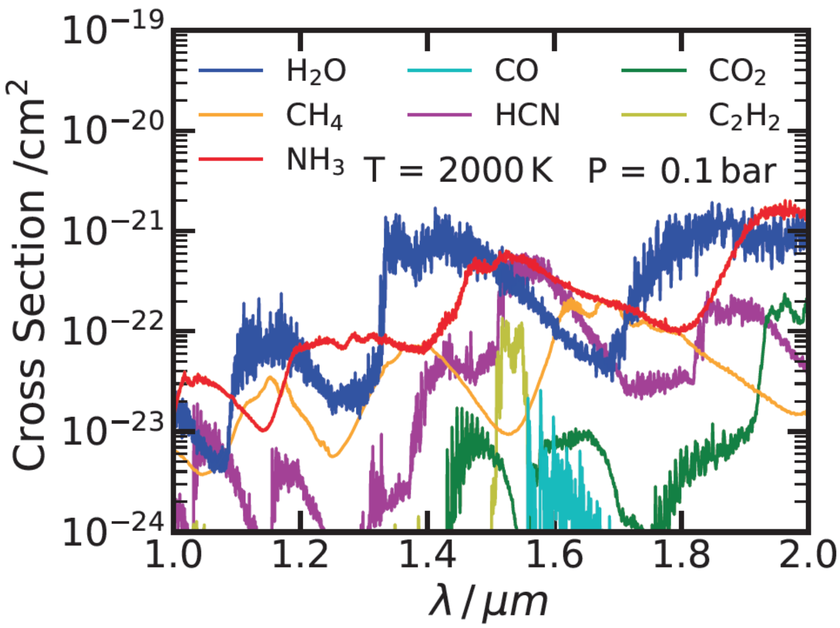


Figure 1.4: Relevant opacity sources in typical hot Jupiter emission spectroscopy. Taken from Gandhi & Madhusudhan (2018), this only covers WFC3 wavelength range. Note these are cross-sections, not opacities, meaning it does not take abundance into account. For example, H_2O and NH_3 have similar cross sections, but water is almost always at a greater abundance and thus typically dominates opacity.

Recently, GCMs — which couple fluid dynamics and radiative transfer — have been used to estimate the self-consistent 3-D structure of hot Jupiters via the meteorology primitive equations (Lewis et al., 2014; Kataria et al., 2015; Carone et al., 2020). These revealed that an eastward-shifted hotspot is the norm, and that large day-night temperature contrasts are typical. This is consistent with the Cowan & Agol (2011) inference of poor heat recirculation efficiency in hot Jupiters based on high dayside temperatures. GCMs are computationally expensive, and so 1-D temperature-pressure (T-P) profiles are often used in practice, especially with atmospheric retrievals which require many model evaluations. These rely on the decent approximation that the integrated dayside hemisphere can be described by a single T-P profile, which probes the change in temperature with height. They are typically parameterized to capture three profile shapes: isothermal, decreasing (temperature decreases with height), or a thermal inversion.

Of particular interest are stratospheric thermal inversions, where temperature increases with height. Hubeny et al. (2003) hypothesized that, given that TiO and VO are observed in brown dwarfs, and given their relatively high opacity in the optical at low pressures, that they would absorb the intense stellar irradiation more quickly than IR opacity could radiate, causing temperature to increase to maintain radiative equilibrium. Fortney et al. (2008) built off this by classifying planets by both temperature and TiO/VO presence, and predicted TiO and VO to be gaseous and spectroscopically active above 1600K. However, this classification is complicated by the non-ubiquity of TiO/VO evidence and the many mechanisms to take it out of the atmosphere. It could fall victim to a cold trap (vertical (Spiegel et al., 2009) or nightside (Parmentier et al., 2013)) for temperatures roughly below 1900K. It could be photodissociated by the UV light from an active host star (Knutson et al., 2010). On the opposite end, it is predicted to be thermally dissociated at temperatures above 3200 K (Lothringer et al., 2018). This leaves a relatively small "goldilocks zone" for TiO and VO to be spectroscopically active. Lothringer et al. (2018) predicted that regardless of TiO/VO, metallic atom and ion opacity would be enough in planets above 3200 K to drive a thermal inversion. Though this would be more akin to a thermospheric inversion (which is typical, though not often probed in exoplanet transits) being pushed down to observable pressures rather than a stratospheric inversion. Finally, Mollière et al. (2015) hypothesized that in atmospheres with bulk carbon-to-oxygen ratios (C/O) close to one, the water abundance would drop resulting in limited IR opacity, and so the lack of efficient cooling would drive an inversion.

Another interesting question relating to hot Jupiter climates is the paucity of water vapor in emission spectroscopy. Water has been detected as an absorption feature several times, typically in cooler hot Jupiters (T_{eq} < 1500 K) (Crouzet et al., 2014; Kreidberg et al., 2014b; Line & Parmentier, 2016; Beatty et al., 2017). However, it has only been tenatively observed in emission in two hot Jupiters (Haynes et al., 2015; Evans et al., 2017). The thermal structure dictates the type of feature: decreasing thermal profiles appear as absorption dips (since the higher opacity water band is sampling Planck function higher in the atmosphere where it is cooler), inversions appear as emission bumps, and isothermal atmospheres appear flat. Given that ultra-hot Jupiters are expected to host thermal inversions, and that water is well mixed in hot Jupiter atmospheres (Madhusudhan, 2019a), emission features were predicted to be more common. In completely featureless spectra explanations range from the precision being too low (Wilkins et al., 2014) to an isothermal region of the atmosphere being sampled (Nikolov et al., 2018). When water is not seen in conjunction with a CO detection, a high C/O is inferred (Stevenson et al., 2014), since oxygen would be sequestered in CO and a limited amount would be available for water (Madhusudhan, 2012; Moses et al., 2013). More recently, Lothringer et al. (2018) and Parmentier et al. (2018) hypothesized that, in ultrahot Juipters, molecular dissociation and H- opacity become significant. At high enough temperatures, H- opacity can mask water opacity and water can be thermally dissociated, resulting in water features being muted in both the WFC3 and *Spitzer* bandpasses. Additionally, these changes in opacity impact the thermal structure of the atmosphere, cuasing water to only be spectroscopically active in the deep, isothermal layers of the atmosphere. The net result is a featureless blackbody spectra with only CO being a potential feature in the MIR since its strong molecular bonds prevent it from being thermally dissociated.

For context, I share Figure 4 from Parmentier et al. (2018), which details relevant opacity sources in emission in the WFC3-*Spitzer* wavelength range at a typical ultra-hot Jupiter temperature (3100 K).

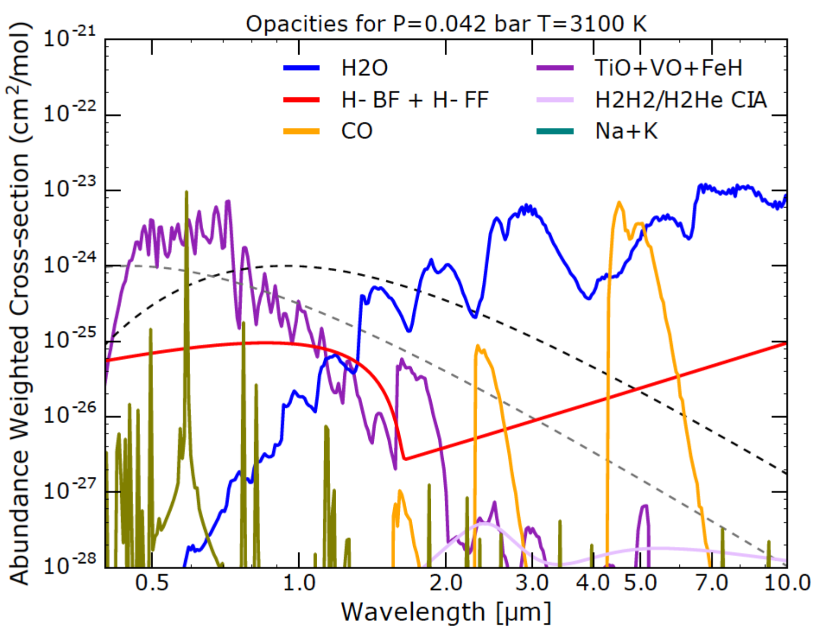


Figure 1.5: Relevant opacity sources in typical ultrahot Jupiter emission spectroscopy. Taken from Parmentier et al. (2018).

1.4 Hot Jupiter Composition and Chemistry

Bulk density measurements (mass and radius of entire planet, not just the atmosphere) reveal that hot Jupiters are H_2/He-dominated, similar to Jupiter (Lissauer & Stevenson, 2007). This matches predictions of core-accretion formation theory, where a rocky planetessimal core accretes enough mass to trigger runaway gas accretion (Pollack et al., 1996). This formation model predicts a bulk mass-metallicity trend amongst planets (similar to the solar system; Mordasini, 2014), and that prediction was born out in observations (Thorngren et al., 2016). The formation mechanism of hot Jupiters is not perfectly known. The three most prominent theories are *in-situ* formation (form at their current location), disk-migration (form far from host star but migrate inwards through planetary disk), and disk-free migration (form far away, become perturbed onto elliptical orbit after disk dissipates and migrate inward via tidal dissipation) (Fortney et al., 2021). In the core-accretion paradigm, proto-hot Juptiers cannot accrete enough gas to reach observed masses and radii — the feeding zone is too small. Inward migration is more likely, and its possible both formation mechanisms are common. For example, some planets have orbits aligned with their host stars spin (naively expected for disk-migration), but others have a significant misalignment (naively expected for disk-free migration). The formation history of a planet, though highly stochastic, still causally impacts the planet's eventual atmosphere, and so the atmosphere may contain hints to formation history (Mordasini et al., 2016).

Due to their formation mechanism, planets are thought to approximate the elemental abundances of their host stars. This is only to the first order though, since disk-migration and the evolution of the disk itself (e.g, snow-lines) impact the eventual atmospheric properties of an accreting planet (Öberg et al., 2011; Madhusudhan et al., 2014a; Mordasini, 2014). On top of that, the relative gas/ice accretion, pebble accretion, and pebble *drift* all directly impact the atmospheric metallicity (ratio of "metals"[1] to hydrogen; typically O/H in planets' atmospheres) and C/O ratio (Mordasini et al., 2016; Madhusudhan, 2019a). In particular, pebble drift, which describes inward-drifting pebbles sublimating as they cross snow-lines and enriching the metallicity of gas, can allow for metal-enriched planets with a variety of C/O ratios (Booth et al., 2017).

C/O ratios are a common parameterization of exoplanet atmospheres (Madhusudhan, 2012; Moses et al., 2013). This is because 1) C and O are the two most common metals 2) most spectroscopically active species in the optical to MIR are C- or O-based and 3) the abundances profiles in chemical equilibrium differ drastically as C/O ratio approaches and crosses one. It is estimated in exoplanet atmospheres either from chemical equilibrium assumptions or directly if a C-bearing and O-bearing molecular abundance is determined. Though C/O ratios are often debated and there are very few definitive C-bearing molecule detections (not coincidentally because common molecules CO and CO_2 have strongest cross-sections by *photometric Spitzer* points), hot Jupiters are primed for abundance determination. Their temperatures are high enough to vaporize most species and to keep them aloft and

[1]Astronomers define metals as any element heavier than hydrogen and helium. This contrasts with the chemical/physical definition (an electrical conductor). For example, though carbon and oxygen are classified as non-metals in the periodic table, they are considered metals in an astronomical context.

well-mixed in the atmosphere. In fact, the relative coldness of the Solar System gas giants means that we know more about water abundance in planets hundreds of parsecs away than we do about those in the Solar System (though recent *in-situ* probes have helped even the playing field, such as *Galileo* (e.g, von Zahn et al., 1998; Owen et al., 1999) and *Juno* (e.g, McComas et al., 2017; Li et al., 2020)).

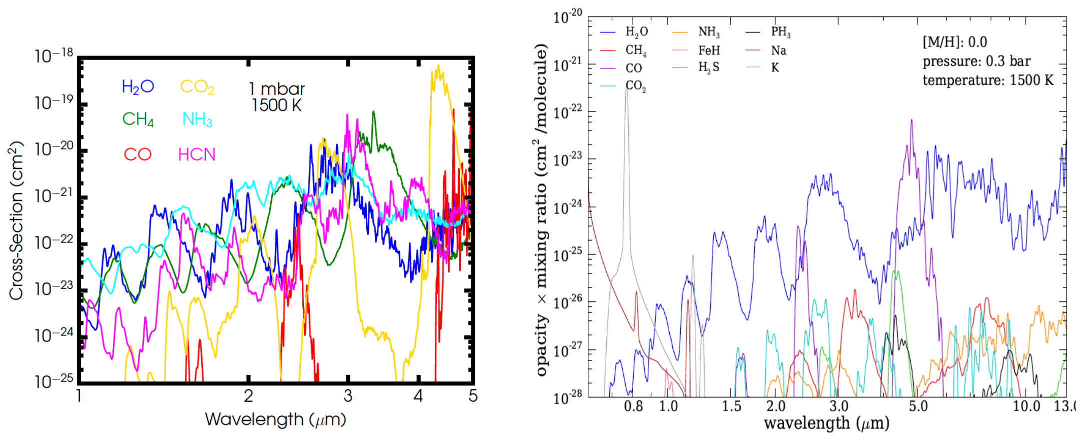


Figure 1.6: Relevant opacity sources in typical hot Jupiter transit spectroscopy. Taken from Pinhas et al. (2018) (left) and Kreidberg (2018) (right). Note the Pinhas et al. (2018) values are cross-sections, meaning they do not take the abundance of each species into account. The Kreidberg (2018) values are abundance-weighted opacities for a representative atmosphere, and cover optical wavelengths. Potentially important opacities TiO and VO are absent since they are expected to condense out of the atmosphere below 1600K.

The cross-sections of relevant molecules for typical transiting exoplanet temperatures are shown in Figure 1.6, taken from Pinhas et al. (2018); Kreidberg (2018). Many of these species have been detected in hot Jupiter atmospheres: water (Fraine et al., 2014; Evans et al., 2016, and many more), Na (Charbonneau et al., 2002; Nikolov et al., 2014, etc.), K (e.g, Sing et al., 2011), TiO (Sedaghati et al., 2017, though contested), and CO_2 (Morley et al., 2017). These detections can lead to abundance constraints, which — under the assumption of that the atmospheres is in a state of chemical equilibrium — can be used to estimate the amount of elemental C or O (and thus C/O) in an atmosphere. Abundance constraints are much more reliable in multi-instrument spectra, since optical and IR spectra can break degeneracies between abundance and clouds, hazes, or the reference pressure (Line & Parmentier, 2016; Welbanks & Madhusudhan, 2019).

Similar to eclipse spectroscopy, in which water features were not observed as frequently as predicted, WFC3 transit observations commonly found water features to be smaller in size than expected based on the slant path of stellar irradiation through the atmospheric limb (Deming et al., 2013; Stevenson, 2016; Fu et al., 2017). Water features have been found to typically cover only 2 scale heights instead of the 5–10 predicted for a saturated spectral feature in a clear atmosphere (Madhusudhan, 2019a). Though the cause was not clear, this effect was not necessarily unexpected, as early models predicted that atmospheric clouds or hazes could result in the muting of atomic and molecular gaseous absorption features (e.g, Brown, 2001; Fortney, 2005). Still, viable alternative explanations such as oxygen depletion (i.e, a high C/O ratio; Madhusudhan, 2012; Crouzet et al., 2014) or a high mean molecular weight (Line & Parmentier, 2016) are also plausible.

Sing et al. (2016) presented ten optical-to-IR spectra to argue that hazes and clouds were responsible for the majority of muted features. However, the cause is still unclear. A recent population study by Welbanks et al. (2019) argued that although atoms Na and K are typically supersolar, water is generally depleted as compared to Solar System extrapolations, even when accounting for clouds (Figure 1.7). Separately, Line & Parmentier (2016) demonstrated that, without optical data, there is a degeneracy between partial cloud coverage (i.e, clouds only covering a fraction of the planet's limb) and abundance. A high mean molecular weight can also describe the subdued feature sizes. Clouds (typically treated as grey opacity condensates at a fixed height) and hazes (which manifest as a steep slope increasing in the optical) are a unique problem since their opacity tends to mute or even hide gaseous features. Though there are many ideas on the make-up on these clouds and hazes (e.g, perovskite, corundum, or silicate clouds and hydrocarbon soot hazes in hot Jupiters), the microphysics is complicated and the source of clouds remains elusive (Kreidberg, 2018). These are often modeled parametrically when retrieving atmospheric properties (Zhang et al., 2019).

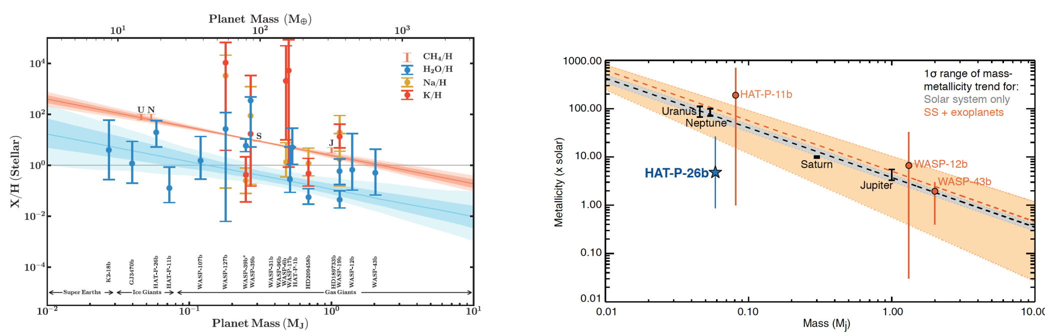


Figure 1.7: Extrapolating the Solar System mass-atmospheric metallicity using several different tracers for metallicity, from Welbanks et al. (2019) (left) and Wakeford et al. (2017) (right). Welbanks et al. (2019) find supersolar Na and K, but a consistent relative underabundance of water. Wakeford et al. (2017) suggested exoplanets generally follow the Solar system trend. This is highly dependent not only on retrieval method, but also on molecular abundance-to-metallicity conversion method.

Similar to bulk properties, a mass-atmospheric metallicity trend has been proposed for exoplanets (Kreidberg et al., 2014b; Wakeford et al., 2017; Mansfield et al., 2018). This would match what is seen in the Solar System planets and is predicted — albeit with significant intrinsic scatter — by formation models (Fortney et al., 2013). Figure 1.7 shows two recent observational investigations of this trend from

Welbanks et al. (2019) and Wakeford et al. (2017). While Welbanks et al. (2019) does a systematic study (re-deriving abundances for each planet with the same retrieval methodology), there is much variation on abundance constraints between the two studies (e.g, the median HAT-P-11b abundance differs by a factor of 1250 between the two plots). Additionally, as Heng (2018) points out, converting water to O/H to compare to solar is a pressure and temperature-dependent (e.g, there is not a single solar water abundance for which to compare), so derivation of metallicities may be flawed. Still, populating this plot (and cross-checking multiple methodologies) is useful in determining if such a trend is real, and this is possible through transit spectroscopy.

1.5 Earth-sized Exoplanet Atmospheres

Earth-sized planets (sometimes referred to as super-Earths or mini-Neptunes) are the most commonly discovered planet type over the last several years (Bean et al., 2021). Though often grouped together, there is a well studied radius valley that splits this planets into two categories: $R > 1.75 R_\oplus$ (sub-neptunes) and $R < 1.75 R_\oplus$ (Earth-sized planets) (Fulton et al., 2017). In this dissertation, I primarily discuss the Earth-sized planets. The most interesting Earth-sized planets (ESPs) are found around M-dwarfs, the most common star in the galaxy, because they provide the best opportunity for identifying a terrestrial atmosphere due to their small radii (Charbonneau et al., 2009; National Academies of Sciences & Medicine, 2018; Benneke et al., 2019b).

This archetype of planet is comparable to Earth in size and is the same order-of-magnitude mass, making them likely terrestrial. The key difference is that they orbit their (cooler) host stars at distances roughly 3% of Earth's orbit. Despite this lower host star temperature, the orbital distance is small enough to make their equilibrium temperatures and incident stellar irradiation greater relative to Earth's. Still, these are most direct gateway to biosignatures and understanding habitable planets, who are their cooler cousins (pun intended).

In this planetary regime, we really only know about Earth and Venus. There is a question about how common each atmosphere type is: is Earth unique and Venus the norm, or vice versa, or neither? Or is the Mars — with no significant atmosphere — the norm? Understanding the range and degree of ubiquity of atmospheres for terrestrial planets contextualizes our Solar System.

Unlike the massive hot Jupiters in Chapters 3 and 4, atmospheres are not definitively the norm for rocky planets. Looking at our own Solar System, only 50% of the rocky planets possess obvious, thick atmospheres (sorry Mars). Atmosphere detections are even rarer for ESPs (Pidhorodetska et al., 2021). This is partially due to atmospheric escape (Watson et al., 1981). In the core-accretion model, isolation mass cores form from the build up of solids into planetessimals (Safronov & Zvjagina, 1969; Wetherill & Stewart, 1993; Pollack et al., 1996). These planetessimals collide violently and do not accrete any primordial atmosphere if the isolation mass is not reached until disk gas has already dissipated. Even if an atmosphere was accreted, giant impacts (like the one that formed the moon) would likely remove primordial gas (Howe et al., 2020). However, Misener & Schlichting (2021) argue that super

Earths both form in the presence of disk gas and maintain residual H/He from their primordial atmosphere. This is because the initial removal of the majority of the primordial atmosphere decreases the cooling time scale of the remaining atmosphere enough for it to efficiently cool before being stripped. Further, this primordial H/He atmosphere would influence any secondary atmosphere (vulcanism/outgassing after H_2 is stripped, or volatile deposits from comets (Swain et al., 2021; Mugnai et al., 2021)) and potentially lower the mean molecular weight, making the planet more amenable to atmospheric characterization. Additionally, (Howe et al., 2020) argues that an alternative formation mechanism — pebble accretion — would allow super Earth cores to assemble quickly enough to accrete gas before the disk dissipates. Impact erosion, the typically dominant atmosphere loss process, would be much less effective in the pebble accretion paradigm.

A primordial H_2-dominated atmosphere for an ESP is thus plausible and should not be dismissed *a priori*. More likely, however, are secondary atmospheres. Swain et al. (2021) argued that, on the terrestrial exoplanet GJ1132b, a secondary H_2 atmosphere was possible via hydrogen dissolving in vulcanic magma, allow it to be stored, then later being released back into the atmosphere due to tidal heating. Observationally, Moran et al. (2018) found that a hydrogen-rich atmosphere with high altitude clouds can explain observations of several terrestrial planets in the TRAPPIST system (Gillon et al., 2017; de Wit et al., 2018).

Expected signal size is another reason for the lack of a convincing atmospheric detection for an ESP. Atmospheres are most clearly detected via atomic or molecular absorption features; atmospheres may exist, but the signal size is such that features are undetectable with the precision of current ground-based telescopes and HST WFC3 (e.g, Diamond-Lowe et al., 2018; Pidhorodetska et al., 2021). There is also a selection bias, since the majority of characterizable terrestrial planets orbit M-dwarfs, which emit a greater UV flux than Solar-type G stars. This intense UV irradiation makes low mean molecular weight, H_2-dominated atmospheres — which produce the largest features and are the easiest to detect — unlikely. Additionally, the increased stellar variability of M-dwarfs — and the fact that water absorption is clearly present in their spectra — also complicates atmospheric detection (Rackham et al., 2018; Deming & Sheppard, 2017). High-altitude clouds may also mask atmospheres by making spectra appear as featureless flat lines (Diamond-Lowe et al., 2018). Figure 1.8, taken from (Kempton et al., 2017), provides opacities for several species at 1000 K. This is above the typical equilibirum temperature of around 500 K, but it is a better approximation of important species opacities than the previous opacity figures.

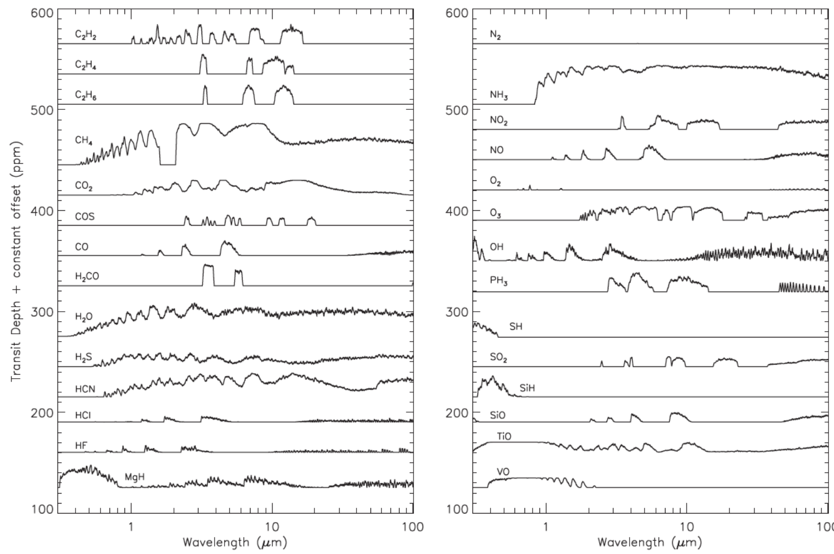


Figure 4. Spectra for atmospheres made up entirely of individual molecules. Each of the molecules with opacity data provided with `Exo-Transmit` is included on this plot. The spectra are calculated for a 2 $R_\oplus$ planet orbiting a 1 $R_\odot$ star with an Earth-like surface gravity of 10 m s^{-2}. The atmosphere has an isothermal temperature of 1000 K.

Figure 1.8: Approximation of relevant opacity sources in typical Earth-sized planet (ESP) transit. This is Figure 4 from Kempton et al. (2017).

GJ1214b (Berta et al., 2012), HD97658b (Kreidberg et al., 2014a), and the TRAPPIST system (de Wit et al., 2018) are ESPs which have all been observed with WFC3, and all display flat spectra. Though there is no molecular detection, this is at least informative since it can rule out a cloud-free, hydrogen dominated atmosphere. Swain et al. (2021) claimed a secondary H_2 atmosphere with an HCN and CH_4 detection in GJ1132b, but follow up data analysis by Mugnai et al. (2021) and Libby-Roberts et al. (2021) found a flat spectrum, consistent with earlier optical observations by Diamond-Lowe et al. (2018). Flat spectra are the norm for ESPs, since essentially the only easily detectable atmosphere is a clear, hydrogen-dominated one. Flatness could be caused by clouds, photochemical hazes, a high mean molecular weight secondary atmosphere (similar to Venus), even a giant spot or faculae on an active host star, which can act to cancel out spectral features (Rackham et al., 2018). The ability of ESPs to host atmospheres — and their detectability — is an active question in the super-Earth sub-field.

1.6 Outline

This dissertation takes a broad approach, analyzing both transit and emission spectroscopy in different sections of parameter space. This includes a range of planets from cool, Earth-sized planets (T$\sim$500K, R=0.8$R_\oplus$) up to massive, ultra-hot Jupiters (T$\sim$2500K, M=10$M_{\rm Jup}$). I emphasize clarity and sensitivity tests in data analysis and light curve fitting. I also emphasize properly contextualized statistics and how to combine data from different instruments. Deep dives into individual planets are important to properly account for uncertainty in model choice, to contextualize how results depend on assumptions, and to accurately represent results. This dissertation explores how different modeling assumptions impact results, both within a single model paradigm and between paradigms.

In Chapter 2, I derive and the transit spectra for 5 visits of two likely rocky exoplanets, L9859b and L9859c. I detail my custom analysis pipeline `DEFLATE`, which is the basis of all WFC3 data analysis in this thesis. I investigate the likelihood that the structure in the spectra are indicative of the first convincing rocky planet atmospheres, and if it is plausible for stellar activity to implant that structure onto a flat transmission spectrum. In Chapter 3, I perform a detailed multi-instrument, multi-retrieval transit analysis of the hot Jupiter HAT-P-41b. I explore dozens of retrieval assumptions and relate any disparities directly to — sometimes equally valid — modeling choices. The analysis reveals a 5σ water detection and a significantly super-stellar atmospheric metallicity in almost every single retrieval model. Finally, in Chapter 4, I analyze the thermal structure of two exoplanets, both on the boundary of potentially important chemistry in gravity-temperature parameter space. I find a thermal inversion for ultra-hot WASP-18b, water absorption in hot WASP-19b. I explore how both results provide insight into dominant physical processes on hot Jupiter atmospheres.

Chapter 2: Investigating a Potential Terrestrial Atmosphere in the L98-59 Multi-planet System

2.1 Introduction

The effort to understand the atmospheres of terrestrial exoplanets is a major goal of exoplanet science and a priority for JWST. However, the total number of currently characterizable Earth-sized exoplanets is in the single-digits (Pidhorodetska et al., 2021). Though there have been no convincing atmospheric detections, HST WFC3 observations of potentially rocky, Earth-size planets have been informative. Notably, they have ruled out cloud and haze-free hydrogen-rich atmospheres (de Wit et al., 2018), and shown that cloudy H_2 rich atmospheres or heavier molecules (e.g, CO_2 or H_2O) could explain observed transit spectra (Moran et al., 2018; Wakeford et al., 2019; Mugnai et al., 2021; Libby-Roberts et al., 2021).

L9859 is a small ($R=0.31R_\odot$) and bright ($T_{\text{eff}}=3400$K) M3V-dwarf roughly 10 parsecs away. TESS recently discovered it to be the host star in a multi-planet system, making it the second closest transiting multi-planet system to Earth (Kostov et al., 2019). Kostov et al. (2019) discovered three Earth-sized planets, and follow up HARPS RV observations (Cloutier et al., 2019) derived masses and concluded that the two innermost planets — L9859b ($R=0.8R_\oplus$, $M<1M_\oplus$, T_{eq}=550K) and L9859c ($R=1.35R_\oplus$, $M=2.4M_\oplus$, T_{eq}=470K) — have bulk densities consistent with terrestrial planets. Their orbital periods are both on the order of a few days, placing them at an orbital distance of roughly 3% of Earth's orbital distance. Given this close orbit, the planets receive much more irradiation than Earth, and are in the "Venus Zone" (Pidhorodetska et al., 2021).

Cloutier et al. (2019) calculated the atmospheric detection index (Kempton et al., 2018) as greater than the previous highest for a pre-TESS terrestrial planet, up to 1.6× that of GJ1132b. The expected signal size from two of the planets — inner-and-smaller b, and outer-and-bigger c — thus merited follow up observations and 5 HST WFC3 transits were awarded to a collaborator, PI Tom Barclay. Pidhorodetska et al. (2021) highlighted that HST WFC3 is the most favorable instrument to conduct transit spectroscopy of the L9859 in the near-term.

The atmospheres are worth investigating. Pidhorodetska et al. (2021) further classified potentially characterizable atmospheric scenarios for L9859c. Of course, it could have no atmosphere — or at least a non-detectable one. Pidhorodetska et al. (2021) showed that 1 transit of L9859c and 4 transits for L9859b with HST WFC3 are informative: the precision could potentially detect a H_2-dominated atmosphere, and a water feature in a steam atmosphere, respectively. They also put forward runaway greenhouse (CO_2 dominated, similar to Venus; Kasting, 1988), and O_2-dessicated as plausible atmospheres, though neither would be detectable with HST WFC3's precision. Further, TESS monitoring showed no obvious stellar flares or activity (Kostov et al., 2019), potentially simplifying analysis (Wakeford et al., 2019).

However, see Section 2.5.2 for full stellar activity discussion.

Though Kostov et al. (2019) postulated that an H-rich atmosphere akin to larger gas giant planets was unlikely due to atmospheric escape, Pidhorodetska et al. (2021) showed that L9859c could retain a secondary H_2/He-dominated atmosphere (since its equilibrium temperature, depending on albedo, is plausibly lower than the hydrogen escape temperature 510 K). Pidhorodetska et al. (2021) also demonstrated that atmospheric escape would prevent the lower gravity L9859b from retaining an H_2-dominated atmosphere, though higher mean molecular weight atmospheres such as runaway greenhouse gas (CO_2 dominated, like Venus) or steam (H_2O-dominated) atmospheres are possible (Kopparapu et al., 2013). Further, Pidhorodetska et al. (2021) predicted a typical L9859b water feature in a clear, steam atmosphere would be detectable in as few as 3 transits.

In this chapter I analyze 4 transits of L98b (R=0.8 $R_\oplus$, M$\lesssim$1 $M_\oplus$, T=600 K) and 1 of L98c (R=1.35 $R_\oplus$, M=2.4 $M_\oplus$, T=515 K). I emphasize data and light curve analysis, which must be reliable in order for scientific interpretation to be at all meaningful. For every transit and eclipse analysis in this dissertation, I use a version of my custom HST WFC3 data and light curve analysis pipeline, nicknamed `DEFLATE` (**D**ata **E**xtraction and **F**lexible **L**ight curve **A**nalysis for **T**ransits and **E**clipses)[1]. It originated as sparse IDL programs, which I converted to Python2 programs, which I converted to Python3 programs, and finally developed on Github to be a public, open source analysis package. I fully explain it in this section since the small signal sizes leave no margin for error, especially for single-visit transits like L9859c. A

[1]https://github.com/AstroSheppard/WFC3-analysis

majority of my research was dedicated not only to performing these analyses, but also verifying their outputs to determine how sensitive results can be to modeling — and even data reduction — assumptions. Such detail is especially necessary for single-visit observations.

I also perform an exploratory analysis on the resulting transit spectra. Using the Bayesian atmospheric retrieval tool (PLanetary Atmospheric Transmission for Observer Noobs; PLATON; Zhang et al., 2019), I use the Bayesian evidence to determine the significance of any atmospheric detection and the characteristics of that atmosphere. I estimate impact and prevalence of stellar variability to explore potential contamination.

Section 2.2 describes the HST WFC3 observations. Section 2.3 describes my custom, publicly available pipeline that I used to preprocess the HST WFC3 data. Section 2.4 describes the parametric marginalization light curve analysis component of this pipeline and derives the transmission spectrum for each planet, and provides a series of diagnostics to ensure the results are valid. I explore the importance of several modeling assumptions, prior volumes, and systematic parameterizations. In Section 2.5, I perform an exploratory analysis on the resulting spectra, including the potential impact of stellar activity. In Section 2.5.2, I discuss the scientific interpretations and suggest future work. Appendix A provides supplemental figures.

2.2 Observations

HST WFC3 was not initially designed for the long exposures of bright sources necessary for observing exoplanet light curves. Therefore, its default "stare" mode is only able to find meaningful information for planets with large enough transit signals that small exposure times are adequate to make out a signal. Otherwise, the detector saturates too easily. First suggested by McCullough & MacKenty (2012) and applied by Deming et al. (2013), spatial scan mode is an observing technique which allows for longer exposure times (and thus SNRs) by spreading out photons on the detector spatially — into a 2-D spectrum — then combining them into a 1-D spectrum during data processing. Observations are either done bidirectionally, taking in light while alternating scan direction (typically denoted forward and reverse), or unidirectionally (only taking in light in one direction, then resetting to the starting point). For transit spectroscopy, this process is guided by fine guidance sensor (FGS) control (in addition to gyroscopes), which minimizes wavelength drift on the detector throughout a scan. The downside of this technique is that it complicates data analysis, but this is a small price to pay for a much greater SNR. It is also important to note that since HST orbits the Earth with a period of 95.47 minutes, it can only continuously monitor in $\sim$40 minute intervals, leading to gaps in the light curve between these observed "orbits".

The observations were taken in the IR channel of WFC3 with the first order of the G141 grism (R=130, dispersion=4.65nm pixel^{-1}) for low-resolution slitless spectroscopy in the 1.1–1.7μm wavelength range. Since the observations are all from the same proposal (Program GO-15856; PI Barclay), the observing and spatial scan parameters are extremely similar between visits. The data set comprises four transits of L9859b and a single transit of L9859c for a total of five visits. Each visit consists of four HST orbits of alternating forward and reverse scans for ~ 100 exposures per visit. Each scan is 69.6 seconds long and produces one exposure, and each exposure contains 4 non-destructive reads. The scan rate on the sky is 0.496 arsec/s ($\sim$4 pixels/s on the detector), and this rate allows for fine-guidance-censor (FGS) control. In total, the scan results in a 2-D spectra spread out among 290 rows of pixels per exposure, with a median signal of about 2.7e4 electrons per pixel. Additionally, WFC3 observed the system for a single exposure per visit with the F130N photometric band for wavelength calibration purposes. Table 2.1 summarizes the observation and spatial scan parameters. It is worth emphasizing that L9859b Visit 01 and the single L9859c visit occurred hours apart on the same day.

The top image of Figure 2.2 shows an example exposure of L9859b, Visit 02. The x-axis is the wavelength axis, and the spatial scan is parallel to the y-axis. Brighter pixels are indicative of greater flux. The straight line to the left of the 2-D spectra "rectangle" is the zeroth order spectrum.

2.3 Data Preprocessing and Analysis

Before physical properties can be inferred from the observations, the data must be processed to ensure instrumental relics are minimized and that the incoming photons are from the transiting system. This "preprocessing" is ubiquitous in

Table 2.1: L9859 Transit Observation Details

Observation Detail	L9859b				L9859c
	Visit 00	Visit 01	Visit 02	Visit 03	Visit 00
Date	02/09/20	04/07/20	09/28/20	11/25/20	04/07/20
Time of First Scan (MJD)	58888.4033	58946.8385	59120.3195	59178.9002	58946.0459
Number of Exposures	103	104	104	102	104
	Common to All Visits[a]				
Program ID and PI	15856, Barclay				
Scan Type	Bidirectional				
Detector Subarray Size	512x512				
Reads per Exposure	5				
Number of HST Orbits	4				
Expsoure Time	69.6 seconds				
Scan Rate	0.496 arcsec s^{-1}				
Scan Length	290 rows (34.5 arcsec)				
Signal Level	2.7e4 electrons per pixel				

[b] Given that these observations are all of the same star, many of the observation parameters are the same for each visit.

telescope data analyses and is often handled by an automated pipeline. For WFC3 G141 transit observations, the `cal-wfc3` pipeline automates several calibrations. However, grism observations done in spatial scan mode require custom reductions beyond those done by the automated calibration pipeline, which was designed assuming stare-mode observations. I describe these custom analysis routines in this section.

Figure 2.1 illustrates the data analysis process. While the `cal-wfc3` end-product *flt.fits* is not useful for spatial scan observations, several of the calibration steps are unaffected and can still be validly performed by the pipeline. The `cal-wfc3` pipeline is described in detail in the WFC3 Data Handbook[2] (Gennaro & et al., 2018). I summarize its contribution to my analysis below.

First, the pipeline initializes a data-quality array of known bad pixels (e.g, hot,

[2] https://hst-docs.stsci.edu/wfc3dhb

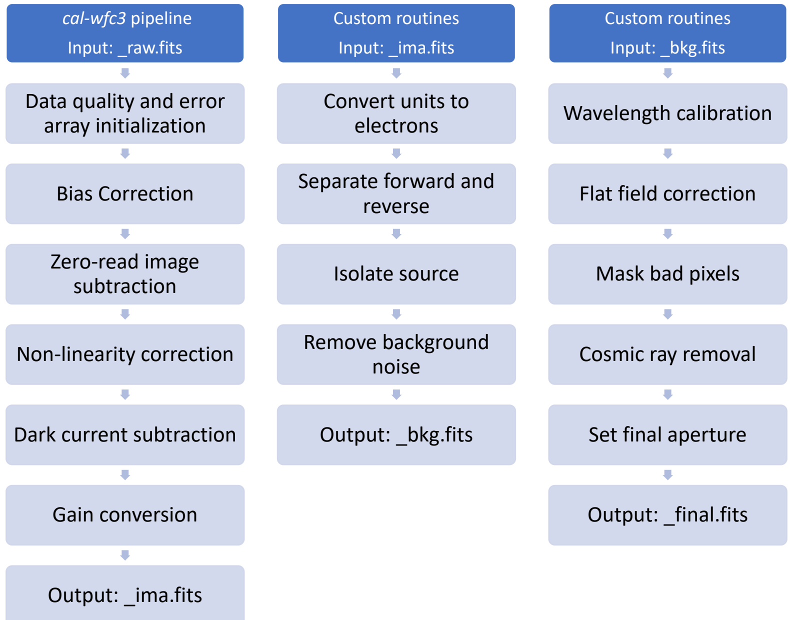


Figure 2.1: Data Preprocessing Flow Chart: Custom data analysis is necessary for spatial scan data, and is shown in columns 2 and 3.

unstable, or saturated pixels). These are identical across all exposures in an observation. It also provides a flux error array for each exposure that accounts for Poisson noise, read noise, and propagates uncertainty from the remaining reduction procedures. Next, it removes variable bias level in an exposure using non-photosensitive reference pixels on edge of detectors. The pipeline then performs a zero-read correction: The process of resetting pixels between exposures takes a non-zero amount of time (about 3 seconds), and the signal accumulated during this time (before the exposure "officially" starts observing) is subtracted from the exposure. `cal-wfc3` corrects for the known non-linear response of the detector, and removes the instrument's (150 K) thermal radiation-caused dark current. Finally, the pipeline applies a gain conversion of about 2.3 electrons/DN to convert recorded counts to electrons. Notably, it doesn't apply a flat-field correction, as the specific flat-field value at each pixel is wavelength-dependent, and `cal-wfc3` does not perform wavelength calibration.

This process results in a science (flux or electron) frame, flux error frame, and data quality frame for each readout in the exposure. This collection of 2-D frames constitutes the *ima.fits* outputs of the pipeline. These *ima* files, available via the HST Mast archive[3], are the beginning of my custom data analysis.

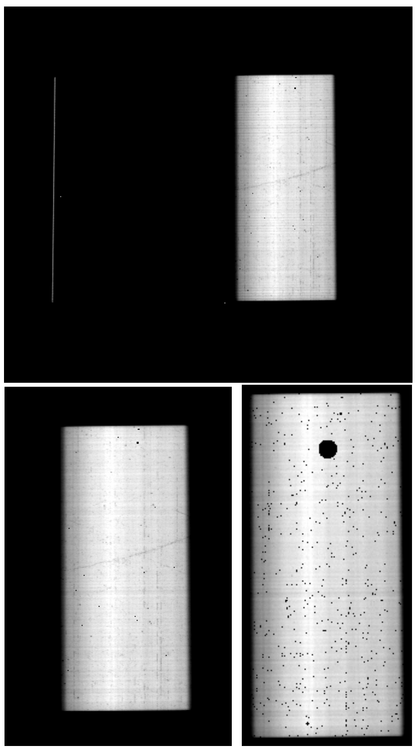


Figure 2.2: The fits file image of an example exposure at various data processing stages. The grism disperses light along the x-axis (wavelength axis), and spatial scan mode spreads out light along the y-axis (spatial axis). Brighter white areas indicate greater flux. **Top:** The raw `ima` fits file, which includes the first-order spectrum (rectangle) as well as the zeroth order spectrum (line). **Bottom left:** Exposure after background removal and initial aperture setting. **Bottom right:** Completely processed exposure. The black dots (including the large blob) are known bad pixels and are given zero weight in light curve analyses.

The custom data analysis described in this chapter is an expanded version of that described in Sheppard et al. (2021), nicknamed `DEFLATE` (**D**ata **E**xtraction and **F**lexible **L**ight curve **A**nalysis for **T**ransits and **E**clipses). The overarching process is illustrated in Figure 2.1. I first download the *ima.fits* files and separate the forward and reverse scans for independent processing, since the spatial scans tend to be

[3]https://archive.stsci.edu/hst/

offset in the spatial direction by several rows, unnecessarily complicating aperture determination. I then isolate the first-order spectra for each direction with a rough user-defined box, removing the zeroth-order spectrum and other sources. Then, if necessary, I convert the units for each pixel from electrons/s to total electrons recorded. These basic reductions set the stage for the more complicated processing steps. There are several valid approaches to each, but reasonable approaches general yield consistent results. I step through each process here, as well as general issues that may arise.

Background removal: `DEFLATE` removes the background noise in each exposure using the "difference reads" method (Deming et al., 2013). While it is possible to use a scaled version of a master sky background file to remove specific background patterns (Gennaro & et al., 2018), this method takes advantage of the multiple readouts within each exposure to remove background in a purely data-defined way. As the instrument scans, it records (reads out) the flux at several pre-determined intervals, and these "mini-exposures" are called reads. The L9859 system uses 5 reads, which appear as increasingly long rectangles, over a roughly 70 second exposure time. While the rows in each read collect flux from the source for a fraction of the scan ($\sim$14 seconds), it collects background the entire scan. By subtracting consecutive reads, I isolate photons observed in that particular 14 second interval and completely remove any photons observed in the non-source region of the observation. The steps are as follows:

- Subtract read n-1 from read n to create a difference frame

- Find the maximum flux via the median of the 5 rows with the greatest flux (to avoid potential cosmic ray issues).
- Find the centroid of the difference frame, and conservatively define the source as all consecutive rows with at least 1% of the maximum flux.
- Mask the source, and find the median of all other pixels.
- Subtract the median from the entire image.
- Set all pixels outside the mask to be zero.
- After done for all reads, sum the difference frames to create a background-removed science exposure.

As a final step, I propagate the uncertainty due to this background subtraction by adding it in quadrature, since the new count for each pixel is $F_{new} = F_{old} - F_{bkg}$. The difference-reads method lowers the likelihood of cosmic rays impacting the data (since the location of the source on the detector has no bearing on cosmic rays, any ray that hits a non-source pixel during the observation is automatically zeroed out). It also allows for resolving the source from companions or other field sources in the case of overlapping scans — as is the case for HAT-P-41b in Chapter 3.5.2 — since the individual difference frames do not overlap. `DEFLATE` saves the products of this method as *bkg.fits* files and separates the files by spatial scan direction. An example *bkg.fits* file is shown in the bottom left panel of Figure 4.2.

The next set of analyses convert the *bkg.fits* file to the final.fits files used in light curve analyses. For bi-directional scans, each step is performed on the forward and reverse scans independently. First, `DEFLATE` converts the detector pixels to wavelength.

Wavelength calibration: Due to distortions, the pixel-to-wavelength calibration (i.e, wavelength solution) depends on the exact X and Y position on the detector, and so it varies between observations. Still, it is a roughly linear conversion that follows the following set of equations (Wakeford et al., 2013):

$$\lambda_{(X_{\text{ref}},Y_{\text{ref}})} = \lambda_{\text{ref}} = a_0 + a_1 * X_{\text{ref}} \tag{2.1}$$

$$\lambda_{\text{pixel}} = \lambda_{\text{ref}} + Y_{\text{dispersion}} * (X_{\text{pixel}} - X_{\text{shift}}) \tag{2.2}$$

The reference coordinates $(X_{\text{ref}}, Y_{\text{ref}})$ are determined by the photometric images taken at the beginning of each visit. Coefficients for converting this reference pixel to a reference wavelength (a_0, a_1) were determined empirically by Kuntschner et al. (2009, Table 5). The wavelength of light recorded by a particular pixel is dictated by the dispersion for the Y-coordinate of the reference pixel ($Y_{\text{dispersion}}$) and the intrinsic offset (X_{shift}, in pixels) between the location of the filter image and the grism-dispersed light. $Y_{\text{dispersion}}$ and X_{shift} are constrained, but spatial scan mode complicates those values. Consequently, `DEFLATE` follows best practice and fits for these values by comparing an observed out-of-transit spectrum (thus, a stellar spectrum) to a stellar model. To mimic how the instrument would detect the star, the model spectrum is the product of an ATLAS stellar model (Castelli & Kurucz, 2004) and the G141 grism sensitivity curve.

The ATLAS model spectra cover a wide range of temperature (3500–50000 K), metallicity (-5.0–+1.0 [Fe/H]), and gravity (logg 0.0–5.0) with increments of 250 K, 0.1, and 0.5, respectively. A perfect fit is not necessary here: the wavelength solution necessary for the model to approximate low-resolution WFC3 spectrum is guided primarily by broad stellar feautres (e.g, the 1.28μm Paschen beta line) and the steep edges of the G141 sensitivity curve. As such, I select the closest stellar model to that of the star. For the L9859 system, this is T_{eff}=3500 K, logg=5.0, [Fe/H]=-0.5 (Table 2.2). This outputs three components: a wavelength grid, a line-opacity flux at each point in the grid, and a continuum-opacity flux at each point in the grid. For flexibility, and due to the non-exactness of stellar properties compared to the model, `DEFLATE` combines the line and continuum fluxes as ($\alpha\times$Line + Continuum), essentially allowing the strength of the stellar lines to vary, though typically I fix α to 1.

After converting to the appropriate units, the flux of the model is $\mathrm{G141}_{\mathrm{sensitivity}} \times (\alpha \times \mathrm{F}_{\mathrm{lines}} + \mathrm{F}_{\mathrm{continuum}})$ at each wavelength λ_{model}. The data are the total flux at each pixel column for pixel numbers 0, 1, ..., N and converted to wavelength via equations 2.1 and 2.2. This conversion is dependent on parameters $\mathrm{Y}_{\mathrm{dispersion}}$ and $\mathrm{X}_{\mathrm{shift}}$. For each fit iteration, `DEFLATE` uses those parameters to calculate the wavelength for each pixel, then interpolates the (high-resolution) model to determine a model flux at each pixel's wavelength, and it minimizes this model-data flux difference. `DEFLATE` uses a weighted least-squares estimator (KMPFIT; see Section 2.4) to fit the model to the data, allowing parameters $\mathrm{Y}_{\mathrm{dispersion}}$, $\mathrm{X}_{\mathrm{shift}}$, and optionally α to vary. I set a uniform prior on $\mathrm{Y}_{\mathrm{dispersion}}$ to contain the reasonable possible values based on the $Y_{\rm ref}$–$Y_{\rm dispersion}$ relationship determined by Figure 6 of Kuntschner et al. (2009).

Figure 2.3 shows the result of one such fit. The L9859 system is a rare case where fixing $\alpha = 1$ results in a relatively poor fit (top left subfigure). However, allowing the strength of the line flux to vary results in an excellent fit ($\chi^2_{red} \sim$ 1; top right) and, notably, the same wavelength solution. The line-strength may differ due to the mismatch between the exact T_{eff}, logg, and metallicity, and that of the model, or it may be due to opacity modeling choices. I emphasize that this is not common for my wavelength calibrations by showing the calibration for a different system (WASP-79, T$\sim$ 6500 K) in the bottom panel.

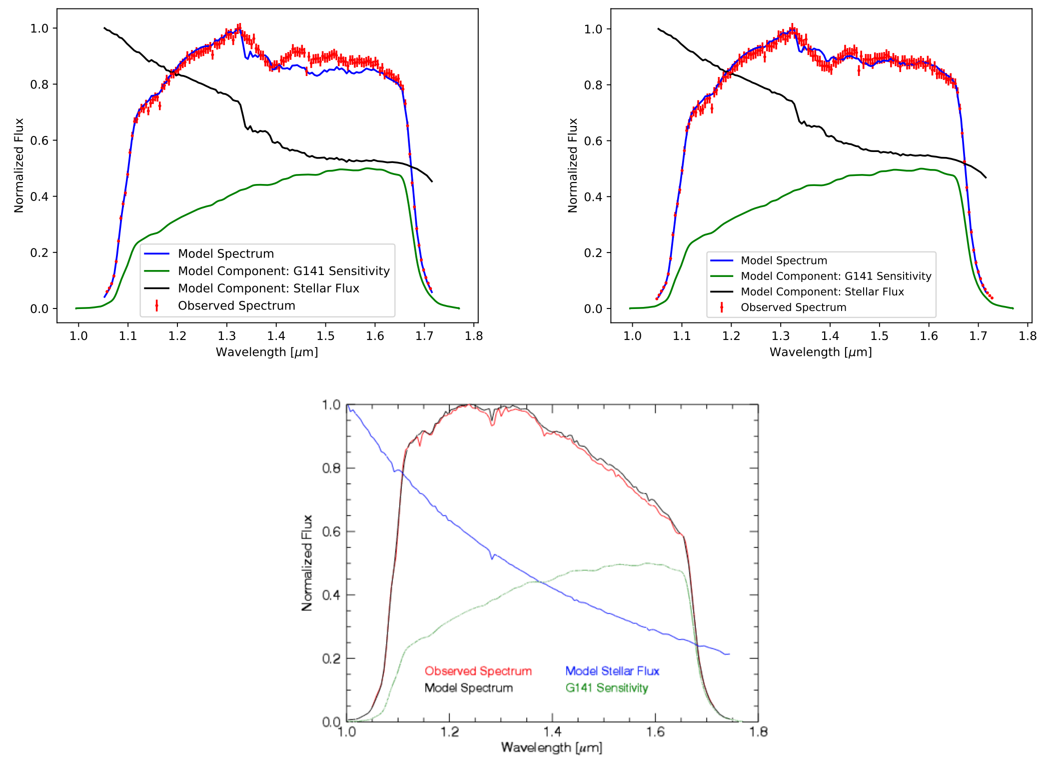


Figure 2.3: Example wavelength calibration for three scenarios. **Top Left:** L9859 with fixed stellar model line-flux strength. **Top Right:** L9859 with scaled line-flux contribution. Though a better fit, the wavelength solution is nearly identical to the top-left case. **Bottom:** Same process (fixed line flux) for the simpler spectrum of the hotter WASP-79.

I determine the wavelength solution for each of the five observation visits, using both the forward and reverse light curves to inform the calibration. The forward and reverse scans tend to be offset vertically from one another by a small amount, and the difference in wavelength solution is never more than 3% and typically around 1.5%. Practically this is negligible. It is around 2–3 Å (for bin sizes of roughly ~ 300 Å) and well within the size of a pixel (i.e, subpixel shift). For low-resolution spectroscopy, this is sufficient: the most important features in this wave band are very broad combinations of millions of lines, insensitive to differences of a few angstroms. I note that the wavelength solution is not significantly impacted by exact stellar model choice, or error scaling, or line-strength scaling.

Flat-field: The wavelength provides a scaling factor for the flat of each pixel (i.e, intra-pixel sensitivity). `DEFLATE` combines this factor with the downloadable flats files to divide out the flat-field from both the data and the error array (to propagate uncertainty). This flat-field removal is responsible for the disappearance of the "streak" from the bottom left (*bkg.fits*) panel to the bottom right (*final.fits*) panel of Figure 4.2.

Bad pixels: `DEFLATE` handles the `cal-wfc3`-flagged "bad" pixels, which are identical across all exposures, by giving them zero-weight. It accomplishes this by converting the data to masked `numpy` arrays and setting a mask for any flag value greater than zero. The exception is flag 2048, which indicates that a pixel took in flux during the zero-read frame. This is already accounted for in the pipeline reduction, so these pixels are not unreliable. It is possible to interpolate flux values at these pixels, but I prefer the zero-weight method since it requires fewer assumptions. The zero-weight pixels make up roughly 2% of all pixels in an exposure and appear as black dots (and a blob) in the bottom right image of Figure 4.2.

Cosmic ray removal: `DEFLATE` uses a corrected median time filter to flag cosmic rays. Each pixel is compared to itself in each exposure of the observation, and anomalously bright pixels are flagged as cosmic rays and set to the median value of that pixel over time. Before applying the filter, `DEFLATE` corrects for three known processes that lead to flux variations for a given pixel: time-dependent instrumental effects (see Section 2.4), the transit/eclipse itself, and inconsistent spatial scan rates.

To the first order, each pixel is impacted similarly by time-dependent instru-

Table 2.2: L9859 Stellar and Orbital Properties

Type	Parameter	L9859	
Stellar	Radius [$R_\odot$]	0.314 ± 0.009	
	Mass [$M_\odot$]	0.293 ± 0.02	
	$T_{\rm eff}$ [K]	3429 ± 157	
	$\log g_s$ [cgs units]	4.91 ± 0.004	
	[Fe/H]	-0.5 ± 0.5	
	Distance [pc]	10.619 ± 0.003	
	$\log R'_{HK}$	$-5.40 \pm .011$	
	$P_{\rm Rot}$ [days]	78 ± 13	
Type	Parameter	L9859b	L9859c
Orbital	R_p/R_s	0.0234 ± 0.0009	0.0396 ± 0.0009
	a/R_s	$16.2^{+0.8}_{-1.0}$	$22.5^{+1.1}_{-1.4}$
	i [Degrees]	88.7 ± 0.8	89.3 ± 0.5
	T_c [BJD-2457000]	1366.1701 ± 0.0007	1367.2755 ± 0.0004
	$\log g_p$ [cgs units]	< 16.1	13.0 ± 2.3
	P [days]	$2.253 \pm 2 \times 10^{-5}$	$3.691 \pm 1.4 \times 10^{-5}$
Planet	R_p [$R_\oplus$]	0.80 ± 0.05	1.35 ± 0.07
	M_p [$M_\oplus$]	< 1.01	2.42 ± 0.35
	$T_{{\rm eq},A=0}$ [K]	610 ± 15	520 ± 15
	a [AU]	0.0233 ± 0.0017	0.0324 ± 0.0024

NOTE — Stellar spectral values from TIC-v8 (Stassun, 2019). Planet mass and stellar activity parameters from RV paper (Cloutier et al., 2019). Orbital and remaining planet parameters from TESS transit discovery paper (Kostov et al., 2019).

mental effects and transits, so this is easily corrected by normalizing each pixels by the median of its row, column, or exposure. I choose to normalize by row due to potential inconsistent scan rates. This refers to the tendency of the WFC3 to occasionally experience "hitches" where it lingers on a given spatial coordinate for slightly too long, then quickly scans through the next few rows to "catch up" to where it should be. A non-uniform scan rate has no effect of the observed spectrum — the same number of electrons are detected for a given column regardless. Still, the vertical distribution of electrons in a given column would be impacted, leading to certain rows appearing as very bright (and adjacent rows appearing as unusually dim). Thus, dividing each pixel by the median of its row prevents `DEFLATE` from flagging entire rows as cosmic rays.

`DEFLATE` uses a double-sigma cut: first it applies an 8σ cut and corrects any extreme outliers, then, in case that an extremely bright cosmic ray was distorting the standard deviation, it applies a second 5σ cut to correct the remaining energetic particles. Less than 0.5% of all pixels in L9859b's and L9859c's observations are impacted by cosmic rays, and typically only a few pixels per exposure are impacted.

Aperture: `DEFLATE` follows a simple procedure to define a light curve extraction aperture. It first defines the maximum flux of an exposure as the median of the five rows with the greatest flux. The edge of the box is set to the outermost row and column with a median value of greater than 3% of the maximum flux. This relatively low cut-off captures the entire first-order spectrum and minimizes the impact of vertical shifts. This method maximizes the photons observed from the source and avoids over-processing the data. After determining the science aperture, `DEFLATE` saves the data as *final.fits* files. An example *final.fits* file for L9859b is shown in the bottom right panel of Figure 4.2.

The end result of my `DEFLATE` data analysis is an isolated 2-d spectrum of the planetary system that can easily be collapsed in the spatial direction (via summing over columns) to form a 1-d spectrum of flux — with a well-defined error — at each exposure. In other words, the end product is a time series of flux (light curve) at each wavelength. From these time series, I derive a transit (or eclipse) depth at each wavelength to form a spectrum, which is then compared to chemical and physical models to give insight into the planet's atmosphere.

2.4 Light Curve Analysis

2.4.1 Modeling the Light Curves

The processed data are a time series of 2-D spectra, which is the integrated flux as seen through the instruments lens. This time series can be interpreted as the "true" transit light curve muddled by systematic effects due to the instrument. To extract the transit depth, I need to decouple the instrumental effects to approximate the "true" transit event.

Modeling a transit light curve has two major components: modeling the physical transit, and modeling the non-astrophysical instrumental effects related to how the solid state CCD detector collects flux, i.e. the systematics. WFC3 observations commonly exhibit several systematic effects. The most prominent are a hook/ramp feature due to charge-trapping, a visit-long decrease in flux, a "breathing" effect based on changing temperatures during HST's orbit, and a wavelength jitter effect (e.g, Berta et al., 2012; Wakeford et al., 2016a; Zhou et al., 2017; Tsiaras et al., 2018, among many others). These features vary in magnitude between different observations in non-obvious ways. Solid-state physics is complicated, and there is no encapsulating physically-motivated model to describe all of these effects (though recently individual features have been modeled more successfully, e.g. Zhou et al. (2017)). Instead of using inherent properties of the detector, these features are typically removed using empirical methods (Gibson, 2014a; Nikolov et al., 2014; Haynes et al., 2015). In this chapter — and for all light curve analyses in this dissertation — I use a novel version of parametric marginalization to derive transit parameters from WFC3 observations.

Here I explain the "parametric" in parametric marginalization. Instrumental effects are commonly parameterized by easy-to-calculate, auxiliary properties as proxies for the underlying physics. Useful properties were determined empirically from trial and error from many HST WFC3 observations (e.g, Nikolov et al., 2014). For example, there is a clear decrease in flux over the course of a single observation. This is clear in the top panel of both white light curves in Figure 2.4: the average flux in the last orbit is less than the average flux in the first orbit. There is no clear-cut way to predict this decrease from solid-state physics; however, there is clear correlation between the time of an exposure and its flux. I convert observation time to the planet's orbital phase for computational convenience, and I can parameterize this slope by the phase of the planet's orbit (θ) to account for this systematic effect.

The same logic is applied to other instrumental effects to determine relevant auxiliary parameters: both the “hook” effect and the HST breathing effect correlate to HST orbital phase (ϕ), and wavelength-jitter complications correlate to the horizontal shift of each exposure (as determined by the cross correlation between exposure i and the first exposure; δ). These three auxiliary parameters account for every potential instrumental effect previously observed in WFC3 light curves. It is intuitive and efficient to approximate the form of these parameters as simple polynomial expansions (Gibson, 2014b). However, since these are empirical methods, the order of polynomial to use for each auxiliary parameter is not known *a priori*. Further, since different datasets are not impacted by systematics in a consistent way, that polynomial order may not be consistent between observations. Marginalization addresses these issues.

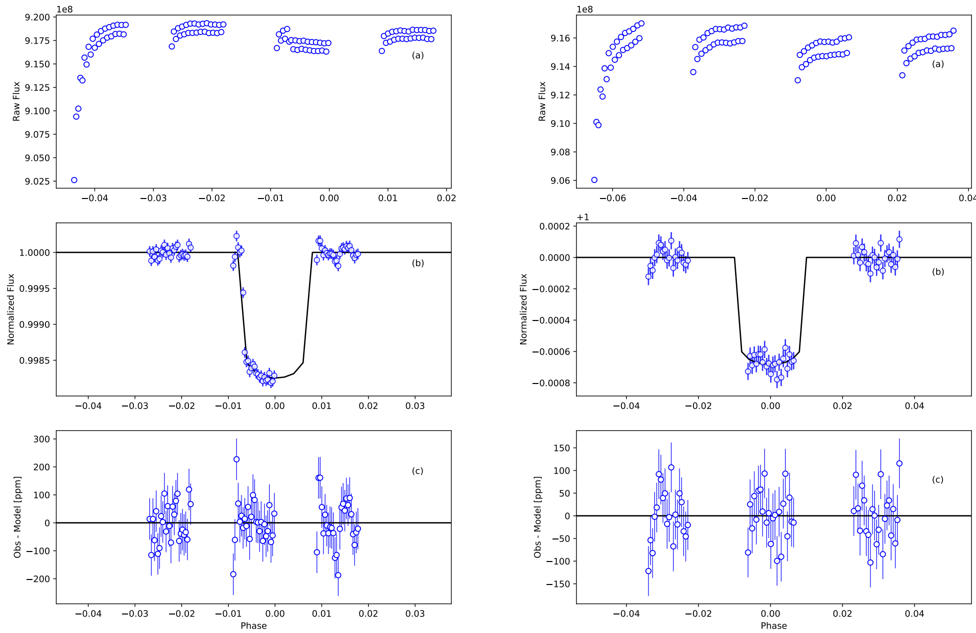


Figure 2.4: Visualization of white light curve fit for the highest weighted systematic model for L9859c and L9859b visit 03. Panel (a) shows the band-integrated light curve. Panel (b) shows the de-trended light curve as well as the best fitting transit model. Note that this is illustrative — the instrumental effects and transit model parameters are fit for simultaneously. Panel (c) shows the residuals between the data and the best-fitting model. Note how the systematic effects, such as the severity of the orbit-long exponential ramp, differ between the two observations.

Parametric marginalization is a form of Bayesian model averaging, conceptually introduced to exoplanet light curves by Gibson (2014b) and first applied to WFC3 transit spectroscopy by Wakeford et al. (2016a). It first defines a *feature space*, which is a grid of systematic models to be tested. Then, instead of selecting the "best" systematic model, it assigns an evidence-based weight to each and then marginalizes over the systematic models (i.e., takes a weighted average). In the extreme case that every model is equally likely, then every model has an equal weight the the marginalized parameters are just averages. The other extreme case — where one model in an excellent fit and the rest are terrible — would give the good model ∼100% of the weight, reducing to model selection.

Marginalization has a few advantages over model selection. First, it is applicable to many different data sets, which will allow for a comparison of planetary spectra without necessitating customization or that differences are due to different model choices. Second, it provides physical insight into what systematic effects are present in different data sets, which could lead to a better understanding of the prevalence and driving forces behind those effects. Third, it intrinsically accounts for the uncertainty in model selection and shows the sensitivity of results to the model choice. In normal model selection, all results are conditional on the model choice, which leads to overconfidence (Gibson, 2014b). The uncertainty in derived transit parameters is then based both on the uncertainty on that parameter *conditional* on a model *and* the scatter in parameter values *between* different models. In this way, it functions as a less flexible version of Gaussian processes (GP Gibson, 2014a) that utilizes physical insight to provide better constraints than GP and is more easily interpretable.

A necessary requirement for marginalization is that least one model is "correct", i.e. able to describe the systematic effects. It is important to have a flexible enough systematic grid to consistently meet that criteria while balancing computational expense. I find, similar to Wakeford et al. (2016a), that fourth order polynomials strike that balance. I use a grid of models include of to four powers of HST phase, four orders of wavelength shift, and 5 forms of a orbital phase-dependent visit-long slope (none, linear, quadratic, exponential, and log). Each higher power includes all lower powers (e.g, 3rd order HST phase is $a_0\times$HST + $a_1\times$HST^2 + $a_2\times$HST^3), and there are no cross terms. This results in a grid of 125 systematic models (5 possible HST powers × 5 possible shift powers × 5 possible slope parameterizations). There are an additional two parameters: separate normalization constants for the forward (A_f) and reverse scans (A_r). It is typical for the two directions to be offset, though that is the primary effect and they can still be fit simultaneously.

It is computationally difficult to fully sample the parameter space of all 125 models using Markov Chain Monte Carlo (MCMC) samplers, so I instead fit each model using `KMPFIT`[4] (Terlouw & Vogelaar, 2015), a Python implementation of the Levenberg-Markwardt least squares minimization algorithm, to more quickly determine parameter values and uncertainties. Wakeford et al. (2016a) found that uncertainties derived from these two methods typically agree within 10%. My own comparison of MCMC and KMPFIT fits also finds excellent agreement, and that KMPFIT (for a single model) tends to *overestimate* uncertainty relative to MCMC. I then weight each model by its Bayesian evidence — approximated by the Akaike information criterion (Akaike, 1974) — and marginalize over the model grid (assuming a prior that each model is equally likely) to derive the light curve parameters and uncertainties while inherently accounting for uncertainty in model choice.

[4] `https://github.com/kapteyn-astro/kapteyn/`

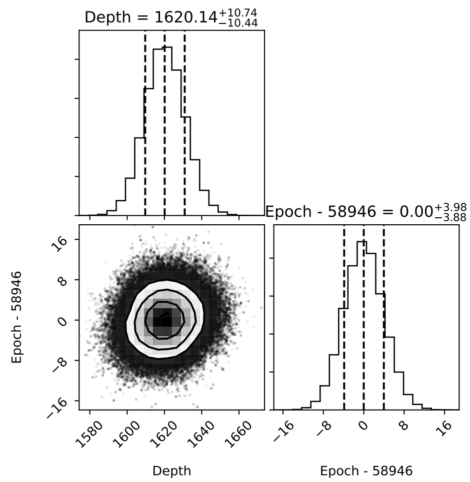


Figure 2.5: Corner plot of astrophysical parameters for MCMC fit of highest weighted systematic model. The model converged and derived uncertainties slightly less than KMPFIT's.

`DEFLATE` uses `BATMAN`[5] transit models (Kreidberg, 2015) for the physical component of the light curve model. The only orbital parameters that are potentially constrainable from HST WFC3 data are transit depth (R_p/R_s), center of transit time (T_0), and occasionally orbital density (a/R_s), inclination (i), and a linear limb-darkening coefficient (c_0). I typically assume nonlinear limb darkening (LD) and derive the coefficients by interpolating the 3-D values from Magic et al. (2015) to the central wavelength of WFC3 (1.4 μm). These coefficients are fixed for light curve fitting, as HST's poor phase coverage could not possibly constrain the shape of transit well enough to converge on four coefficients. Magic et al. (2015) only provides coefficients for stars hotter than 4000 K, so for cooler M-dwarfs (like L9859) I instead use values from either Claret & Bloemen (2011) or Claret et al. (2012). The model from Claret & Bloemen (2011) uses ATLAS stellar models (Castelli & Kurucz, 2004) and spans different stellar [Fe/H] values, while Claret et al. (2012) uses PHOENIX stellar models (Baron et al., 2010) with more up-to-date opacities (but no metallic-

[5]`https://github.com/lkreidberg/batman`

ity flexibility). I default to the PHOENIX models, but I check to make sure results are not sensitive to LD source. For Earth-like planets with smaller transit signals I also test a linear LD law with the coefficient being a fittable parameter.

The most complex possible model is:

$$\text{Depth} = T(R_p/R_s, T_0, a/R_s, i, c_0) \times S(A_f, A_r, f(\phi), f(\theta), f(\delta)) \tag{2.3}$$

Here, T() is the BATMAN orbital model and S() is the systematic model.

With the models and fitting methods defined, I briefly summarize the spectral derivation process. First, I fit the white light curves. This provides a sanity check on the data, maximizes photons for deriving wavelength-independent properties such as inclination and a/R_s, and captures the structure of residuals for each systematic model, if present. Determining the residuals allows for further de-trending of spectral curves via white light residual removal (Mandell et al., 2013; Haynes et al., 2015). The shape of the residuals are assumed to be constant with wavelength, though the amplitude is allowed to vary. This allows for removal of any wavelength-independent red noise from spectral bin curves at the penalty of slightly increasing the white noise. Note that the band-integrated uncertainty is sufficiently small relative to spectral light curve uncertainty such that the added noise has only a minor effect. The spectral light curve fits are extremely similar to the white light fits. The only differences are the incorporation of residuals, re-calculating LD coefficients and wavelength shift for each bin, and always fixing T_0, a/R_s, and i to their white light values.

Table 2.3: Derived transit depth and time for every L9859 broadband light curve.

Observation	Transit Depth [ppm][a]	T_c [BJD-245700]
L9859b Visit 00[b]	643 ± 17	1888.8920 ± 0.0001
L9859b Visit 01	655 ± 23	1947.4712 ± 0.0002
L9859b Visit 02	658 ± 16	2120.9607 ± 0.0004
L9859b Visit 03	637 ± 16	2179.546 ± 0.004
L9859c	1620 ± 24	1946.7068 ± 0.0001

[a] $(R_p/R_s)^2$

[b] Variance-weighted average of all four L9859b visits is $648 \pm$ 20 ppm

Since this is empirical, it's important that no transit depth/spectral feature is sensitive to a loosely supported assumption. Therefore, `DEFLATE` is highly customizable, allowing for many changes to test if spectrum or depth significantly change based on certain assumptions. For example, including residuals in the spectral derivation is optional, allowing the user to easily check if residuals significantly improve fits, or if they overfit the data.

2.4.2 White light Results

I fix the orbital period, inclination, and a/R_s to their discovery paper values (Table 2.2) in the light curve analyses of each visit, which ensures consistent orbital parameters are used across different datasets (Kostov et al., 2019). I also only allow linear visit-long slopes, since the typical L98 dataset only has three usable orbits covering a small amount of the out-of-transit baseline. As is common practice, I ignore the systematic-dominated first orbit in the white light analysis; however, the use of common-mode detrending provides the option of including that orbit in the spectral light curve analysis. Similarly, the planet b light curves are more

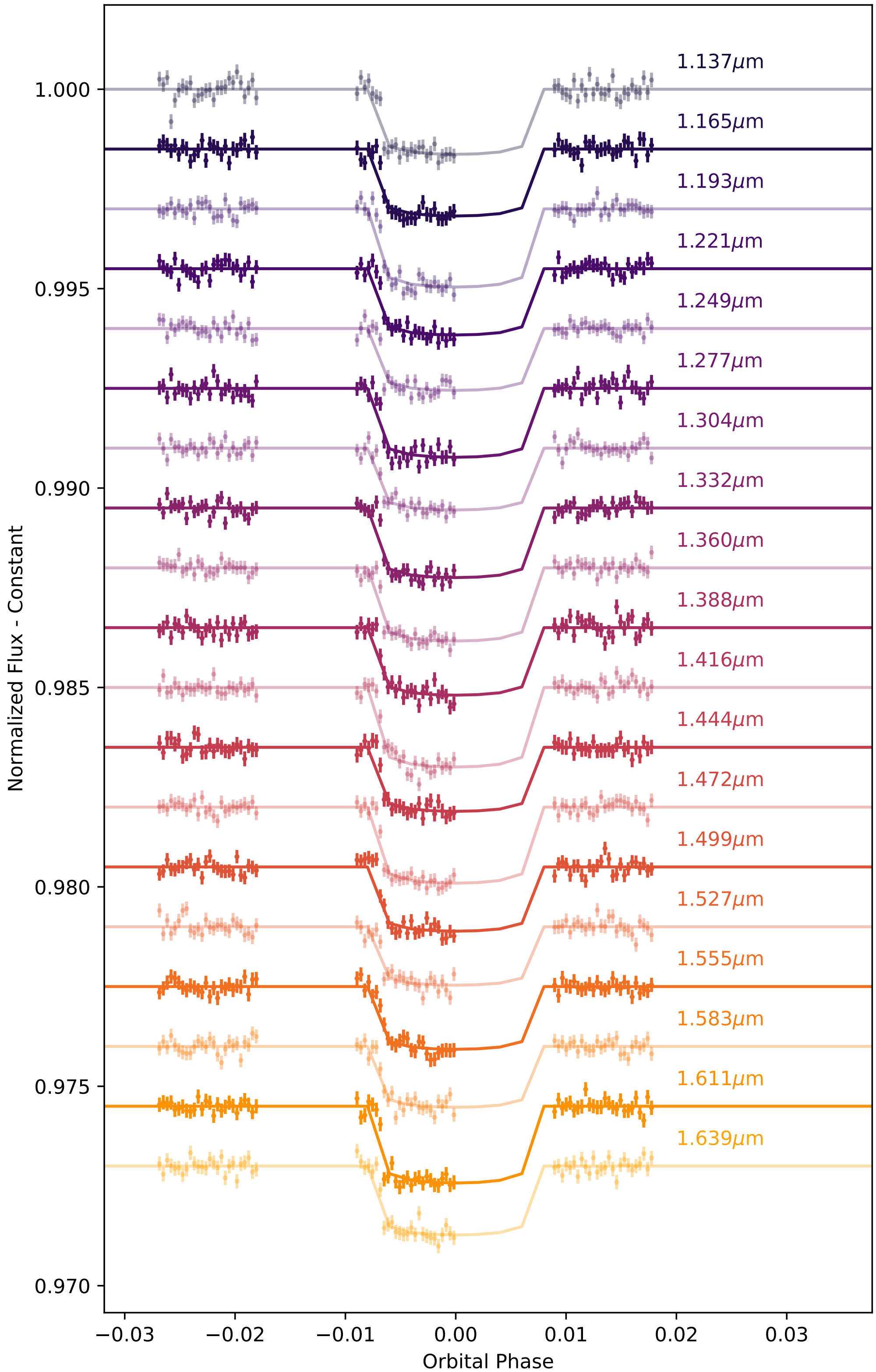


Figure 2.6: Spectral light curves for L9859c, visit 00.

severely impacted by the orbital ramp, and I follow common practice and ignore those data points for my planet b analysis. The raw light curve, the light curve with instrumental systematics removed, and the residuals from the highest-weight systematic model are shown in Figure 2.4. The derived transit depths and center-of-transit times (T_c) are given in Table 2.3. I derive the white light depth to be 1620 ± 24 ppm for L9859c and 648 ± 20 ppm for L9859b (the RMS between all four visits is 9 ppm, showing excellent agreement). The derived depths are insensitive to model assumptions, varying no more than 20 ppm if linear LD is fit for, or if a a/R_s is fit for, or if a quadratic visit-long slope is assumed. The reduced chi-squared of each fit is around 1.2, which is typical of HST white light curves.

To further validate these results, and to make sure the derived uncertainties are reasonable, I fit the highest-weighted systematic model of L9859c with MCMC (`emcee`; Foreman-Mackey et al., 2013). I also show the corner plot (Foreman-Mackey, 2016) of astrophysical parameters in Figure 2.5. Validation of convergence and a full corner plot are provided in the Appendix A. The posterior of the transit depth is Gaussian. The two methods are in excellent agreement, down to the ppm: both find a depth of 1620ppm and the MCMC uncertainty is 98% of the KMPFIT uncertainty (note that the marginalized white light uncertainty is greater due to accounting for systematic model uncertainty)

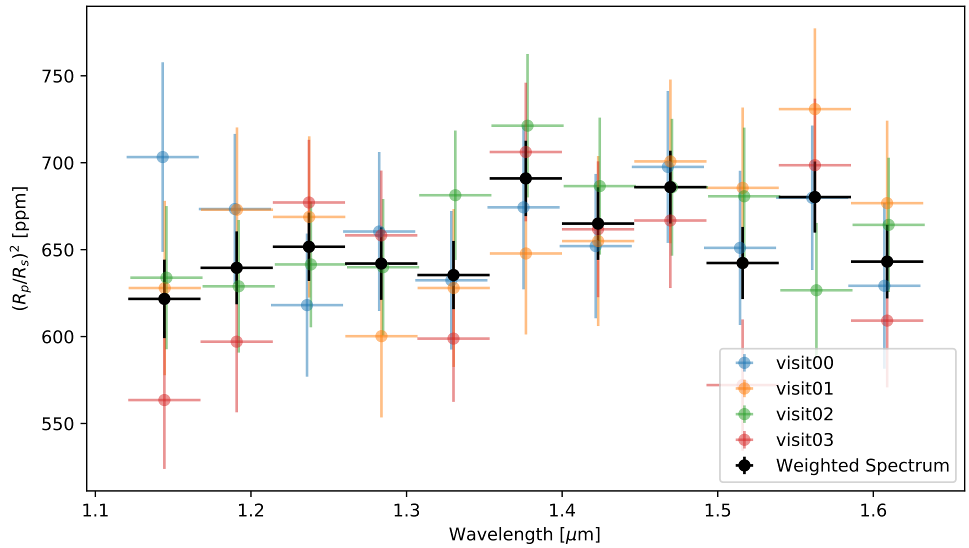


Figure 2.7: Derived transit spectra for each of the four observations of L9859b. The inverse variance-weighted is shown in black. The variability between spectra generally agrees with the uncertainties, with the exception of the bluest bin (1.14μm).

2.4.3 Transit Spectra Derivation

To derive the transit spectrum, I bin the 1D spectra from each exposure between the steep edges of the grism response curve (1.1–1.6 μm), deriving a flux time-series for each spectral bin. I test several bin widths since the long scan observations (close to 300 rows) are more at risk of wavelength blending (Tsiaras et al., 2016), which will effect larger bins less than smaller ones. I find no difference for L9859c (Figure 2.8), and choose 6 pixel (0.0279μm) bins to maximize resolution without drowning the signal in noise. For L9859b, the signal is smaller, so I choose larger 10 pixel bins (0.0464μm) to increase SNR. I derive the spectra for each visit of L9859b separately, then combine them using a variance-weighted average (Figure 2.7). The spectra are given in Table 2.4. The spectral light curves for L9859c

Table 2.4: Transmission Spectra of L9859 planets

Planet	λ [μm]	Depth [ppm]	Planet	λ [μm]	Depth [ppm]
L9859b[a]	1.123–1.169[b]	622 ± 23	L9859c[c]	1.123–1.151	1561 ± 56
	1.169–1.215	640 ± 21		1.151–1.179	1609 ± 55
	1.215–1.262	652 ± 20		1.179–1.207	1686 ± 54
	1.262–1.308	642 ± 21		1.207–1.235	1600 ± 52
	1.308–1.355	635 ± 20		1.235–1.263	1616 ± 51
	1.355–1.401	691 ± 22		1.263–1.291	1539 ± 59
	1.401–1.447	665 ± 21		1.291–1.318	1585 ± 49
	1.447–1.494	686 ± 21		1.318–1.346	1572 ± 51
	1.494–1.540	642 ± 21		1.346–1.374	1658 ± 53
	1.540–1.587	680 ± 20		1.374–1.402	1628 ± 56
	1.587–1.633	643 ± 21		1.402–1.430	1693 ± 56
				1.430–1.458	1697 ± 55
				1.458–1.485	1721 ± 53
				1.485–1.513	1665 ± 52
				1.513–1.541	1549 ± 54
				1.541–1.569	1662 ± 54
				1.569–1.597	1625 ± 54
				1.597–1.625	1739 ± 55
				1.625–1.653	1635 ± 54

[a] Variance-weighted average of four visits.
[b] Bin size = 0.0464μm, resolution ~ 30
[c] Bin size = 0.0279μm, resolution ~ 50

are shown in Figure 2.6. The same figure for L9859b are shown in the Appendix A.

2.4.3.1 WFC3 Transit Spectrum Verification

Marginalization is only reliable if at least one model is a good representation of the data (Gibson, 2014b; Wakeford et al., 2016a). I therefore checked the goodness-of-fit of the highest-weighted systematic model for each light curve using both reduced χ^2 and residual normality tests. Further, I explored if red noise is present in the light curve residuals, as that can bias inferred depth accuracy and precision (Cubillos et al., 2017).

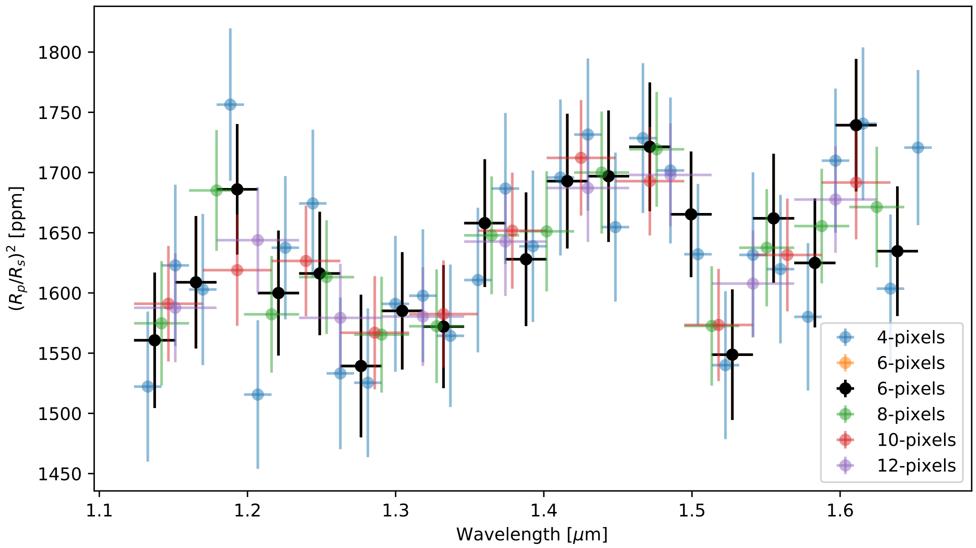


Figure 2.8: Marginalization-derived transit spectrum for L9859c at different resolutions. The shape of the spectrum is not sensitive to the spectral bin size.

Though χ^2 cannot prove that a model is correct, it can demonstrate that the fit of a particular model is consistent with that of the "true" model with "true" parameter values (Andrae et al., 2010). Therefore it is an informative goodness-of-fit diagnostic, and it is particularly useful due to its familiarity and simplicity. The "true" model with "true" parameters will have a reduced χ^2 of one with uncertainty defined by the χ^2 distribution. For both the band-integrated and spectral light curves (~ 60 degrees of freedom), this results in an acceptable reduced χ^2 range of roughly 0.66–1.4.

The band-integrated analysis ($\chi^2_\nu = 1.2$) and all spectral bins (median $\chi^2_\nu = 0.9$) fall within this range. The exception is the 1.499 μm light curve, which heighest-weighted model fit has a reduced χ^2 of 0.59. This low value indicates that the uncertainties in this light curve are overestimated. This is likely due to incorporating

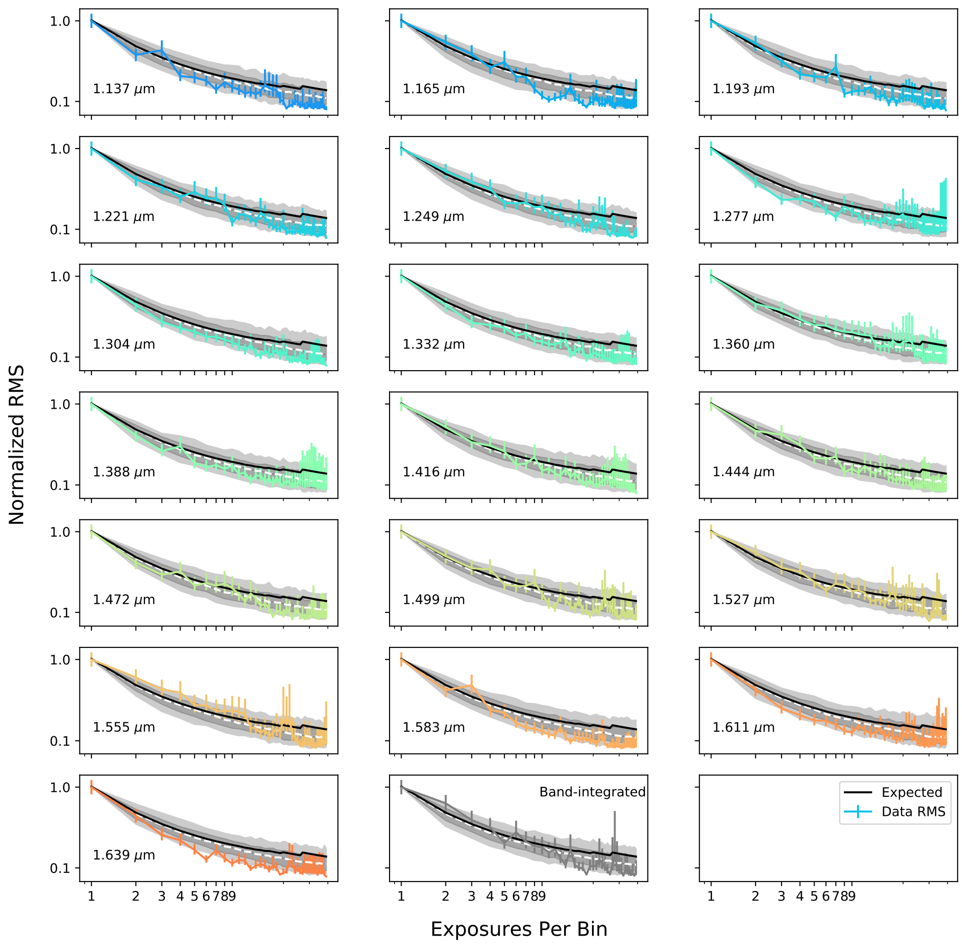


Figure 2.9: Correlated Noise Diagnostic Figure for L9859c. Bin RMS analysis for each spectral bin (see Section 2.4.3.1). The RMS of the data residuals are shown by the colored lines. The solid black line is the theoretical trend from Cubillos et al. (2017). The dashed-white line is the median value from simulated pure white noise residuals. The grey region is the 1 and 2σ range for the simulated white noise residuals, which more fully contextualizes if the data are consistent with white noise.

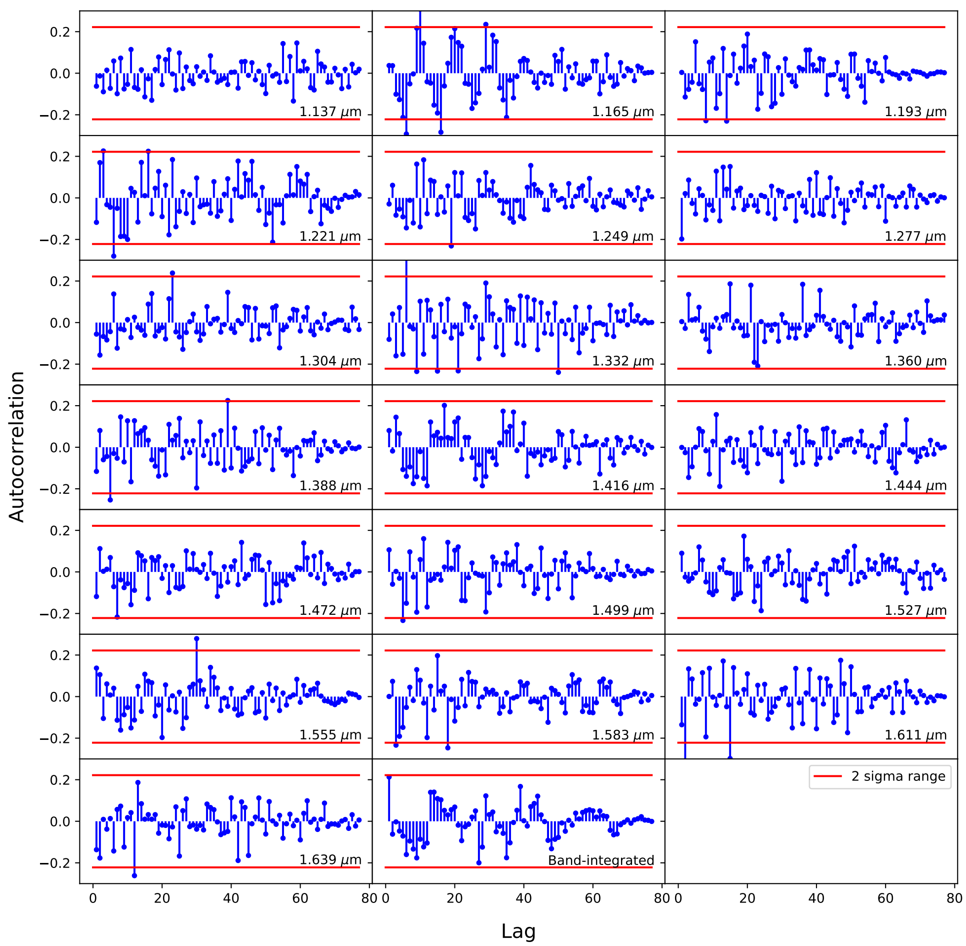


Figure 2.10: Correlated Noise Diagnostic Figure for L9859c. The blue lines and dots show the autocorrelation function (as a function of lag) for each spectral bin, with lag 0 left out for clarity. The solid red lines indicate the 2σ range of autocorrelations: autocorrelation value within these lines are not considered significant. Though difficult to quantify, significant structure is indicative of a time-correlation in the residuals.

white light residuals, which both inflate uncertainties and can potentially interpret random white noise as structure. However, it is not flagged by the normality or correlated noise analyses (described below), and fitting the light curve without incorporating white light residuals finds a consistent depth with a more reasonable $\chi^2_\nu = 0.9$. Therefore, I include it in the transit spectrum. For the other spectral bins, the reduced χ^2 values provide no evidence against validity of the derived transit depths and uncertainties.

A residual normality test checks if the residuals for a model are Gaussian-distributed to determine goodness-of-fit, since this is expected for the "correct" model. Like reduced χ^2, a normality test cannot prove that a model is correct, but can only diagnose incorrect models. I use the `scipy` implementation of the common Shapiro-Wilk test for normality (Shapiro & Wilk, 1965), and determine for which light curves the highest evidence model has normality ruled out at the 5% significance level. At a sample size of around 75 this is by no means rigorous, but it is still a useful heuristic for flagging potentially problematic light curve models.

Normality is rejected at the 5% significance level only for the 1.14 μm spectral bin residuals. Normality is ruled out due to a single outlier in the time-series. When this exposure is ignored, I recover a consistent depth and uncertainty and the residuals are consistent with normality. Further, ignoring residuals again recovers almost the exact same depth without any normality flags. I therefore keep this exposure in the analysis.

Finally, I test for correlated noise in the residuals following the time-average methodology of Cubillos et al. (2017) (also see Pont et al. (2006)) and using MC3[6]. Noise can be thought of as the sum of a purely white (random) noise and a time-correlated (red) noise: $\sigma_{total} = \sqrt{(\sigma_w^2/N + \sigma_r^2)}$ (Pont et al., 2006). As randomly distributed residuals with mean=0 (i.e, if uncorrelated white noise is dominant uncertainty source) are averaged in time, the scatter in the points decreases proportional to $\sigma_w/\sqrt{N}$. If red noise is significant, then the time averaging only decreases noise until it flattens out at σ_r. One can test for the impact of red noise by time-averaging the residuals and comparing the resulting RMS function to theoretical expectations of white noise. For example, first average each point with its neighbor, then recalculate the RMS of those averaged points. Though this method is not necessarily rigorous for HST due to the relatively small number of exposures, it is still a practical diagnostic. I improve upon this method by simulating normally-distributed "residuals" with the same standard deviation as the actual residuals, and putting them through the same method. I note the 1 and 2σ bands for random, pure white noise residuals to contextualize the results for the actual residuals. Interestingly, the median result (white-dashed line) disagrees slightly with MC3's predicted result (solid black line) at a very small amount of bins. For every bin, the residuals are consistent with random white noise for every bin size. I find no evidence of correlated noise (Figures **??** and 2.10).

I also visualize correlated noise by looking at the autocorrelation function of the residuals. This method is not purely quantitative, but can provide another look at potential structure in the residuals. The red lines in Figure 2.10 indicate the 2σ

[6]`https://github.com/pcubillos/mc3`

line — roughly indicating "significant" correlations at that lag. Note that a few lags passing this line is not problematic, since 2-sigma events happen roughly 5% of the time and I am sampling many bins. This is less quantitative, but autocorrelation functions that appear too "structured" can be, unsurprisingly, indicative of structured noise. An example might be 1.165μm, which appears to be a decreasing sinusoid. However, the human eye is great at picking out patterns, and structure below significance is less problematic. Different derivation assumptions give the same transit depth at this bin without structured residuals, and it passes the other red noise test, so I trust the depth. Red noise analysis figures for all of planet b's visits — which also show no evidence of correlated noise — are shown in the Appendix A.

With the caveats noted above, marginalization does an excellent job in fitting the spectral light curves. Together, these tests support the validity of the derived transit depths and uncertainties.

2.5 Exploratory Analysis of Potential Atmospheres

In this section, I investigate if the apparent structures in the spectra of L9859c and L9859b are statistically significant and indicative of atmospheres. Only a low mean-molecular weight atmosphere could produce a feature the size of the potential 1.4μm water feature (roughly 100ppm) (Kreidberg, 2018). Similarly, an estimate for L9859b assuming a water vapor-dominated atmosphere gives an expected signal (for 5 scale heights) of 30ppm, on par with an apparent feature there. However, it is not obvious that a molecular feature would be a significantly improved fit over a straight line (indicative of no atmosphere). Further, stellar activity can potentially contaminate transit spectra and mimic molecular features.

I use the open source retrieval tool PLATON (Zhang et al., 2019) to retrieve atmospheric parameters for planet c and investigate the likelihood of an atmosphere on either planet.

2.5.1 Likelihood and Characteristics of a Hydrogen-rich Atmosphere on Rocky Planet L9859c

The PLATON retrieval works as described in Chapter 3.6.1. It assumes an H_2-He/dominated atmosphere in chemical equilibrium. It does not account for photochemistry. Still, it is useful in contextualizing the spectrum and investigating the likelihood of an H_2-dominated atmosphere on L9859c.

Table 2.5 describes the parameters their priors for the PLATON atmospheric retrieval. I allow planet radius, C/O, metallicity, temperature, and cloudtop pressure to vary, and assume an isothermal temperature profile. C/O and metallicity dictate the elemental ratios in the atmosphere, which are input with temperature into a chemical equilibrium code (ggchem; Woitke et al., 2018) to determine the abundance of every species at every pressure layer. I also fit for stellar radius and planetary mass as nuisance parameters, in order to propagate the uncertainties forward. Each chemical parameter is given a prior set by computational limits (most notably T_{min}=300 K), and the mass/radius priors are set by literature values (Kostov et al., 2019; Cloutier et al., 2019). The retrieval utilizes nested sampling

Table 2.5: Priors for parameters used in L9859c Retrievals

Parameter	Symbol	Prior Distribution	Default Value[a]
Planet Radius	R_p	$\mathcal{U}(0.68, 2.03)$[b]	1.35 R_{Earth}
Limb Temperature	T	$\mathcal{U}(300, 1100)$[c]	550 K
Carbon-oxygen ratio	C/O	$\mathcal{U}(0.05, 2.0)$	0.53[d]
Metallicity	Z	$\mathcal{LU}(-1, 3)$	1 $Z_\odot$
Planet Mass	M_p	$\mathcal{N}(2.40, 0.35)$	2.4 M_{Earth}
Stellar Radius	R_s	$\mathcal{N}(0.314, 0.01)$	0.314 $R_\odot$
Cloudtop Pressure	P_{cloud}	$\mathcal{LU}(-3, 8)$	1 Pa
Stellar Effective Temperature	T_{star}	Fixed	3429 K
Spot Temperature	T_{spot}	Fixed[a]	2920 K
Spot covering fraction	f_{spot}	$\mathcal{U}(0, 0.5)$	0.1

[a] Default values are point estimates from Kostov et al. (2019), Cloutier et al. (2019), and TIC-v8 (Stassun, 2019) (see Table 2.2).
[b] Range is 50–150% of the default value.
[c] Computational minimum to twice the default value.
[d] Solar C/O

(Skilling, 2004; Speagle, 2020) with 200 live points to sample the parameter space and calculate a Bayesian evidence for the model.

The model is an excellent fit ($\chi^2_{\mathrm{Red}} = 1.15$), and the results of the retrieval are shown in Figure 2.11. These can be interpreted as follows: if I assume that L9859c not only has an atmosphere, but also require that it has a H_2-dominated atmosphere with no disequilibrium processes, then these are the characteristics of that atmosphere. This is an example of the parameter estimation/model selection distinction important to Bayesian statistics (Parviainen, 2018). Under these assumptions, L9859c is best described as a high-metallicity atmosphere (Z$\sim 250\times$Z$_\odot$) with a likely super-solar C/O ratio. This atmospheric metallicity is consistent with predictions from the hypothesized mass-metallicity relationship from the Solar System planets (e.g, Mansfield et al., 2018). The retrieved R$_p$ is consistent with the literature (1.30±0.07 R_{Jup} in this work compared to 1.35±0.07 R_{Jup} in (Cloutier et al., 2019)), and though the median temperature is higher than expected, it is also poorly constrained and consistent with the equilibrium temperature of around 450 K. I note that temperatures above ~510 K would likely lead to atmospheric escape of hydrogen and thus are unlikely. However, given the lack of constraint on temperature, the parameter posteriors and model fit would be unchanged by setting a strict prior that T<510 K. The model infers small water features at 1.4μm and 1.1μm, though there is no statistically significant water detection. I note that the inferred water feature at 1.4μm is roughly 100ppm, which is consistent with predictions of the feature size for a H_2-dominated atmosphere (assuming 4 scale heights; derived from Kreidberg (2018)).

PLATON allows for easy model comparison, since nested sampling naturally calculates the Bayesian evidence of a model. Though relatively useless on its own — the evidence cannot act as an absolute goodness-of-fit metric— the ratio of two evidences provides a straightforward measure of how much more likely one model is in comparison to the other. This ratio is known as the odds ratio ($\mathcal{O}_{12} = \mathcal{Z}_1/\mathcal{Z}_2$), and is directly interpreted as "Model 1 is $\mathcal{O}_{12}\times$ more probable than Model 2". There are also empirically-determined benchmarks for converting $\mathcal{O}$ into more familiar σ-level significance (Trotta, 2008; Benneke & Seager, 2013).

To determine the likelihood of an H_2 atmosphere on L9859c, I compare the evidence of the retrieved fiducial atmosphere to that of a flat line. I model the flat line by fixing a high, grey cloud (manifesting as a straight line in the spectrum) and only allowing planet radius to vary. The resulting fit is shown in the bottom-right

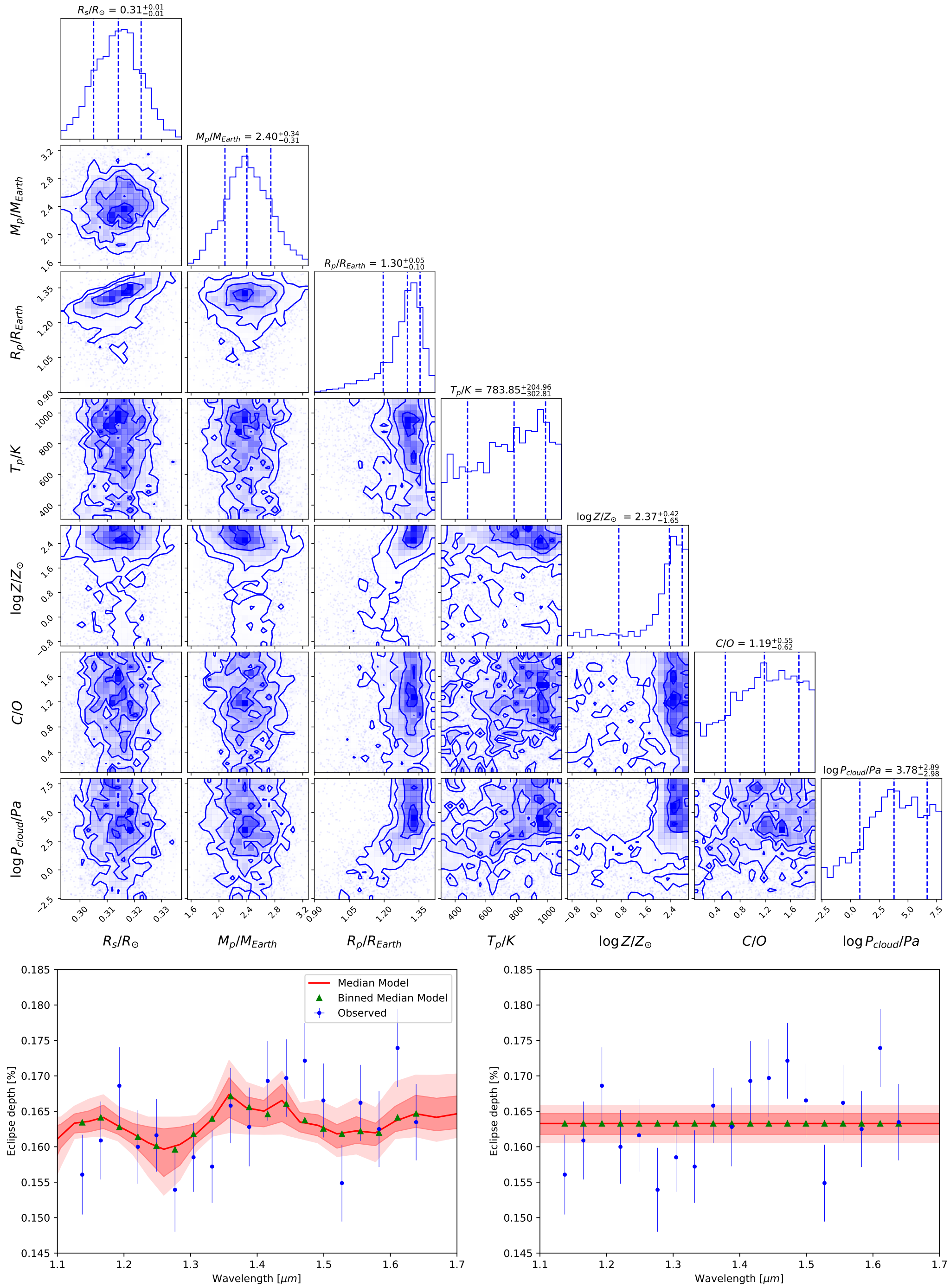


Figure 2.11: Retrieval results for L9859c transit. **Top:** Corner plot for fiducial H_2-dominated atmosphere. **Bottom Left:** Median retrieved model (green triangles) with 1 and 2-sigma contours (red) plotted over data (blue) for fiducial atmosphere model. **Bottom right:** Median retrieved no atmosphere model.

panel of Figure 2.11. The odds ratio between the fiducial model and no atmosphere is 3, indicating that, based on the observed spectrum, it is $3\times$ more likely that L9859c has an atmosphere than not. This corresponds to a "weak" detection of roughly 2.1σ, or about 75% probability. The specific sigma significance should be taken with a grain of salt, since even a small numerical error in $\mathcal{Z}$ — which is common (Speagle, 2020) — could shift the odds ratio slightly below the empirical cut-off. I also compare the Bayesian evidences between the fiducial atmosphere and an atmosphere with no water opacity, finding the odds ratio to be roughly one. This indicates that there is no conclusive evidence of the presence of water vapor in the atmosphere. The evidence of water vapor specifically is weaker than that of an atmosphere since other opacity sources (e.g, NH_3) can "fill-in" for water vapor and capture some of the structure in the observed spectrum.

2.5.2 Discussion and Impact of Stellar Activity

Even if transit signal is real, it may not be indicative of a planetary atmosphere. Instead, it could be due to stellar activity. While spot-crossing events are generally detectable in the light curve data (e.g, Huitson et al., 2013), the impact of unocculted star spots is generally more nefarious. Practically, this refers to the in transit and out of transit stellar baselines differing due to a difference in stellar spot and faculae coverage. The logic is as follows: the baseline flux established out-of-transit is from the entire integrated stellar disk. In transit, the incident light passing through the planet's atmosphere is from a small fraction (less than a few percent) of the star. If the planet's transit chord passes through a relatively immaculate photosphere (i.e, no spots), then the inferred incident light (the integrated stellar disk) and the actual incident light differ significantly enough to overestimate transit depth. Worse, this varies with wavelength and can cause false structure, or "mock" features (Seager & Sasselov, 2000). Importantly, M-dwarf are cool enough to exhibit significant water absorption in their spectra, and solar-like star spots have stronger water absorption, so a planet with no atmosphere could appear to have a water feature if the host star's spot-covering fraction were high enough (Rackham et al., 2018). I first discuss likelihood of stellar activity being relevant to L98, then try to estimate its impact using activity models for M-dwarfs (Rackham et al., 2018).

1–3% of M-dwarfs exhibit spot-covering fractions of at least 10% (Goulding et al., 2012), and the variability amplitude of M-dwarfs is roughly an order of magnitude greater than variability in the Sun (Kopp et al., 2005; Newton et al., 2016). Rackham et al. (2018) modeled how this variability impacted transit spectra for several spot models. For L9859's spectral type (M3V-dwarf), Rackham et al. (2018) find that stellar contamination for common spectral features in the near-infrared (e.g, water) can easily dominate real features. Further, "mimicked" features are possible, with variations in the water band on the order of 200ppm possible for L9859c. For context, the apparent water feature is roughly 100ppm. This is an issue in the similar TRAPPIST-1 system (Rackham et al., 2018).

As a population M-dwarfs are more susceptible to activity, but what about L9859 specifically? The TESS monitoring observations of L9859 did not show significant variability and describe a quiet star (Kostov et al., 2019; Pidhorodetska et al., 2021). However, a recent re-analysis of the TESS data revealed a clear flare, indicating L9859 is likely active (Barclay, private comms). The data set contains 5 transits of a single star, and 4 of the same transit, which does help constrain variability. Looking at Figure 2.7, there is not obvious extreme variability — the distribution of the depths for each bin seem to line up decently well with the estimated uncertainties. Similarly, planet b visit 02 and planet c were observed 14 hours apart, so variability must be very frequent for contamination to impact the two significantly differently. Alternatively, variability is not necessary for stellar contamination. For example, a large amount of solar-type (as opposed to giant) spots distributed uniformly longitudinally would not exhibit obvious variability but could still systematically bias transit spectra. This would be more likely if spots on L9859 are, like the Sun (Mandal et al., 2017), dependent on latitude. For both planets to be systematically affected, they would both need to not occult many spots during transit, which is plausible given that they have very similar inclinations (88.7 and 89.3 degrees, respectively, consistent to 1σ). This could potentially allow for a situation where the chord of both planets' transits consistently do not occult areas densely populated by spots. Overall, though the specific spot (and faculae) model and prevalence are unknown for L9859c, there is enough evidence to accept that some spectral contamination is likely, and that a systematic bias of all 5 transits is possible.

To quantify this, I use the built in stellar activity model in PLATON, discussed in Section 3.7.2.4. Following Rackham et al. (2018), I set the spot temperature for the T=3400 K star to 2920 K. I fit for the fractional spot coverage as an open parameter, only requiring that it stay under the conservative 50%. PLATON weights the contributions from the different temperature regimes via this fractional coverage parameter, which represents the fractional overabundance of spots in the unocculted regions as compared to the occulted regions. From this weighting it derives a wavelength-dependent correction factor, which models stellar activity's impact on the observed spectrum and disentangles planetary features from stellar contamination. I follow the flat line fit from Section 2.5.1 by forcing a high cloud, and fixing every parameter except R_p (and coverage fraction). This approximates a planet with no atmosphere transiting in front of an active star with a disproportionate amount of the unocculted area being star spots.

The results of this fit are shown in Figure 2.12. The model is an excellent fit, with reduced chi-squared=0.9. More importantly, its odds ratio against the "normal atmosphere/non-active star" model is 3.5, indicating it is 3.5× more likely and weakly preferred. For arbitrarily stricter spot coverage priors (f_{spot} <20%), the two evidences are roughly equal, indicating no statistical preference. I performed the same activity retrieval on the weighted L9859b spectrum, and the results are also given in Figure 2.12. Though there is no atmosphere model for comparison, the stellar activity model is able to explain the L9859b spectrum with the same spot coverage fraction distribution as L9859c (χ^2_{red}=0.8). Further, an active star with such high spot coverage would emit a greater amount of XUV rays that would make retaining an atmosphere more difficult. For example, water in a steam atmosphere on L9859b would be more frequently photodissociated, speeding up the process of atmospheric escape (Bolmont et al., 2017). Consequently, I do not claim any

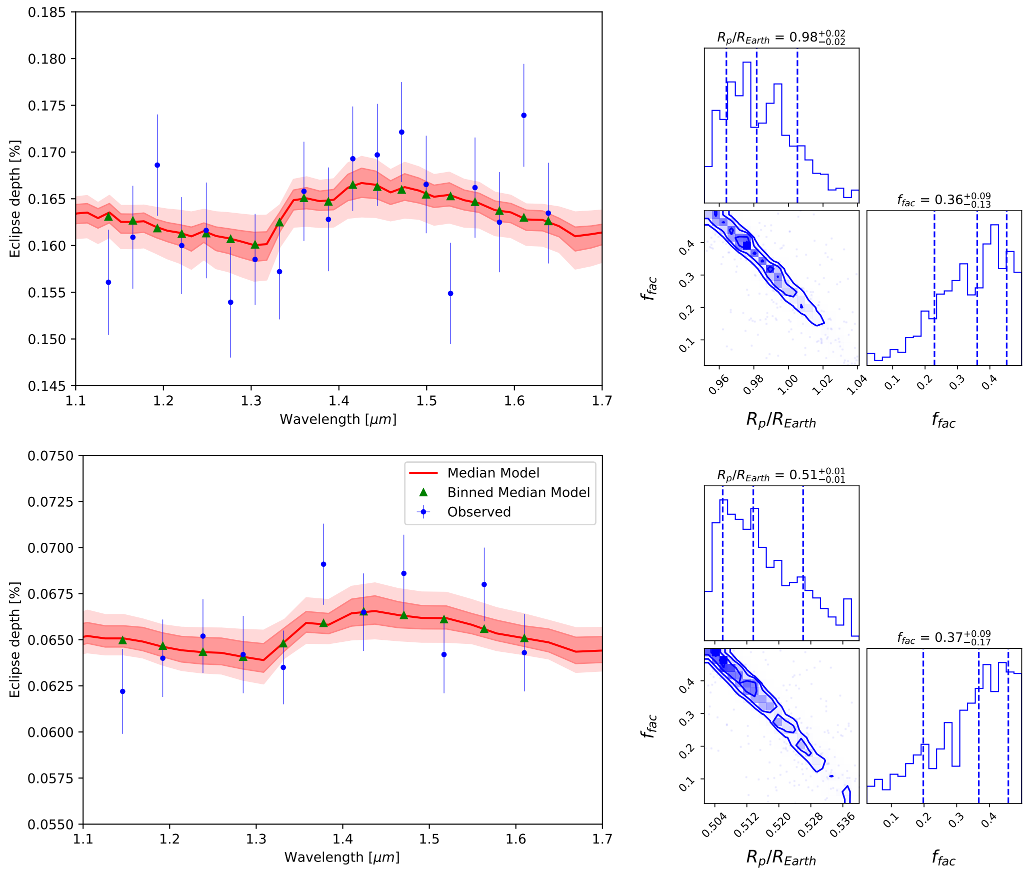


Figure 2.12: Transit retrieval results assuming no atmosphere and an active star with spots. **Top:** Results for L9859c. This model is preferred over atmosphere model by a factor of 3.5. **Bottom:** Results for L9859b. Notably, the same spot coverage fraction is able to produce each planet's spectrum without any planet atmospheres. Radii in these corner plots are not physical or meaningful.

atmospheric detection, and find stellar activity to be the most likely explanation of the spectrum. I note that an atmosphere might exist but be masked by stellar contamination. If the variability of L9859 becomes better understood, then its star spots can be more accurately modeled and disentangled from the spectrum in order to characterize L9859c's atmosphere (Rackham et al., 2018). Wakeford et al. (2019) did this exact procedure for TRAPPIST-1g, concluding that stellar contamination is unlikely for that planet and allowing the authors to rule out a hydrogen-rich atmosphere to greater than 3σ.

Ultimately, L9859c is an interesting follow up planet. The weak preference for the presence of mock water features in both L9859c and L9859b constitute the first observational evidence of mock features due to stellar contamination. Unlike other Earth-sized planets (GJ1132b, TRAPPIST 1d/e/f Diamond-Lowe et al., 2018; Mugnai et al., 2021; de Wit et al., 2018), the observed water feature and the uncertainty in spot coverage do not allow me to rule out a H_2-dominated atmosphere. While I cannot definitively rule out any of the atmospheric scenarios mentioned, L9859b and L9859c are an excellent case study on how stellar activity can impact interpretation in terrestrial planets orbiting M-dwarfs. Another HST WFC3 transit of L9859c and the first transit observation of the probable mini-Neptune L9859d are scheduled and can provide insight into impact and prevalence of activity. The follow-up planet c observation should increase the SNR of the inferred feature if it is real (and decrease it if not real). Further, L9859d is expected to have a hydrogen-dominated atmosphere (Pidhorodetska et al., 2021). If its observed spectrum contains a larger-than-expected water feature, than that is evidence that stellar contamination is manifesting as mock features and increasing the spectral feature size. Alternatively, if the spectrum appears flat (e.g, due to clouds), then it would act as evidence against stellar contamination and instead make an atmosphere on L9859c more likely. Follow-up observations with JWST, which has enough precision to resolve features from even a non-H_2 dominated atmosphere in a single visit (Pidhorodetska et al., 2021), will also be tremendously useful. Fortunately, L9859 will be one of the first systems observed by JWST (Pidhorodetska et al., 2021).

Chapter 3: A Metal-rich Atmosphere for the Inflated Hot Jupiter HAT-P-41b

3.1 Overview

In this chapter, I present a comprehensive analysis of the 0.3–5 μm transit spectrum for the inflated hot Jupiter HAT-P-41b. The planet was observed in transit with Hubble STIS and WFC3 as part of the Hubble Panchromatic Comparative Exoplanet Treasury (PanCET) program, and I combine those data with warm *Spitzer* transit observations. We extract transit depths from each of the data sets, presenting the STIS transit spectrum (0.29–0.93 μm) for the first time. I retrieve the transit spectrum both with a free-chemistry retrieval suite (AURA; collaborator) and a complementary chemical equilibrium retrieval suite (PLATON) to constrain the atmospheric properties at the day-night terminator. Both methods provide an excellent fit to the observed spectrum. Both AURA and PLATON retrieve a metal-rich atmosphere for almost all model assumptions (most likely O/H ratio of $\log_{10} Z/Z_\odot = 1.46^{+0.53}_{-0.68}$ and $\log_{10} Z/Z_\odot = 2.33^{+0.23}_{-0.25}$, respectively); this is driven by a 4.9-σ detection of H_2O as well as evidence of gas absorption in the optical (>2.7-σ detection) due to Na, AlO and/or VO/TiO, though no individual species is strongly detected. Both retrievals determine the transit spectrum to be consistent with a clear atmosphere, with no evidence of haze or high-altitude clouds. Interior modeling constraints on the maximum atmospheric metallicity ($\log_{10} Z/Z_{\odot} < 1.7$) favor the AURA results. The inferred elemental oxygen abundance suggests that HAT-P-41b has one of the most metal-rich atmospheres of any hot Jupiters known to date. Overall, the inferred high metallicity and high inflation make HAT-P-41b an interesting test case for planet formation theories.

3.2 Introduction

Transit spectroscopy has been fundamental in understanding the physics and chemistry of hot exoplanet atmospheres. Transit observations with the Hubble Space Telescope (HST) and the *Spitzer* Space Telescope have been especially fruitful in illuminating the composition and atmospheric structure of close-in planets, starting with the first measurements of sodium absorption (Charbonneau et al., 2002) and the first detection of thermal emission (Deming et al., 2005) for the atmosphere of HD209458b.

The installation of the Wide Field Camera 3 (WFC3) instrument and the refurbishment of the Space Telescope Imaging Spectrograph (STIS) on HST opened up a new era of transit spectroscopy measurements for hot Jupiters. WFC3 has provided the first repeatable and well-validated detections of the presence of water vapor (Deming et al., 2013; Huitson et al., 2013; Wakeford et al., 2013; Mandell et al., 2013), and has opened the field to population studies looking at H_2O abundance and metallicity as a function of stellar and planetary properties (Sing et al., 2016; Tsiaras et al., 2018; Pinhas et al., 2019; Welbanks et al., 2019). The upgraded STIS instrument has been a key contributor in illuminating the critical role that aerosols play in driving the continuum opacity for transit measurements of hot planets (Pont et al., 2013; Nikolov et al., 2014; Sing et al., 2016; Chachan et al., 2019b).

One of the most intriguing topics from these studies is the question of atmospheric metallicity. Studies of individual planets suggested a wide diversity of atmospheric metallicity as a function of planetary mass (e.g., Madhusudhan et al., 2014c; Kreidberg et al., 2014b; Wakeford et al., 2017, 2018). However, recent homogenous statistical analyses of many planets reveal that a paucity of water vapor in hot planet atmospheres is the norm (Barstow et al., 2017; Pinhas et al., 2018; Welbanks et al., 2019). One strategy to investigate the relationships between mass and atmospheric metallicity is to study the best targets within the Saturn and Jupiter mass range, in order to achieve high S/N and leverage the expectation of a high primordial gas fraction and large transit signals.

First discovered in 2012 (Hartman et al., 2012), the inflated hot Jupiter HAT-P-41b (T_{eq}=1940K, P=2.7 days) is a strong candidate to inform these trends. It is among the most inflated hot Jupiters (R=1.69R_{Jup}, M=0.8M_{Jup}), and it orbits a relatively inactive mid-F dwarf (R=1.68$R_{\odot}$, $T_{\rm eff}$ = 6390 K). HAT-P-41b's extended atmosphere and its host star's lack of significant variability make it highly amenable to characterization through transit spectroscopy. Johnson et al. (2017) determined the spin-orbit misalignment of the system to be moderate (-22°), while the host star appears to be part of a multi-stellar system, with a wide-orbit late-type companion discovered at ~ 1000 AU (Hartman et al., 2012; Wöllert & Brandner, 2015; Evans et al., 2016).

Tsiaras et al. (2018) retrieved the WFC3 G141 grism spectrum (1.1–1.7 μm) with τ-Rex (Waldmann et al., 2015a,b), confirming an 4.2σ atmospheric detection and finding no evidence of contributions from either high-altitude clouds or photochemical hazes (e.g., Zahnle et al., 2009a). Tsiaras et al. (2018) also found evidence of and abundance constraints for water vapor ($\log_{10}(X_{H_2O}) = -2.77 \pm 1.09$), though due to narrow wavelength coverage abundance uncertainties are large. Still, they are able rule out upper atmospheric water depletion. Fisher & Heng (2018) built upon this result by retrieving on the same dataset with a focus on cloud opacity and other near-infrared opacity sources (NH_3, HCN). They find a water abundance of $-0.9^{+0.28}_{-1.20}$ which they note is generally consistent with that of Tsiaras et al. (2018). However, this reported abundance is for a cloud-free model with only H_2O and NH_3 as opacity sources, and this simplified treatment is not necessarily directly comparable with more comprehensive atmospheric models. They also find weak evidence of NH_3, though they are unable to favor the NH_3 and H_2O model over a model with grey clouds and H_2O.

Wide spectral baselines provide the potential for a more complete and constrained understanding of atmospheric properties (Benneke & Seager, 2013; Griffith, 2014; Welbanks & Madhusudhan, 2019). For example, Line & Parmentier (2016) demonstrated how individual WFC3 spectra are unable to constrain mean molecular weight due to a degeneracy with partial clouds. Furthermore, for a fully homogeneous cloud cover the cloud-top pressure is degenerate with the chemical abundance (e.g., Deming et al., 2013). Welbanks & Madhusudhan (2019) showed that optical data help alleviate such degeneracies and improve the precision with which planetary radius, cloud properties, and molecular/atomic abundances are inferred. As a practical example, Sing et al. (2016) utilized optical-to-infrared spectra to jointly constrain cloud, haze, and chemistry parameters for a sample of ten hot Jupiters. STIS data have been specifically useful in constraining the atmospheric metallicity of giant exoplanets (Wakeford et al., 2017, 2018; Chachan et al., 2019a).

Bayesian spectral retrievals are the most reliable way to interpret exoplanet spectra, due to their flexibility in describing diverse exoplanet atmospheres and their ability to evaluate the full posterior distribution of a forward model's parameters (Madhusudhan, 2018). This allows for understanding not only the properties of an exoplanet's atmosphere, but also the uncertainties on those properties. Consequently, such retrieval codes are common in atmospheric characterization (Madhusudhan & Seager, 2009, 2010a; Lee et al., 2012; Benneke & Seager, 2013; Line et al., 2013; Amundsen et al., 2014; Waldmann et al., 2015b; Barstow et al., 2017, and many others). Nested sampling (Skilling, 2004) is a particularly powerful Bayesian sampler, as it naturally determines the Bayesian evidence of the fitted model (with the posterior distribution being a byproduct), which is necessary for correct model comparison (e.g., justifying more complicated models, reporting correct detection significances).

Despite their ubiquity, each retrieval code is necessarily unique given the assumptions and modeling choices that must be made. Though these retrievals generally agree, subtle discrepancies can lead to different conclusions for the same data

(Kilpatrick et al., 2018; Fisher & Heng, 2018; Barstow et al., 2020). Examples include different chemical parameterizations (i.e., enforcing chemical equilibrium), cloud parameterizations, opacity sources, and prior assumptions. Therefore, it is important to understand the effect of modeling assumptions on the retrieved atmospheric parameters (e.g., Welbanks & Madhusudhan, 2019). Testing a suite of models for a given retrieval code — and, even better, different modeling paradigms altogether — more accurately captures the uncertainty in the atmospheric parameters. It is important to be transparent about the assumptions made in a retrieval analysis in order to best contextualize and understand the results.

In this chapter I derive the 0.3–5 μm transit spectrum of HAT-P-41b using transit observations from HST/STIS, HST/WFC3, and *Spitzer* (Sec. 3.3). Section 3.4 characterizes both the variability (incorporating both X-ray and visible photometric monitoring; Sec. 3.4.1) and the parameters (Sec. 3.4.2) of the host star. Section 3.5 describes the data analysis to derive the transit spectrum. Sec. 3.6 describes that we use two different retrieval methods. First, I use a chemical-equlibirum framework (PLATON Zhang et al., 2019, Sec. 3.7), and those results are described in Sec. 3.8. We also explore a more flexible free-chemistry retrieval using the AURA framework (Pinhas et al., 2018, Sec. 3.9). The two retrieval analyses were independently done by different members of the team to allow for an unbiased comparison. I conclude that a high, supersolar atmospheric metallicity (on the order of 30–200× solar O/H) best describes the observed spectrum, and — though median retrieved values differ — this result is not sensitive to model assumptions. Sec. 3.10 discusses the comparison between retrievals (Sec. 3.10.1, the comparison to interior modeling constraints (Sec 3.10.2, and the implications for planet formation (Sec 3.10.3). Sec. 3.11 provides a summary of my conclusions.

3.3 Observations

3.3.1 HST

We observed one transit of HAT-P-41b with the WFC3 instrument on HST and three transits with the STIS instrument as part of the PanCET Program (ID 14767, P.I. Sing). The WFC3 observations were taken on October 16, 2016 using the G141 prism, which covers a wavelength range of approximately 1.1-1.7 μm with a spectral resolving power of R$\sim$150. The STIS observations were taken with the G430L and G750L grisms, which cover a wavelength range of approximately 0.3-1.0 μm with a spectral resolving power of R$\sim$500. The STIS data were acquired on September 4 2017 (G430L, visit 83), May 7 2018 (G430L, visit 84) and June 11 2018 (G750L, visit 85). For each visit, the target was observed for 7 hours over five consecutive HST orbits. An HST gyro issue prevented the acquisition of the third orbit for visit 85.

For the WFC3 observations, data were taken in spatial scan spectroscopic mode with a forward scanning rate of 0.065 arcsec s^{-1} along the cross-dispersion axis, resulting in scans across approximately 46 pixel rows. The observations utilized the 256 $\times$ 256 pixel subarray and the SPARS-10 sampling sequence, with 12 non-destructive reads (NSAMP = 12) resulting in a total integration time of 81 seconds for each exposure. HST obtained a total of 17 exposures in the first HST orbit following acquisition and 19 exposures in each subsequent HST orbit. Typical peak frame counts were ~33,000 electrons per pixel, which is within the linear regime of the WFC3 detector.

For the STIS observations, each visit consisted of 5 orbits, with $\sim 45\,\text{min}$ gaps due to Earth occultations. We utilized the wide 52×2 arcsec slit to minimize slit light losses and an integration time of ~253 sec for each exposure, for a total of 48 spectra for each visit. Data acquisition overheads were minimized by reading-out a subarray of the CCD with a size of 1024×128 pixels.

3.3.2 Spitzer

The Spitzer Infrared Array Camera (IRAC) observations were taken in January and February 2017 as part of Program 13044 (P.I. D. Deming). A single transit of HAT-P-41b was observed in each of the IRAC1 ($3.6\,\mu$m) and IRAC2 ($4.5\,\mu$m) channels. Each transit was preceded by a 30-minute peakup sequence that also mitigates the steepest portion of a temporal ramp due to the detector. The transit was observed over ~12 hours, with equivalent in-transit and out-of-transit coverage. 338 exposures were obtained for each transit, and each exposure consisted of 64 subarray frames of 32x32 pixels, using an exposure time of 2.0 seconds per frame.

3.3.3 Photometric Monitoring Observations

To better diagnose the likelihood of stellar variability impacting the transit spectrum, I complemented the transit observations with monitoring observations at both visible (AIT) and X-ray wavelengths (XMM-Newton). XMM-Newton observed HAT-P-41 on 2017-April-07 with an overall 17 ks exposure time (Proposal ID 80479, P.I. J. Sanz-Forcada). The target was not detected in any of the EPIC detectors; I discuss the implications of this in Section 3.4.1.

A collaborator obtained nightly ground-based photometric observations of HAT-P-41 during its 2018 and 2019 observing seasons with the Tennessee State University Celestron 14-inch (C14) automated imaging telescope (AIT) located at Fairborn Observatory in the Patagonia Mountains of southern Arizona (see, e.g., Henry, 1999; Eaton et al., 2003). The AIT is equipped with an SBIG STL-1001E CCD camera; observations were made through a Cousins R filter. Details of the observing, data reduction, and analysis procedures are described in Sing et al. (2015).

They collected a total of 207 successful nightly observations (excluding a few isolated transit observations) over the two observing seasons. The observing activities at Fairborn must come to a halt each year during the southern Arizona rainy season, which typically lasts from approximately July 1 to September 10. Since HAT-P-41 comes to opposition around July 18, each observing season is broken into two intervals, which we designate as intervals A and B. Information for a portion of the AIT observations are shown in Table 3.1; the full table is available in the electronic edition of *ApJ*.

Table 3.1: AIT photometric observations of HAT-P-41

Hel. Julian Date (HJD − 2,400,000)	Delta R (mag)	Sigma (mag)
58175.0255	−0.56702	0.00122
58176.0214	−0.56986	0.00109
58180.0102	−0.56636	0.00052
58181.0158	−0.56444	0.00293
58182.0022	−0.56479	0.00185
58184.9980	−0.56390	0.00182

NOTE – This table shows a sample of the full set of observations.

3.4 Stellar Properties

3.4.1 Analysis of Stellar Variability

The results of analysis of the AIT photometric observations (Sec. 3.3.3) are given in Table 3.2. The low numbers of observations in 2018 B, 2019 A, and 2019 B are the result of the unusually cloudy weather at Fairborn for the past two years. This cloudy weather pattern continues to the present. Column 4 of the table gives the standard deviation of the individual observations with respect to their corresponding seasonal mean. The standard deviations range between 0.00224 and 0.00394 mag for the four observing intervals. This is near the limit of the nightly measurement precision with the C14, as determined from the constant comparison stars in the field. Periodogram analyses of the four intervals reveal no significant periodicities. The scatter in the seasonal means given in column 5 is consistent with the expected photometric precision considering the small number of observations in the last three intervals and the marginal photometric conditions prevalent at

Table 3.2: Results of the analysis of photometric monitoring observations for HAT-P-41

Observing Season	N_{obs}	Date Range (HJD − 2,400,000)	Sigma (mag)	Seasonal Mean (mag)
2018 A	110	58175–58295	0 00224	-0.56496 ± 0.00021
2018 B	34	58386–58451	0.00291	-0.56979 ± 0.00051
2019 A	42	58577–58657	0.00264	-0.56966 ± 0.00041
2019 B	21	58756–58802	0.00394	-0.56602 ± 0.00088

Fairborn Observatory over the past two years. Therefore, HAT-P-41 appears to be constant on night-to-night and year-to-year timescales to the limit of the telescope's precision.

Additionally, HAT-P-41 was not detected in any of the XMM Newton's EPIC detectors. Given the distance of the object I can set an upper limit of $L_{\rm X} = 1 \times 10^{29}$ erg s^{-1} on the stellar X-ray luminosity. This implies a value of $\log L_{\rm X}/L_{\rm bol} < -5.2$, indicating that the star has a moderate activity level at most (Wright et al., 2011).

The photometric observations of HAT-P-41 describe a relatively quiet star. Furthermore, the Ca II chromospheric activity index (S = 0.18 Duncan et al., 1991) and the corresponding estimated parameter flux $\log R'_{HK}$ (-5.04) for HAT-P-41's spectral type (B-V = 0.29) are not indicative of high activity (Hartman et al., 2012; Noyes et al., 1984) and may indicate instead a basal-level activity (Isaacson & Fischer, 2010). Rackham et al. (2019) show that for all but the most active F-dwarfs variability does not result in any detectable change to the transit spectrum. Specifically, the impact of potential complications such as false TiO/VO detections, false water detections, and optical offsets are all determined to be less than $\sim$ 10ppm. Therefore, I conclude that stellar variability is unlikely to contaminate HAT-P-41b's transit spectrum.

3.4.2 Stellar Parameters

Inferred atmospheric planetary parameters are directly dependent on host star parameters. For my analyses, I incorporate the stellar parameters from TESS Input Catalog — version 8 (TIC-8; Stassun et al., 2019; Stassun, 2019). TIC-8 provides reliable stellar parameters for planetary host stars based primarily on Gaia Data Release 2 (GDR2) point sources (Gaia Collaboration et al., 2016, 2018). The algorithm for HAT-P-41's parameters is as follows: distance is first derived from Gaia DR2 parallax, using a correct inference procedure (Bailer-Jones et al., 2018). HAT-P-41's galactic longitude (-10.6) puts it in a region where uncertainty on reddening makes determining effective temperature from Gaia photometry difficult. As a result, a spectroscopically-derived effective temperature (from the PASTEL catalog Soubiran et al., 2016) is preferred. The stellar radius and mass are then self-consistently derived from the distance and effective temperature (Andrae et al., 2018). Finally, $\log g_s$ is calculated from the stellar radius and mass.

It is important to recalculate R_p, M_p, and semi-major axis a based on the R_s value from TIC-8, since those are derived in the discovery paper assuming a certain value for R_s. As a simple example, R_p is derived by constraining R_p/R_s in transit and multiplying by R_s. To re-derive the planetary parameters, I follow the methodology of Stassun et al. (2017). The resulting values and 1-σ ranges are shown in Table 3.3 along with the values from the discovery paper (Hartman et al.,

Table 3.3: System Parameters for HAT-P-41

Parameter	TIC-8[a]	Discovery Paper[b]
$R_s[R_{Sun}]$	$1.65^{+0.08}_{-0.06}$	$1.683^{+0.058}_{-0.036}$
$M_s[M_{Sun}]$	$1.32^{+0.25}_{-0.16}$	1.42 ± 0.047
$\log g_s$ [cgs units]	$4.12^{+0.11}_{-0.06}$	4.14 ± 0.02
$T_{s,\mathrm{eff}}$ [K]	6480^{+100}_{-100}	6390 ± 100
$R_p[R_{Jup}]$	$1.65^{+0.08}_{-0.07}$	$1.685^{+0.076}_{-0.051}$
$M_p[M_{Jup}]$	$0.76^{+0.14}_{-0.12}$	0.80 ± 0.10
ρ_p [g cm^{-3}]	0.21 ± 0.05	0.20 ± 0.03
$T_{p,\mathrm{eq}}$ [K]	1960^{+40}_{-35}	1940 ± 38
a [AU]	$0.0418^{+0.0021}_{-0.0019}$	0.0426 ± 0.0005
Distance [pc]	348 ± 4.5	344^{+12}_{-8}

[a] Provided by or derived from Tess Input Catalog Stassun et al. (2019)
[b] Hartman et al. (2012)

2012).

I favor the TIC-8 stellar parameters over the discovery paper values (derived using isochrones and high-resolution spectroscopy; Hartman et al., 2012) since they are based on more recent and comprehensive data. I emphasize that the two sets of parameters are consistent to better than 1-σ, and using the discovery paper values has no impact on the conclusions of this chapter.

A planet's composition is directly linked to its host star's composition. Brewer & Fischer (2018a) determined the stellar abundance of 15 different elements for HAT-P-41 as part of the Spectral Properties of Cool Stars (SPOCS) catalog. Table 3.4 gives the abundances, relative to solar, for the relevant elements. Brewer & Fischer (2018a) find an effective temperature, metallicity, and $\log g_s$ consistent with

Table 3.4: HAT-P-41 Host Star Elemental Abundances

Elemental ratio	Abundance (log solar unit)
[O/H]	0.37 ± 0.04
[C/H]	-0.08 ± 0.03
[Na/H]	0.17 ± 0.01
[Ti/H]	0.22 ± 0.01
[V/H]	0.09 ± 0.03
[Al/H]	0.07 ± 0.03
[M/H]	0.18 ± 0.01
[C/O]	-0.45 ± 0.05

NOTE—All values from Brewer & Fischer (2018b)

both TIC-8 and the discovery paper. HAT-P-41 is a metal-enriched star, and notably has an elemental oxygen abundance of 2.3× solar. Carbon is the only depleted element at $\sim 0.8\times$ solar, resulting in a subsolar C/O ratio of 0.19 (0.36× solar).

3.5 Data Analysis

3.5.1 STIS

3.5.1.1 Data Reduction

The STIS data analysis procedures follow the general methodology detailed in Nikolov et al. (2014, 2015). A collaborator commenced analysis from the flt.fits files, which were reduced (bias-, dark- and flat-corrected) using the latest version of the CALSTIS pipeline and the latest calibration frames. We used median combined difference images to identify and correct for cosmic-ray events in the images as described by Nikolov et al. (2014). We found ~ 4 percent of the detector pixels were affected by cosmic ray events. We also corrected pixels identified by CALSTIS as bad with the same procedure, which together with the cosmic ray identified pixels resulted in a total of $\sim$14 percent interpolated pixels.

We performed spectral extraction with the IRAF procedure APALL using aperture sizes in the range from 6 to 18 pixels with a step of 0.5. The best aperture for each grating was selected based on the resulting lowest light-curve residual scatter after fitting the white light curves. We found that aperture sizes 13.5, 13.5 and 10.5 pixels satisfy this criterium for visits 83, 84 and 85, respectively.

We cross-correlated and interpolated all spectra with respect to the first spectrum to account for sub-pixel wavelength shifts in the dispersion direction. The STIS spectra were then used to extract both white-light spectrophotometric time series and custom wavelength bands after summing the appropriate flux from each bandpass.

The raw STIS light curves exhibit instrumental systematics on the spacecraft orbital time-scale, which are attributed to thermal contraction/expansion (referred to as the 'breathing effect') as the spacecraft warms up during its orbital day and cools down during orbital night. We take into account the systematics associated with the telescope temperature variations in the transit light-curve fits by fitting a baseline function depending on various parameters.

3.5.1.2 Light Curve Analysis

White and spectroscopic light curves were created from the time series of each visit by summing the flux of each stellar spectrum along the dispersion axis. We fit each transit light curve using a two-component function that simultaneously models the transit and systematic effects. The transit model was computed using the analytical formulae given in Mandel & Agol (2002), which are parameterized with the mid-transit times (T_{mid}), orbital period (P) and inclination (i), normalized planet semimajor axis (a/R_*), and planet-to-star radius ratio (R_p/R_*).

Stellar limb-darkening was accounted for by adopting the four parameter non-linear limb-darkening law with coefficients c1, c2, c3 and c4, computed using a three-dimensional stellar atmosphere model grid (Magic et al., 2015). We adopted the closest match to the effective temperature, surface gravity, and metallicity values for HAT-P-41 determined by Hartman et al. (2012).

As in their past STIS studies, my collaborator applied orbit-to-orbit flux corrections by fitting for a fourth-order polynomial to the spectrophotometric time series phased on the HST orbital period and a linear time term. We also used a low-order polynomial (up to a third degree with no cross terms) of the spectral displacement in the dispersion and cross dispersion direction. The first exposures of each HST orbit exhibit lower fluxes and have been discarded in the analysis. Similar to our past HST STIS analyses (Nikolov et al., 2014; Sing et al., 2016), we intended to discard the entire first orbit to minimize the space craft thermal breathing trend, but found that for two of the three HST visits, a few of the exposures taken toward the end of the first orbit can be used in the analysis.

We then generated systematics models that spanned all possible combinations of detrending variables and performed separate fits using each systematics model included in the two-component function. The Akaike information criterion (AIC;

Akaike, 1974) was calculated for each attempted function and used to marginalize over the entire set of functions following Gibson (2014b). My choice to rely on the AIC instead of the Bayesian information criterion (BIC; Schwarz, 1978) was determined by the fact that the BIC is more biased toward simple models than the AIC. The AIC therefore provides a more conservative model for the systematics and typically results in larger or more conservative error estimates, as demonstrated by Gibson (2014b). Marginalization over multiple systematics models assumes equal prior weights for each model tested.

For the white light curves, we fixed the orbital period, inclination and a/Rs to the values reported in Table 3.5 and fit for the transit mid-time and planet-to-star radius ratio. We find central transit times T_C[MJD]$= 58000.6958 \pm 0.0029$ (visit 83), T_c[MJD]$= 58245.85414 \pm 0.00039$ (visit 84), T_c[MJD]$= 58280.87484 \pm 0.00036$ (visit 85). We derive the white light transit depths to be 10200 ± 104 ppm and 10320 ± 85 ppm for G430L and G750L, respectively.

For the spectroscopic light curves, a common-mode systematics model was established by simply dividing the white-light curve by a transit model (Berta et al., 2012; Deming et al., 2013). We computed the transit model using the orbital period, inclination and a/Rs from Table 3.5 and the central times for each orbit from the whit-light analysis. The common-mode factors from each night were then removed from the corresponding spectroscopic light curves before model fitting.

We then performed fits to the spectroscopic light curves using the same set of systematics models as in the white-light curve analysis and marginalized over them as described above. For these fits, Rp/Rs was allowed to vary for each spectroscopic

Table 3.5: Transit parameters for HAT-P-41b

Parameter	Value
R_p/R_s	0.1028 ± 0.0016
a/R_s	$5.44^{+0.09}_{-0.15}$
i [Degrees]	87.7 ± 1.0
T_c [BJD]	$2454983.86167 \pm 0.00107$
$\log g_p$ [cgs units]	2.84 ± 0.06
P [days]	$2.694047 \pm 4 \times 10^{-6}$

All values from discovery paper, Hartman et al. (2012)

channel, while the central transit time and system parameters were fixed. We assumed the non-linear limb-darkening law with coefficients fixed to their theoretical values, determined in the same way as for the white-light curve. The detrended spectrophotometric light curves are shown in Figures B.1, B.2, B.3. The derived STIS transit spectrum is shown along with the entire transit spectrum in Table 3.6.

3.5.2 WFC3

3.5.2.1 Data Reduction

The WFC3 data reduction closely follows the data processing described in detail in Section 2.3. I briefly summarize it here. I download the "ima" data files from the HST archive, and remove background contamination following the "difference reads" methods of Deming et al. (2013), which allows us to easily resolve and remove potential contamination from the nearby companion (Evans et al., 2016). I determine a wavelength solution by taking the zero-point from the F140W photometric

Table 3.6: HAT-P-41b Transit Spectrum

Instrument	λ [μm]	Depth [ppm][a]	Instrument	λ [μm]	Depth [ppm]
STIS G430L[b]	0.290–0.350	10091 ± 230	STIS G750L	0.711–0.731	10499 ± 364
	0.350–0.370	10006 ± 249		0.731–0.750	10038 ± 302
	0.370–0.387	10397 ± 198		0.750–0.770	10142 ± 224
	0.387–0.404	10273 ± 163		0.770–0.799	10124 ± 281
	0.404–0.415	9999 ± 137		0.799–0.819	10618 ± 317
	0.415–0.426	9980 ± 185		0.819–0.838	9910 ± 239
	0.426–0.437	10324 ± 145		0.838–0.884	10252 ± 238
	0.437–0.443	10428 ± 287		0.884–0.930	10217 ± 298
	0.443–0.448	10245 ± 168	WFC3[e]	1.122–1.141	10297 ± 107
	0.448–0.454	10411 ± 165		1.141–1.159	10620 ± 115
	0.454–0.459	10367 ± 192		1.159–1.178	10347 ± 130
	0.459–0.465	10227 ± 145		1.178–1.196	10479 ± 113
	0.465–0.470	10640 ± 164		1.196–1.215	10497 ± 111
	0.470–0.476	10429 ± 165		1.215–1.233	10644 ± 111
	0.476–0.481	10453 ± 144		1.233-1.252	10289 ± 100
	0.481–0.492	10584 ± 125		1.252–1.271	10360 ± 107
	0.492–0.498	10224 ± 192		1.271–1.289	10488 ± 122
	0.498–0.503	10289 ± 166		1.289–1.308	10405 ± 97
	0.503–0.509	10459 ± 149		1.308–1.326	10396 ± 94
	0.509–0.514	10519 ± 183		1.326–1.345	10581 ± 92
	0.514–0.520	10506 ± 168		1.345–1.364	10684 ± 105
	0.520–0.525	10429 ± 186		1.364–1.382	10622 ± 113
	0.525–0.531	10558 ± 128		1.382–1.401	10477 ± 112
	0.531–0.536	10247 ± 174		1.401–1.419	10689 ± 98
	0.536–0.542	10451 ± 180		1.419–1.438	10535 ± 108
	0.542–0.547	10476 ± 177		1.438–1.456	10626 ± 113
	0.547–0.552	10422 ± 206		1.456–1.475	10564 ± 121
	0.552–0.558	10794 ± 153		1.475–1.494	10686 ± 113
	0.558–0.563[c]	9221 ± 279		1.494–1.512	10799 ± 129
	0.563–0.569	10444 ± 188		1.512–1.531	10551 ± 124
STIS G750L[d]	0.526–0.555	10356 ± 220		1.531–1.549	10566 ± 111
	0.555–0.575	10497 ± 278		1.549–1.568	10515 ± 134
	0.575–0.594	10317 ± 202		1.568–1.587	10436 ± 110
	0.594–0.614	10061 ± 236		1.587–1.605	10492 ± 145
	0.614–0.633	10386 ± 142		1.605–1.624	10340 ± 122
	0.633–0.653	10542 ± 219		1.624–1.642	10338 ± 142
	0.653–0.672	10418 ± 276		1.642–1.661	10331 ± 138
	0.672–0.692	10182 ± 225	*Spitzer* IRAC1	3.2–4.0	10191 ± 102
	0.692–0.711	10168 ± 192	*Spitzer* IRAC2	4.0–5.0	10679 ± 145

[a] Transit depth = $R^2_{\text{planet}}/R^2_{\text{star}}$
[b] Typical STIS G430L bin size = 0.0055 μm (median resolution $\sim$ 350)
[c] Outlier bin strongly affected by systematics and ignored in retrieval analyses
[d] Typical STIS G750L bin size = 0.0196 μm (median resolution $\sim$ 130)
[e] WFC3 bin size = 0.0186 μm (median resolution $\sim$ 75)

observation and fitting for the wavelength coefficients that allow an out-of-transit spectrum to match the appropriate ATLAS stellar spectrum (Castelli & Kurucz, 2004).

I then divide the background-subtracted "ima" science frame by the WFC3 flat-field calibration file, and return the dark-, bias-, and flat-field-corrected flux array, in units of electrons. The uncertainty of the flux at each pixel is taken from the "ima" file's error frame, which accounts for gain-adjusted Poisson noise, read noise, and noise from dark current subtraction. This is further adjusted via error propagation for the added uncertainty from background removal and flat-field correction. I use the "ima" file's data quality frame to mask pixels (i.e., give zero weight) that are flagged as bad in every exposure in the time series. I then correct for cosmic rays using a conservative time series sigma-cut of 8-σ, while accounting for changes in flux that occur due to uneven scan rates and the transit itself, and set the affected pixels to the median value of that pixel in the time series. The average amount of pixels either impacted by cosmic ray events or flagged as bad pixels is 1.8% of all pixels in a exposure. The reduced exposures are summed over the spatial scan direction to give a 1-D spectrum at each observation time.

3.5.2.2 Light Curve Analysis

The light curve analysis follows that described in detail in Section 2.4. I briefly summarize it here. I use a similar marginalization light curve analysis as Sheppard et al. (2017), applied to transit curves. This is a Bayesian model averaging method, first described by Gibson (2014b) and applied by Wakeford et al. (2016a), with further de-trending by use of band-integrated (white light) residuals in spectral light curve fitting (Mandell et al., 2013; Haynes et al., 2015). I first analyzed the band-integrated light curve to simplify the spectral light curve analysis, then I fit the spectral light curves to derive the WFC3 transit spectrum.

For the transit model I assume nonlinear limb darkening and derive the coefficients by interpolating the 3-D values from Magic et al. (2015) to the central wavelength of WFC3 ($1.4\,\mu$m). The limb darkening derivation is consistent with that used in the STIS analysis. I only fit for transit depth and central transit time, since the incomplete coverage of HST makes it difficult to improve constraints on other transit parameters, such as a/R_s or inclination. I fix these values in the light curve analyses of each instrument (STIS, WFC3, and *Spitzer*), which ensures consistent orbital parameters are used when analyzing different datasets. The transit and system parameters are shown in Table 3.5.

I fit the each systematic model in the grid, weight each model by its Bayesian evidence — approximated by the Akaike information criterion (AIC) — and marginalize over the model grid to derive the light curve parameters and uncertainties while inherently accounting for uncertainty in model choice. The normalized raw light curve, the de-trended light curve, and the residuals from the highest-weight systematic model are shown in Figure 3.1. I derive the white light depth to be 10490 ± 51 ppm.

To derive the transit spectrum, I bin the 1D spectra from (1.12–$1.66\,\mu$m), and use bins of width $0.0186\,\mu$m (4 pixels) to maximize resolution. I note that the

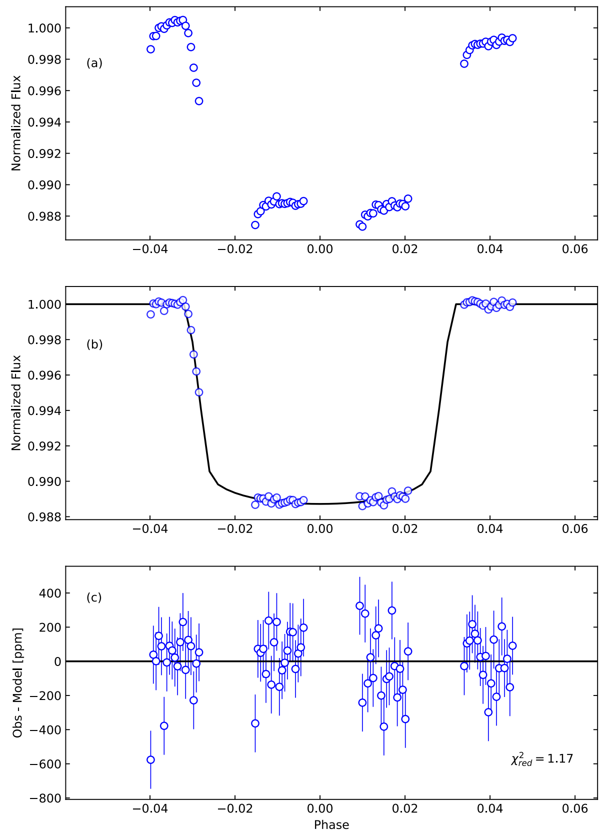


Figure 3.1: Top panel: pre-processed band-integrated light curve for WFC3 observations. This is the band-integrated flux versus planet phase derived from the reduced data. The first orbit is excluded, as it is dominated by instrumental systematics. Middle panel: band-integrated light curve divided by the highest weighted systematic model (i.e., the detrended light curve). Bottom panel: residuals between light curve data and highest-weighted joint transit and systematic model. The reduced χ^2 for the highest weighted model is 1.17, which is consistent with the model being a good fit for 67 degrees of freedom.

atmospheric retrieval is not sensitive to the choice of bin size. The wavelength range, transit depth, an depth uncertainty for each WFC3 bin is shown in Table 3.6.

The shape of my derived spectrum is in excellent agreement with the literature spectrum (Tsiaras et al., 2018), though it is shifted to higher depths by $\sim 90\,\text{ppm}$, indicating that I derive a larger white light depth. This difference persists even if I derive the spectrum without using white light residuals. This could plausibly be due to different limb darkening treatments or different systematic modeling choices. This difference emphasizes the importance of considering offsets between instruments in retrieval analyses (Sec. 3.6.1.2). Further, the white light depth is subject to the choice of orbital parameters, which are typically fixed in spectroscopic light curve fits. I run sensitivity tests to determine that accounting for orbital parameter uncertainties increases the scatter between HST STIS and HST WFC3 by roughly 60 ppm for HAT-P-41b. I capture this scatter by including it in WFC3's depth uncertainty, increasing it from 50 ppm to 80 ppm.

As a check, I performed retrievals using the published spectrum from Tsiaras et al. (2018) in combination with the derived STIS and *Spitzer* depths and found differences well within the 1-σ uncertainties for the retrieved parameters. The major results and conclusions in this chapter are not sensitive to the WFC3 spectrum choice.

I verify the derived spectrum following the process outlines in Section 2.4.3.1. First, I check χ^2. For both the band-integrated and spectral light curves (66 and ~ 88 degrees of freedom, respectively), the acceptable reduced χ^2 range is roughly 0.7–1.3. The band-integrated analysis ($\chi^2_\nu = 1.17$) and all but one of the spectral bins (median $\chi^2_\nu = 0.9$) fall within this range. The only light curve that doesn't fall in this range (1.299 μm) has a reduced χ^2 of 0.56. Like L9859c, this bin is not flagged by the normality or correlated noise analyses, and fitting the light curve without incorporating white light residuals finds a consistent depth with a more reasonable $\chi^2_\nu = 0.71$. Further, the derived depth and uncertainty exhibit good agreement with the published (Tsiaras et al., 2018) transit depth at this wavelength (accounting for the white light offset). Consequently, I include it in the transit spectrum.

Next, I check normality. Normality is rejected at the 5% significance level only for the band-integrated residuals and the 1.243 μm spectral bin residuals. In both cases, normality is ruled out due to a single outlier in the time-series. Normality is rejected at 3% significance for the band-integrated residuals, due entirely to the first exposure in the first orbit. When this exposure is ignored, I recover a consistent depth and uncertainty and the residuals are consistent with normality, and so I keep this exposure in the analysis.

A possible cause of the spectral bin's outlier is a minor cosmic ray event or bad pixel that was small enough to both avoid the detection by the sigma-cut and not affect the band-integrated curve, but large enough to impact the much smaller bin flux. Removing this spectral bin from the retrieval had no noticeable effect on the results. Further, the derived depth and uncertainty are consistent with literature values (Tsiaras et al., 2018). I include this bin in the retrieval.

Finally, I test for correlated noise in the residuals using the time-averaging test as in 2.4.3.1. I find no evidence of correlated noise (Figure 3.2).

I emphasize that removing any of the flagged bin spectra has no affect on the retrieval. Together, these tests support the validity of the derived transit depths and uncertainties in the WFC3 bandpass.

3.5.3 Spitzer

3.5.3.1 Data Reduction

The Spitzer data consists of cubes of 64 subarray frames in each band, each of size 32×32 pixels. A collaborator extracted aperture photometry for each frame, totaling 21,632 frames at both 3.6 and 4.5 μm. To extract photometry, we used 11 numerical apertures with radii ranging from 1.6 to 3.5 pixels, and we centered those

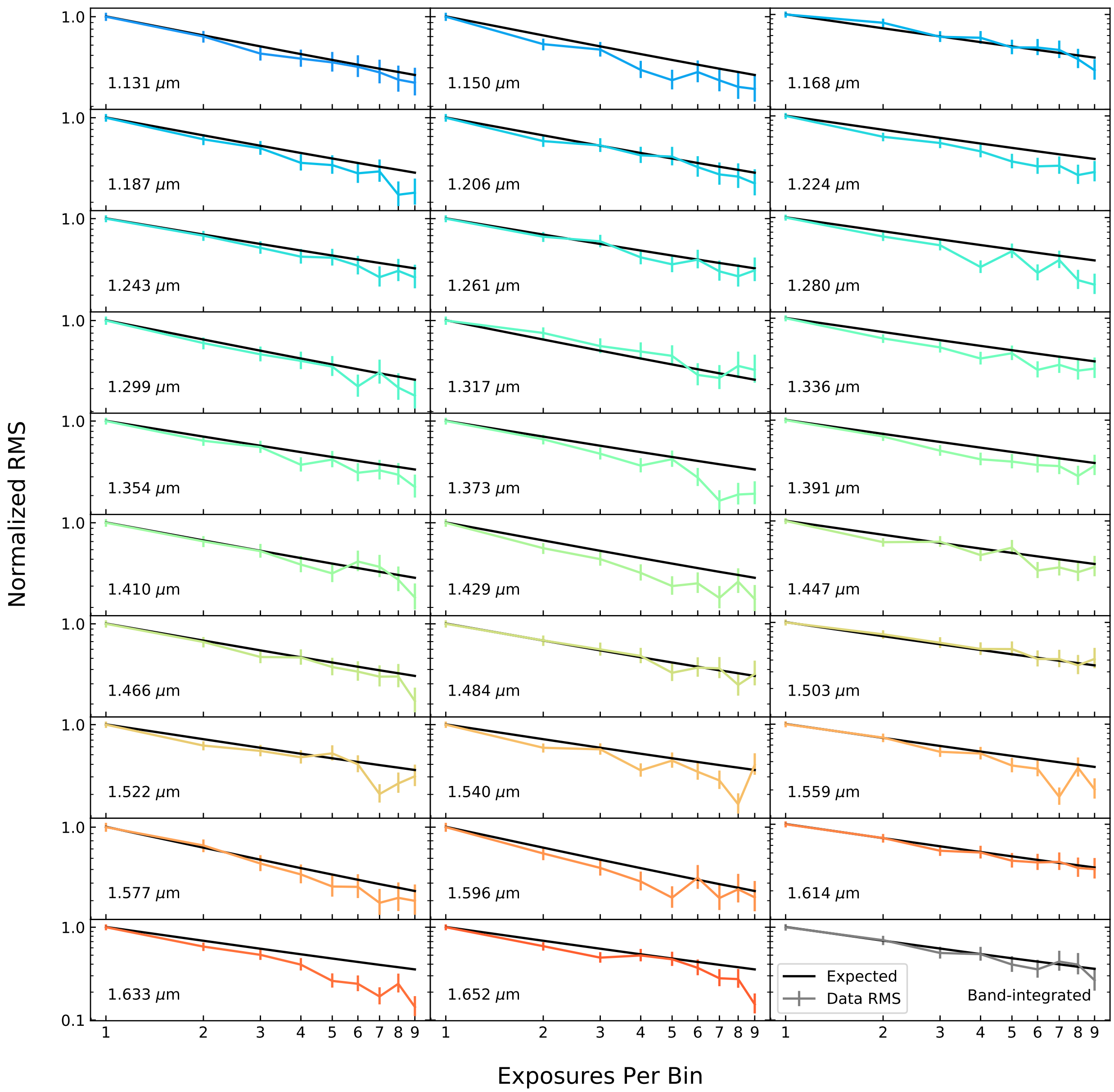


Figure 3.2: Correlated noise analysis for each of the 29 WFC3 spectral bins and the band-integrated light curve. For each bin's light curve, I find the RMS of the residuals for an increasing number of exposures per bin. Pure white noise would scale with the black line, while correlated noise would increase with binning. Though not exact given the gaps between WFC3 data, this is a useful heuristic to search for correlated noise. See Cubillos et al. (2017) and Pont et al. (2006) for more details.

apertures on the position of the host star determined using both a 2-D Gaussian fit to the stellar point spread function, and also a center-of-light calculation. Since HAT-P-41 has a companion star 3.6 arc-sec distant (Hartman et al., 2012; Evans et al., 2016), we measured the flux of the companion scattered into each of our numerical apertures, using the method described by Garhart et al. (2020). We adopted the magnitude difference in the Spitzer bands as deduced by Garhart et al. (2020). Accounting for our different aperture radii than was used by Garhart et al. (2020), we derive dilution correction factors of 1.0171 and 1.0106 at 3.6- and 4.5 μm, respectively. The transit depths are then multiplied by those factors in order to correct for the presence of the companion star.

3.5.3.2 Light Curve Analysis

A collaborator fit transit curves to the 22 sets of photometry at each wavelength (eleven apertures, each with two centering methods). Their default fitting procedure fixes the orbital parameters at the values in the discovery paper by Hartman et al. (2012), fitting only for the central time and depth of the transit. The shape of the Spitzer transits is well matched when fixing the orbital parameters to those values. However, we also explored including the orbital inclination and a/R_s in the fit (see below), those being the orbital parameters that most strongly affect the shape of the transit. We adopt quadratic limb darkening coefficients calculated by least squares for the *Spitzer* bands by Claret et al. (2013), using 2 km/sec microturbulence. We choose the values for $T_{eff} = 6400\,\mathrm{K}$ and $logg = 4.0$, without interpolation. We fix those coefficients in the fitting process, and we deem these choices to be appropriate given that the limb darkening is minimal at these infrared wavelengths. Our fits to the transit account for the intra-pixel sensitivity variations of the Spitzer photometry using pixel-level decorrelation (PLD, Deming et al., 2015), including a linear baseline (ramp) in time. We use a Bayesian information criterion to decide between a linear versus quadratic ramp. The details of the PLD fit are the same as described for secondary eclipses by Garhart et al. (2020), except that we are fitting transits, so we include limb darkening. Briefly, the fitting code bins the photometry and pixel basis vectors to various degrees, and selects the optimal bin size, aperture radius, and centering method, based on the smallest difference from an ideal Allan deviation relation (Allan, 1966). The Allan deviation relation expresses that the standard deviation of the residuals should decrease as the square-root of the bin time. Operating on binned data allows the PLD algorithm to concentrate on the longer time scales that characterize the red noise (and also the transit duration), as opposed to the 0.4-second cadence time of the raw photometry.

We determine the errors on the transit depths and times using an MCMC procedure, with a burn-in phase of 10,000 steps, followed by 800,000 steps to explore parameter space. We calculate multiple chains for each transit, and verify convergence using a Gelman-Rubin (GR) statistic (Gelman & Rubin, 1992). The GR values for the PLD fits are very close to unity, being 1.0027 at 3.6 μm and 1.0004 at 4.5 μm, indicating good convergence. The transit depths and times are listed in Table 3.7.

The derived transit times are in excellent agreement with measurements of the

Table 3.7: HAT-P-41b *Spitzer* Transit Analysis Results

Wavelength	BJD(TDB)	R_p^2/R_s^2 (ppm)[a]
3.6 μm	2457772.20440 $\pm$ 0.00033	10020 $\pm$ 100
4.5 μm	2457788.36860 $\pm$ 0.00032	10568 $\pm$ 135

[a] These are "as observed" transit depths, not corrected for dilution by the companion star. To correct for dilution, multiply the depth by 1.0171 at 3.6 μm, and by 1.0106 at 4.5 μm. The corrected values are shown with the rest of the transit spectrum in Table 3.6.

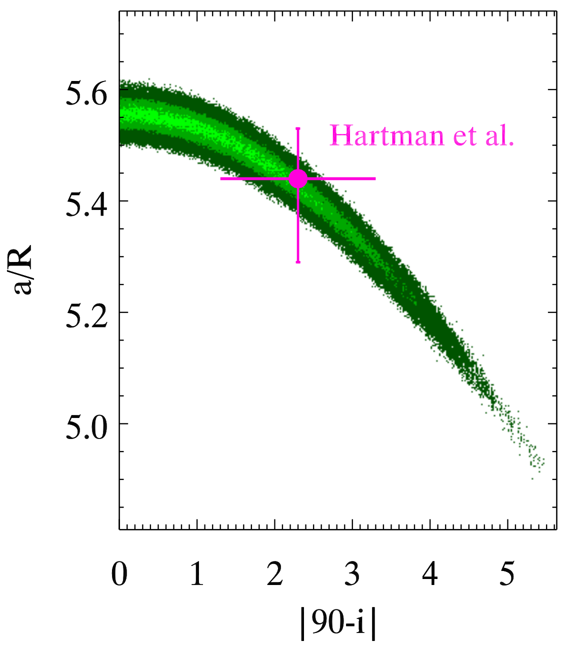


Figure 3.3: Likelihood distribution from the fit to the 3.6 μm transit of HAT-P-41b, based on an MCMC using uniform priors, and shown versus a/R_s and the orbital inclination. These two orbital parameters are degenerate when using only a Spitzer transit, and the values derived by Hartman et al. (2012) are indicated by the point with error ranges. The Spitzer transits at both 3.6- and 4.5 μm are fully consistent with a/R_s and i from Hartman et al. (2012).

same Spitzer transits by Wakeford et al. (2020). Specifically, using our uncertainties, our fitted times differ by 1.1σ and 0.6σ at 3.6- and 4.5 μm, respectively. Wakeford et al. (2020) use 8 HST transits as well as the discovery results and the Spitzer transits to derive a new ephemeris. Our fitted times (Table 3.7) differ from that ephemeris by insignificant amounts (0.2 and 7.8 seconds).

We explored the effect of uncertainties in the orbital parameters, since those can affect the derived transit depth (Alexoudi et al., 2018). Adopting uniform priors for a/R_s and inclination, we find that they are degenerate when fitting only the

Spitzer transits. That degeneracy is illustrated in Figure 3.3, where it is apparent that inclination and a/R_s can trade off to maintain the observed transit duration, and the sharp ingress/egress that characterizes the Spitzer transits. Changing the inclination changes the chord length across the stellar disk, and (when limb-darkening is minimal) that can be compensated by changing a/R_s to maintain the same transit duration. The orbital solution from Hartman et al. (2012) is entirely consistent with the derived likelihood distribution for those parameters, as shown in Figure 3.3 for 3.6 μm (4.5 μm is similar). We therefore freeze the orbital parameters at the Hartman et al. (2012) values when fitting the Spitzer transits.

Figure 3.4 illustrates the transits at 3.6 and 4.5 μm. The residuals from the best-fit model are included in the figure, and the right panel shows the residuals binned over increasing time scales, a so-called Allan deviation relation (Allan, 1966). The slopes of those relations are close to the -0.5 value expected for photon noise.

3.6 Atmospheric Retrieval

There are two common frameworks to retrieve physical parameters from transmission spectra. The first is by assuming chemical equilibrium, where the abundance of a molecule is only dependent on local temperature, local pressure, and global elemental abundances such as O/H and C/H (e.g., Kreidberg et al., 2015). The second is by instead fitting for molecular abundances based on observed spectral features, then determining global elemental abundances from the molecular abundance values (e.g., Madhusudhan, 2012). Since carbon and oxygen-based molecules are typically

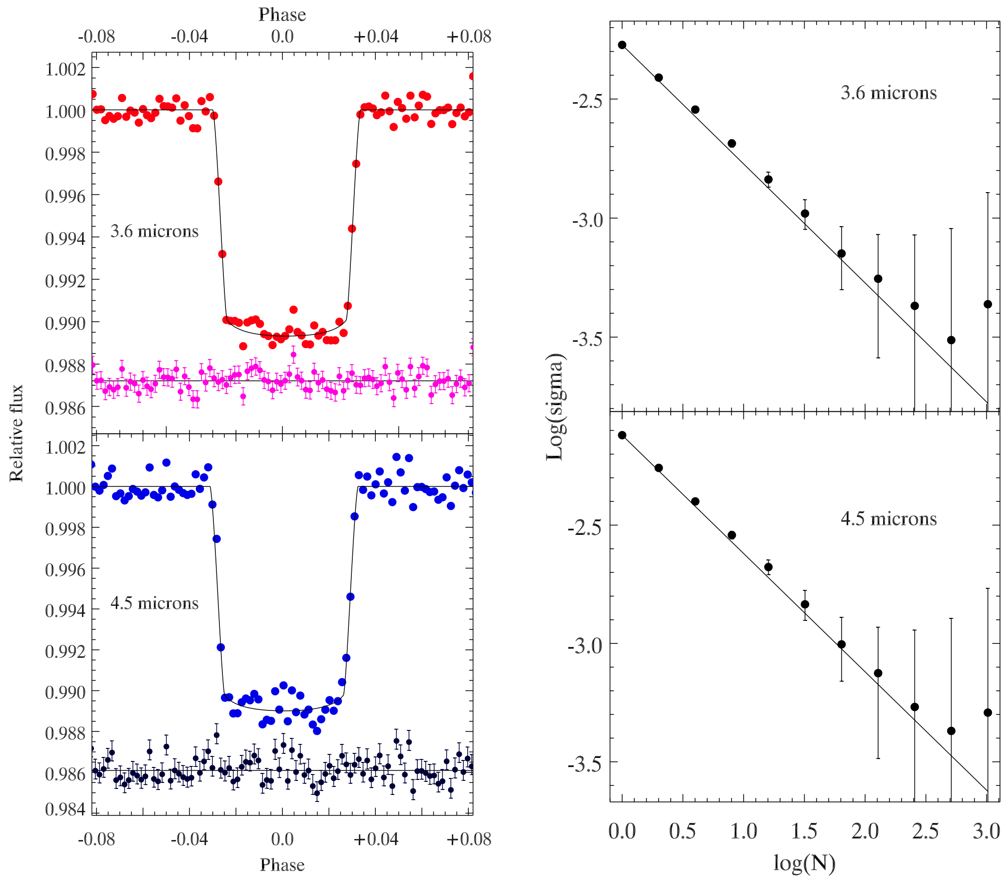


Figure 3.4: Left: Spitzer transit light curves of HAT-P-41b at 3.6 and 4.5 μm after correction of the intra-pixel effects of the detector and temporal ramps. The data are binned to 100 points per transit for clarity of illustration. The residuals (data minus fitted model) are shown below the transit curves, and have error bars added. Right: Allan deviation relations for the binned residuals, i.e. standard deviation of the residuals when the original data are binned over an increasing number of points, N. The solid lines project the single-point ($N = 1$) scatter to larger bin sizes with a slope of -0.5 as expected for photon noise.

the most spectroscopically active species over the wavelengths covered by HST and *Spitzer*, the elemental abundances are commonly parameterized by metallicity — defined as the enhancement of metal elements relative to hydrogen compared to solar values (see Sec 3.6.3 for more detail) — and carbon-to-oxygen ratio (C/O). Some retrievals improve flexibility by allowing other elements — such as sodium or vanadium — to also vary from their solar ratios (Amundsen et al., 2014; Tremblin et al., 2015, 2016; Sing et al., 2016).

The open source code `PLATON`[1] (Zhang et al., 2019) is able to perform quick retrievals which assume chemical equilibrium, whereas `AURA` (Pinhas et al., 2018) is able to capture possible disequilibrium chemistry by not assuming chemical equilibrium. We retrieve the atmospheric parameters with both frameworks to see how interpretations compare, and to explore how sensitive the conclusions are to retrieval assumptions.

3.6.1 PLATON

PLATON is a fast, open-source retrieval code developed by Zhang et al. (2019). Like many retrieval codes, it comprises a forward model and an algorithm for Bayesian inference. Though there are some differences, it essentially uses the same forward model as Exo-Transmit (Kempton et al., 2017). Here, I summarize the forward model: To calculate a spectrum, it first determines the abundances of 34 potentially relevant chemical species for a given atmospheric metallicity and C/O. These include Na and K as well as molecules CH_4, CO, CO_2, HCN, H_2O, MgH, NH_3,

[1] `https://github.com/ideasrule/platon`

TiO, and VO (see Kempton et al. (2017) for complete list). The metallicity and C/O provide elemental abundances, which are combined with a temperature-pressure grid as input into GGchem (Woitke et al., 2018) to compute equilibrium molecular and atomic abundances at each pressure layer in the atmosphere, accounting for the effects of condensation on equilibrium abundances. PLATON allows for a grey cloud deck, below which no light can penetrate, and the abundance-temperature-pressure grid facilitates the determination of total opacity at each pressure layer in the atmosphere which lies above this cloud top. PLATON includes the same opacity sources as Exo-Transmit, and accounts for opacity from gas absorption, H_2-He collision-induced absorption (CIA), and scattering (either parametric Rayleigh scattering or Mie scattering). The forward model converts the opacity-pressure grid to a opacity-height grid using hydrostatic equilibrium with $P_{ref} = 1$ bar, which is then used as an input to a radiative transfer code to determine the uncorrected transit depths. After correcting for possible stellar activity, due to either unocculted spots or faculae, PLATON's forward model outputs the corrected transit spectrum. The largest source of computational uncertainty is opacity sampling error, which is a source of white noise from using a relatively low resolution (R=1000) that cannot resolve individual lines (Zhang et al., 2019). Accounting for opacity sampling for the transit spectrum of HAT-P-41b typically increases the depth uncertainty by 1.5% (2.5ppm), which is sufficiently small such that it does not affect interpretation. For more details on PLATON, see Zhang et al. (2019). I note that the version of PLATON I describe and use in this chapter is Platon 3.1. A newer version, PLATON 5.1, has since been released with additional features as described in Zhang et al. (2020a).

This newer version is used is Section 2.5.1.

My PLATON analysis does not retrieve individual abundances. Instead, it fits for the isothermal temperature, atmospheric metallicity as a multiple of solar values for atomic species, and C/O ratio; the retrieved equilibrium abundances for atomic and molecular absorbers are a natural consequence of those values. This is in contrast with the AURA analysis (Sec. 3.6.2), which retrieves individual molecular and atomic abundances.

The algorithm PLATON uses for Bayesian inference is nested sampling (Skilling, 2004). Specifically, PLATON uses multimodal nested sampling from the Python implementation `nestle`.[2] Like MCMC samplers, nested sampling efficiently samples posterior distributions with dimensionalities typical of atmospheric retrievals (n=5–20), and so it is effective at atmospheric parameter estimation. Unlike MCMC routines, it automatically calculates the Bayesian evidence for a model, which is necessary for model comparison. The Bayesian evidence intrinsically accounts for overfitting by punishing too much model structure and thus determines if extra parameters are warranted. I use this to justify excluding parameters which add structure and do not significantly improve the fit. Nested sampling also has a well-defined stopping criteria, so there is no need to check for convergence. For an excellent write up on this algorithm, especially about using it in practice, see the documentation of `Dynesty`[3] (Higson et al., 2019).

In addition to the standard set included in PLATON's forward model, I added

[2] `https://github.com/kbarbary/nestle`
[3] https://dynesty.readthedocs.io/en/latest/

four new fittable parameters: a partial cloud parameterization and three instrumental transit depth bias parameters (henceforth referred to as instrumental offsets).

3.6.1.1 Partial Clouds

The partial cloud parameter is motivated by work by Line & Parmentier (2016) and MacDonald & Madhusudhan (2017), which showed that if the grey cloud deck were inhomogenous, then the spectrum I observe (D) would be a weighted average of the clear atmosphere transit spectrum (D_{clear}) and the cloudy atmosphere spectrum (D_{cloudy}) with weights given by the cloud fraction (f_c). I implement this as $D = f_c * D_{cloudy} + (1 - f_c) * D_{clear}$. Since high altitude grey clouds are seen in spectra as flat lines, averaging a spectrum with features with this line will shrink the features and can mimic the effect of a high mean-molecular mass, small scale height atmosphere. Thus, including this parameter allows us to account for this possible degeneracy and prevents us from overconfidently claiming a high-metallicity atmosphere.

3.6.1.2 Instrumental Offsets

The instrumental offsets are nuisance parameters that can capture the extent to which transit depths from STIS G430L, STIS G750L, or WFC3 are biased relative to the depths from the other instruments. This is motivated by the use of common-mode corrections in the light curve analysis of each instrument, which can potentially introduce a uniform bias for the depth at each spectral bin for that instrument. Offsets are also able to account for inter-instrumental transit depth scatter introduced by uncertainty in orbital parameters a/R_s and inclination (Section 3.5.2.2).

I explore three offset scenarios. The first is physically-motivated. In this scenario, I use Gaussian priors with sigmas determined by the uncertainties in the band-integrated transit depths to try to reflect the correlated uncertainty that exists between spectral bins for each instrument, whether due to common mode corrections or orbital parameter uncertainties. This essentially propagates the white light depth uncertainty into the retrieval. Derived in Section 3.5, these uncertainties are 105 ppm, 85 ppm, and 80 ppm for STIS G430L, STIS G750L, and WCF3, respectively. The second scenario investigates the potential impact of unknown sources of bias by setting a large, uniform offset prior for each instrument's offset. The third scenario extends this, by setting two large, uniform priors: a WFC3 offset and a single offset for both STIS instruments. The third scenario allows the absolute depths at STIS to vary while preserving the optical spectral shape.

I caution that offsets — especially penalty-free, uniform prior offsets — can cloak missing physics in a model. I do not think they should act as a safety net to achieve a good fit to a spectrum, and the inferred atmospheric properties should be understood in context. However, offsets offer a way to both investigate potential instrumental biases and account for absolute depth uncertainty for each instrument. It is valuable to include offsets as model parameters and marginalize over these possible values in order to understand how the uncertainty in the absolute transit depth for each instrument affects the marginalized posterior distributions of the other model parameters.

3.6.2 AURA

A collaborator complements my analysis of HAT-P-41b by performing retrievals on the STIS, WFC3 and *Spitzer* observations without the assumption of chemical equilibrium. They employ an adaptation of the retrieval code AURA (Pinhas et al., 2018), as described in (e.g., Welbanks & Madhusudhan, 2019).

The code computes line by line radiative transfer in a transmission geometry and assumes hydrostatic equilibrium. We consider a one-dimensional model atmosphere consisting of 100 layers uniformly spaced in $\log_{10}(P)$ from 10^{-6}-10^{2} bar. The pressure-temperature (P-T) profile in the atmosphere is retrieved using the P-T parameterization of Madhusudhan & Seager (2009). The measured radius of the planet R_p is assigned to a reference pressure level in the atmosphere through a free parameter P_{ref}.

The model atmosphere assumes uniform mixing ratios for the chemical species and treats them as free parameters. We consider sources of opacity expected to be present in hot Jupiter atmospheres (e.g., Madhusudhan, 2012) and include H_2O (Rothman et al., 2010), Na (Allard et al., 2019), K (Allard et al., 2016), CH_4 (Yurchenko & Tennyson, 2014), NH_3 (Yurchenko et al., 2011), HCN (Barber et al., 2014), CO (Rothman et al., 2010), CO_2 (Rothman et al., 2010), TiO (Schwenke, 1998), AlO (Patrascu et al., 2015), VO (McKemmish et al., 2016), and H_2-H_2 and H_2-He collision induced absorption (CIA; Richard et al., 2012). The opacities for the chemical species are computed following the methods of Gandhi & Madhusudhan (2017). The CO_2 abundance is restricted to remain below the H_2O and CO abundances as expected at these temperatures for H-rich atmospheres (Madhusudhan, 2012).

We allow for the presence of clouds and/or hazes following the parameterization in Line & Parmentier (2016); MacDonald & Madhusudhan (2017). Non-homogenous cloud coverage is considered through the parameter $\bar{\phi}$, corresponding to the fraction of cloud cover at the terminator. Hazes are incorporated as $\sigma = a\sigma_0(\lambda/\lambda_0)^\gamma$, where γ is the scattering slope, a is the Rayleigh-enhancement factor, and σ_0 is the H_2 Rayleigh scattering cross-section (5.31×10^{-31} m^2) at the reference wavelength $\lambda_0 = 350$ nm. Opaque regions of the atmosphere due to clouds are included through an opaque (gray) cloud deck with cloud-top pressure $P_{\rm cloud}$.

Lastly, we allow for the same three instrumental offset scenarios as described in Section 3.6.1.2. In these model runs, a constant offset in transit depth is applied to the data set of choice. The offset priors for each scenario are given in Table 3.8.

In summary, the AURA retrievals consist of up to 25 parameters: 11 chemical species, 6 parameters for the P-T profile, 1 for the reference pressure, 4 for clouds and hazes, and up to 3 extra parameters for instrumental shifts. Table 3.8 shows the parameters and priors used in the AURA retrievals.

3.6.3 A note on metallicity and C/O

Atmospheric metallicity is a broadly used term that does not always have the same definition or assumptions built into its derivation (Madhusudhan et al., 2014c; Kreidberg et al., 2014b). Here I define C/O and global atmospheric metallicity ex-

Table 3.8: Parameters and priors used in the AURA retrievals.

Description	Parameter	Symbol	Prior Distribution
General	Fractional Abundances	X_i	$\mathcal{LU}(-12,-1)$
	Reference Pressure[a]	P_{ref}	$\mathcal{LU}(-6,2)$
T-P Profile[a]	Zero-point Temperature	T_0	$\mathcal{U}(800,2000)$ K
	Gradients	$\alpha_{1,2}$	$\mathcal{U}(0.02,2.00)$ $K^{-1/2}$
	Gradient Pressures	$P_{1,2}$	$\mathcal{LU}(-6,2)$ bar
	Isothermal Pressure	P_3	$\mathcal{LU}(-2,2)$ bar
Cloud/Haze	Scattering Factor	a	$\mathcal{LU}(-4,10)$
	Scattering slope	γ	$\mathcal{U}(-20,2)$
	Cloudtop pressure	P_{cloud}	$\mathcal{LU}(-6,2)$ bar
	Cloud fraction	$\bar{\phi}$	$\mathcal{U}(0,1)$
Instrument Offsets[b]	STIS	-	$\mathcal{U}(-500,500)$ ppm
	STIS G430L	-	$\mathcal{U}(-500,500)/\mathcal{N}(0,105)$
	STIS G750L	-	$\mathcal{U}(-500,500)/\mathcal{N}(0,85)$
	WFC3	-	$\mathcal{U}(-500,500)/\mathcal{N}(0,80)$

[a] Section 3.6.2 defines the temperature-pressure profile parameters.

[b] Instrumental offsets were employed in a subset of the retrievals and had either uniform or Gaussian priors as explained in section 3.9.2.

plicitly. For AURA, the abundances of different elements are derived independently from the corresponding gaseous absorbers - O/H from oxygen-bearing molecules such as H_2O, Na from gaseous Na, and so on. This retrieval approach allows for different elements to be enhanced or depleted in different quantities. As such, there is no single metric for metallicity in this approach. Nonetheless, as described below, I use the O/H ratio from the retrieved H_2O abundance using AURA as a proxy for metallicity in order to facilitate comparisons with PLATON retrievals in which all the elements are enhanced by a single metallicity parameter.

In PLATON, metallicity is a factor that scales the solar elemental abundances, denoted M/H. The ratios between metals (e.g., Fe/O, Ti/V) are fixed to solar metal ratios, but the solar metal-to-hydrogen ratio is allowed to vary. Thus, all metals are scaled by the same factor. Then, the elemental carbon abundance is determined by C/O $\times$ metallicity. Therefore, only carbon is allowed to differ from its solar ratio compared to other metals. While allowing carbon instead of oxygen to vary is arbitrary, C/O variation is motivated: it is the only metal-to-metal ratio for which predicted molecular abundances are typically sensitive to in the wavelength range covered by HST/*Spitzer* transit spectra. In this paradigm, metallicity is effectively a heuristic for O/H, since that dominates the retrieval both because of molecular opacity (e.g., H_2O, TiO, VO) and because over 99% of the mean molecular weight is due to C, O, and H. In reality I retrieve O/H and C/O, then set all other elements to $X = X_{\odot} * \frac{O/H}{O/H_{\odot}}$. Therefore, PLATON's derived metallicity is reasonably comparable to AURA's derived O/H.

3.7 PLATON Retrieval Analysis

The relative importance of each physical process that affects an observed transit spectrum is not clear ahead of time. PLATON, though less flexible than free-chemistry retrievals or retrievals that allow elemental abundances to vary from solar ratios, is able to quickly perform chemically-constrained retrievals (~30 minute run-time for fiducial model retrieval on full data set). This makes it well suited for testing an array of models, which is important in order to determine how different model assumptions impact the conclusions of a retrieval. To explore this, I choose the fiducial model to be the set of parameters necessary to fully describe the simplest physical processes that I know affect the spectrum: opacity from gas absorption, CIA, and Rayleigh scattering. I then detail the effect of incorporating more complicated physics. In Section 3.7.3, I use both physical and statistical arguments to determine the "best" model. However, it is important to show the sensitivity of the results to each model assumption to not provide overconfident constraints and to be able to predict how new observations might affect the conclusions.

3.7.1 Fiducial Model

The priors for the fiducial model are shown in Table 3.9. The metallicity of the atmosphere, the temperature of the limb, and the C/O ratio are necessary to include in order to calculate molecular and atomic abundances, which determine the gas absorption, CIA, and Rayleigh scattering opacities. While I cannot improve constraints on M_p and R_s, it is best practice to include them as parameters

Table 3.9: Priors for parameters used in all PLATON retrievals.

Parameter	Symbol	Prior Distribution	Default Value
Planet Radius	R_p	$\mathcal{U}(0.83, 2.48)$[a]	1.65 R_{Jup}
Limb Temperature	T	$\mathcal{U}(850, 2550)$[a]	1700K
Carbon-oxygen ratio	C/O	$\mathcal{U}(0.2, 2.0)$	0.53[b]
Metallicity	Z	$\mathcal{LU}(-1, 3)$	1 $Z_\odot$
Planet Mass	M_p	$\mathcal{N}(0.76, 0.14)$	$0.76 M_{Jup}$
Stellar Radius	R_s	$\mathcal{N}(1.65, 0.08)$	1.65 $R_\odot$
Cloudtop Pressure	P_{cloud}	$\mathcal{LU}(-3, 8)$	1 Pa

[a] Range is 50–150% of the default value.
[b] Solar C/O

with Gaussian priors in order to propagate the uncertainties on those measurements (Zhang et al., 2019). Otherwise, I would mistakenly assume that M_p and R_s are precisely known. I also include the cloud top pressure of a grey cloud deck in the fiducial model. I fix the reference pressure to 1 bar and retrieve the planetary radius at that pressure. Note that Welbanks & Madhusudhan (2019) demonstrated that it is justified to assume a reference pressure and retrieve the planetary radius without affecting the ability to constrain atmospheric composition.

Although R_p and M_p are both allowed to vary independently in PLATON retrievals, their uncertainties are not independent: the uncertainties for M_p are derived from $\log(g_p)$ (from transit observations) and R_p (derived from R_p/R_s from transit and R_s from TIC-8). PLATON does not constrain M_p and R_p^2 to match $\log(g_p)$ a priori, however this is only an issue if regions of high likelihood extend to combinations of values that should not be allowed (i.e., more than 3-σ from observed $\log(g_p)$). I re-derive $\log(g_p)$ using values at R_p and M_p at the edge of significant likelihood and find good agreement with the prior, well within 3-σ).

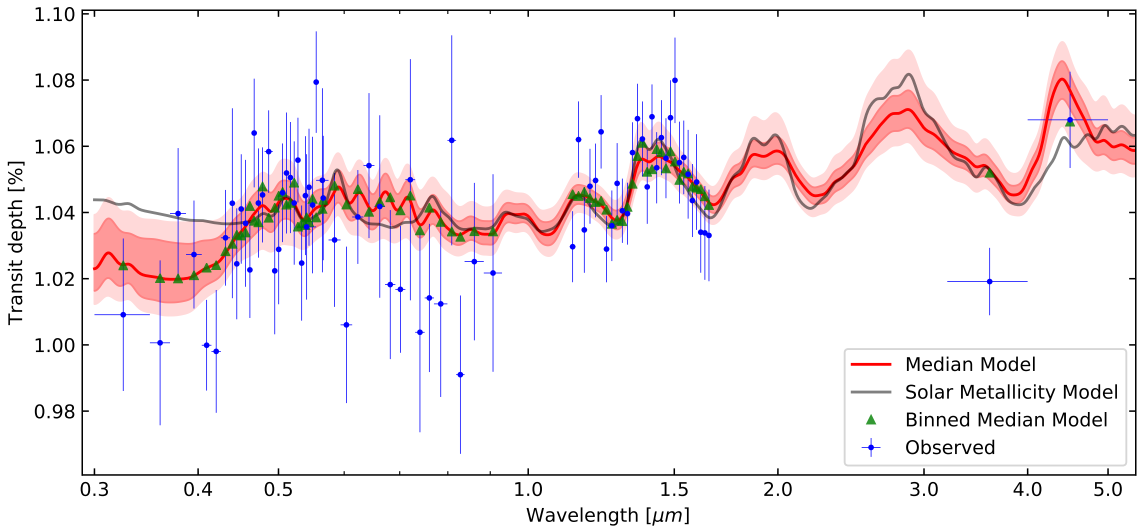


Figure 3.5: Median retrieved model with 1-σ and 2-σ uncertainty contours for the fiducial model. Also shown is the median retrieved model when metallicity is fixed to solar. The median model and uncertainties are derived by generating 100 samples from the correctly-weighted posterior and calculating the depth at each bin for each sample. The contours are given by the 2nd, 16th, 50th, 84th, and 98th-percentile depths at each bin. The continuous model is smoothed with a Gaussian filter with $\sigma = 15$, which approximates the resolution of HST WFC3 (Zhang et al., 2019).

I use uninformative priors where appropriate in order to fully explore the possible parameter space. For quantities that can range over many orders of magnitude, such as the cloud top pressure or the metallicity, this means a log-uniform prior is necessary to avoid bias towards higher values. Otherwise — for T_{limb}, R_p, and C/O — I use uniform priors with limits either set by the functionality of the code (e.g., C/O) or conservatively derived from previous observations. Widening the prior for any parameter in the fiducial model has no significant effect on the result of the retrieval.

The retrieved median fiducial model with uncertainty contours is shown with the observed spectrum in Fig 3.5. The model is an excellent fit (reduced χ^2=1.09;

consistent with the χ^2 of the true model for 70 degrees of freedom to 1-σ). I clearly detect water vapor (>5-σ significance) via the 1.4μm water band in the WFC3 data. The bump in the STIS data is indicative of TiO, and the lack of any optical slope or flat-line indicates that grey clouds and scattering haze do not contribute significant opacity in the planet's spectroscopically active region. The difference between the two *Spitzer* points is attributed to CO_2, though since they are photometric observations I do not resolve any feature.

The posterior probability distribution is represented by the corner plot (Foreman-Mackey, 2016)[4] in Fig 3.6. This figure provides the marginalized posterior distribution for each parameter (with median, 16th, and 84th percentile values indicated by vertical dashed lines), as well as every two-dimensional projection of the posterior (with 0.5-, 1-, 1.5, and 2-σ contours) to reveal covariances. Well-constrained parameters have narrow distributions with clear peaks, and slanted or diagonal shapes are indicative of correlated sets of parameters (e.g., the R_p-R_s shape). I find a super-solar metallicity ($259^{+174}_{-114}\times$ solar metallicity, henceforth $Z_\odot$), a likely sub-solar C/O (C/O < 0.6) that is consistent with stellar C/O (0.19), a clear atmosphere ($P_{cloud} > 0.5$mBar), and $T_{limb} = 1650^{+70}_{-120}$K. The temperature is driven primarily by the STIS data, mostly because PLATON interprets the bump in the STIS data as a metallic oxide, which is only the dominant opacity source above ~1500K in chemical equilibrium. Below $\sim$ 1500K the optical spectrum would be dominated by an atomic sodium line, and this is not seen in the data. The upper limit on C/O is related mainly to the H_2O: in equilibrium chemistry for T~1650K and P~1 bar,

[4] `https://github.com/dfm/corner.py`

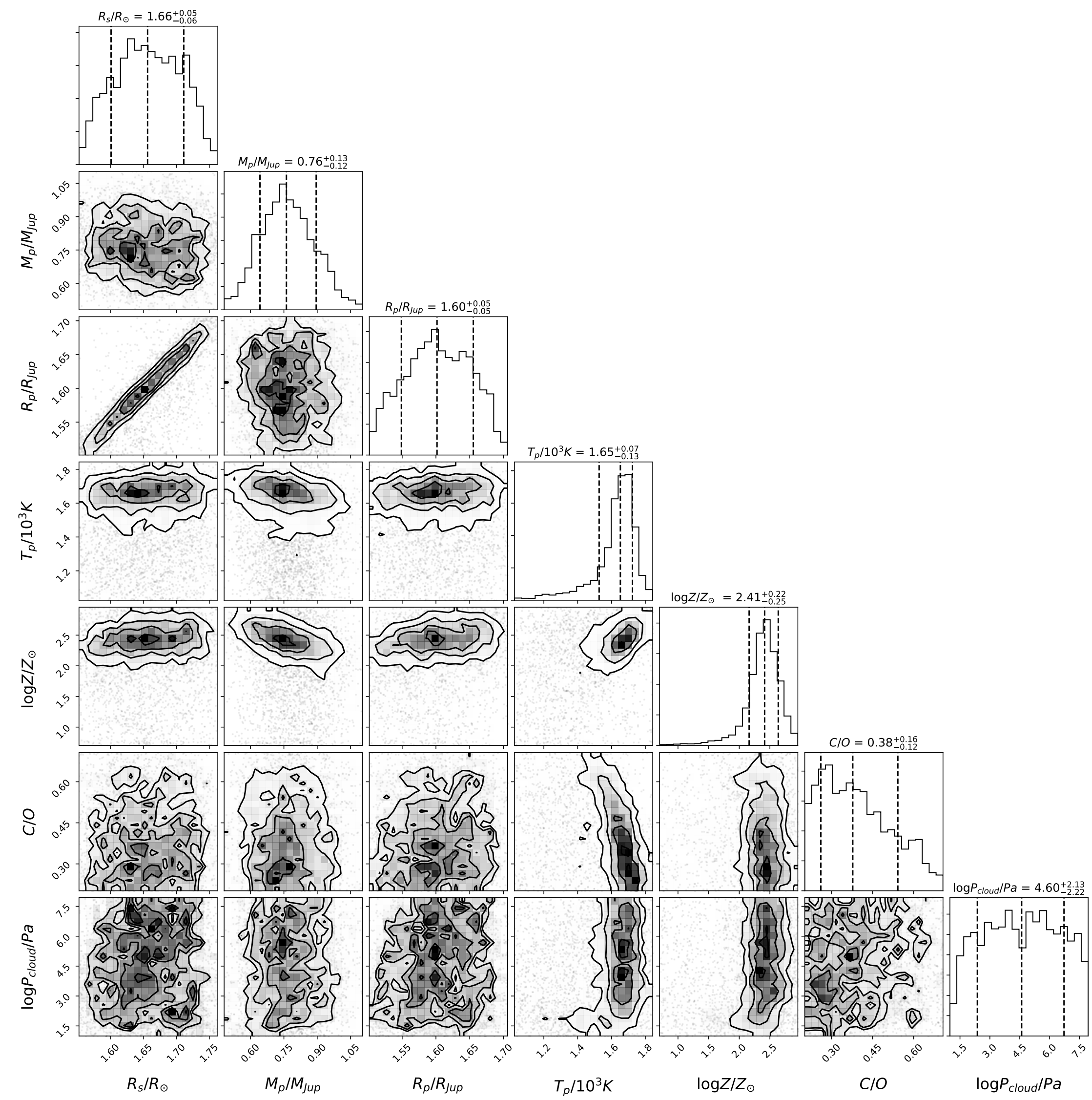


Figure 3.6: Corner plot illustrating posterior probability distributions from PLATON for the fiducial model. The 16th, 50th, and 84th percentile values are indicated by vertical dashed lines and stated in the title of each parameter's 1D marginalized posterior distribution. The contours indicate the joint 0.5-, 1-,1.5-, and $2-\sigma$ levels for each 2D distribution. The 1-σ metallicity range is $\log_{10} Z/Z_\odot = 2.41^{+0.23}_{-0.25}$ (145–437× solar metallicity), the isothermal limb temperature is well constrained around 1650 K, and the C/O ratio is likely subsolar. The spectrum is consistent with a cloud-free atmosphere, and the marginalized posterior distributions of M_p and R_s are dominated by their priors. The slight correlations between T–$\log Z$ and M_p–$\log Z$ are due to their relation in the scale height equation.

H_2O opacity decreases exponentially when C/O > 0.6 (Madhusudhan, 2012). Any model with C/O > 0.6 struggles to capture the water feature and relatively high infrared baseline opacity (compared to the optical) and is thus a poor fit to the data.

The high metallicity is constrained by the size of both the STIS and WFC3 features, as well as the lack of a significant Rayleigh scattering slope. While the metallicity affects chemistry, it is primarily constrained via its effect on the mean molecular mass of the atmosphere. Increasing metallicity increases the ratio of metals to hydrogen by definition, which increases the atmosphere's mean molecular mass. This lowers the atmospheric scale height and consequently decreases the predicted feature size. The equation for approximate feature size (δ_λ; Kreidberg, 2018), where μ is the mean molecular mass, clarifies its dependencies:

$$\delta_\lambda \propto \frac{TR_p}{\mu g_p R_s^2} \propto \frac{TR_p^3}{\mu M_p R_s^2} \tag{3.1}$$

PLATON can decrease the feature size by changing R_p or R_s, but both are well-constrained by the continuum baseline as well as priors and thus are relatively fixed. It can also be lowered by decreasing the temperature, increasing the metallicity (and thus the mean molecular weight), or by increasing M_p. The temperature is strongly constrained by chemistry, and M_p is constrained by previous observations, so only metallicity can vary enough to explain the observed feature sizes. Note that this relationship explains the correlations between M_p, T_p, and metallicity seen in Fig 3.6: as mass increases or temperature decreases, metallicity decreases since a lower value is necessary to achieve the scale height that predicts the observed feature sizes. For reference, the median derived mean molecular weight is 5.8 AMU and the derived scale height is roughly 322km.

At solar metallicity, the model predicts features that are much larger than what the data shows. Consequently, solar metallicity atmospheres in the fiducial model can only explain the observed feature by invoking a cloud to mute the troughs of the features. The median retrieved model for metallicity fixed to solar is shown in Figure 3.5. Note that this model is dispreferred by 3.5-σ, since the cloud leads to a poor fit to the bluest transit depths.

The same metallicity value is an excellent fit for both the water feature in the WFC3 data and the TiO feature in the STIS data. Retrieving only on the STIS data or only on the WFC3 data recovers supersolar atmospheric metallicities. Additionally, due to predicting a greater abundance of CO_2, it is better than low-metallicity solutions at explaining the large change in depth between the *Spitzer* points. Observations from all three instruments support the high metallicity solution.

3.7.2 More Complex Models

In this section I incorporate additional model parameters to explore if more complex physics impacts the inferred atmospheric parameters. I demonstrate the insensitivity of my results to model assumptions.

3.7.2.1 Partial Cloud Coverage

Line & Parmentier (2016) showed that partial cloud coverage (i.e., clouds at the same height but not uniformly covering the limb azimuthally) could mimic the effect of a high mean molecular mass atmosphere for WFC3 spectra. When partial clouds are present, the observed spectrum would be the weighted average of the cloudy and clear spectra. The transit depth of a grey-cloud dominated atmosphere does not vary with wavelength, and so the cloudy spectrum is a straight line. Averaging a clear spectrum with molecular features and a cloudy, flat spectrum reduces the size of the features by an amount proportional to the cloud fraction. Given that I find a significantly supersolar metallicity, I investigate if this possible mean molecular mass-cloud fraction degeneracy affects the results from the fiducial case.

When fit independently and allowing the cloud fraction to vary, both WFC3 and STIS spectra retrievals no longer constrain the metallicity to be supersolar. However, fitting the infrared and optical data jointly breaks this degeneracy, as predicted by Line & Parmentier (2016). Effectively, a low-mean molecular mass and non-uniform cloud solution should be impacted by Rayleigh scattering in the optical data, especially the bluest six wavelength bins. The dominance of gas absorption opacity over Rayleigh scattering opacity in the STIS data disallows this solution, breaking the degeneracy in favor of the high mean molecular mass solution.

However, it is possible that removing assumptions made in the fiducal model — such as fixed Rayleigh scattering or no instrumental offsets — could muddle this decisive degeneracy break and allow for a low-metallicity solution. I investigate this

Table 3.10: Priors for parameters used in more complicated PLATON models.

Parameter	Symbol	Distribution	Default
Cloud fraction	f_c	$\mathcal{U}(0,1)$	1
Scattering slope	γ	$\mathcal{U}(-2,20)$	4
Scattering factor	a_0	$\mathcal{LU}(-4,8)$	1
WFC3 offset	-	$\mathcal{U}(-500,500)/\mathcal{N}(0,80)$	0 ppm
STIS G430L offset	-	$\mathcal{U}(-500,500)/\mathcal{N}(0,105)$	0 ppm
STIS G750L offset	-	$\mathcal{U}(-500,500)/\mathcal{N}(0,85)$	0 ppm
STIS offset	-	$\mathcal{U}(-500,500)$	0 ppm
Stellar Effective Temperature	$T_{\rm star}$	Fixed	6480 K
Faculae Temperature	$T_{\rm fac}$	Fixed[a]	6580 K
Faculae covering fraction	$f_{\rm fac}$	$\mathcal{U}(0,0.1)$	0
Mie Particle Size	$r_{\rm part}$	$\mathcal{LU}(-2,0)$ μm	0.1 μm
Mie Number Density	n	$\mathcal{LU}(1,15)$ m^{-3}	10^5 m^{-3}
Fractional Scale Height	$\frac{H_{\rm cloud}}{H_{\rm gas}}$	$\mathcal{LU}(-1,1)$	1.0

[a] $T_{fac} = T_{star} + 100K$ (Rackham et al., 2019)

below.

3.7.2.2 Parametric Rayleigh Scattering

The fiducial model assumes Rayleigh scattering. In lieu of complicated microphysics, PLATON allows parametric scattering, in which the slope and the magnitude of Rayleigh scattering vary in order to capture the possible signature of many hazes. For a more detailed explanation, see Zhang et al. (2019). Though there is no obvious signature of haze in the optical data (i.e., no linear slope decreasing with increasing wavelength), it is worth exploring if loosening the assumption of exact Rayleigh scattering affects the results.

Allowing the full scattering parameter space (see Table 3.10) has little effect: the clear lack of slope in the STIS data conclusively leads to a haze-free atmosphere. Further, the median scattering factor is 0.01, implying that the data is easiest to fit when opacity from Rayleigh scattering is muted. This complicates the mean molecular weight-cloud fraction ($\mu - f_c$) degeneracy. Lower values of μ are now possible, since the model no longer expects scattering opacity to be important at optical wavelengths. The lower the magnitude of Rayleigh scattering — and the shallower the scattering slope — the lower μ can be. This is because decreases in Rayleigh opacity allow for gas absorption to still be dominant at larger scale heights. As a result, a patchy cloud and low metallicity solution is viable. Though possible, the low μ solution requires a specific combination of cloud top pressure, scattering slope, scattering factor, and cloud fraction, and does not improve the fit. Therefore, it is much less likely than the high metallicity solution. The marginalized posterior probability distribution for metallicity has the same maximum likelihood value as the fiducial model. The difference is that the distribution has a tail extending to lower metallicities (Fig 3.7). The resulting median log metallicity and 1-σ range (as determined by the 16th and 84th percentile values) is $\log_{10} Z/Z_{\odot} = 2.34^{+0.27}_{-0.64}$.

Though all cloud fractions and cloud pressures are allowed, the posterior is consistent with a clear atmosphere due to the likelihood-desert in the upper-left corner of the cloud fraction-cloud top pressure pairs plot: clouds are only seen above the altitude corresponding to the ~ 10 Pa pressure level at fractions below 0.50.

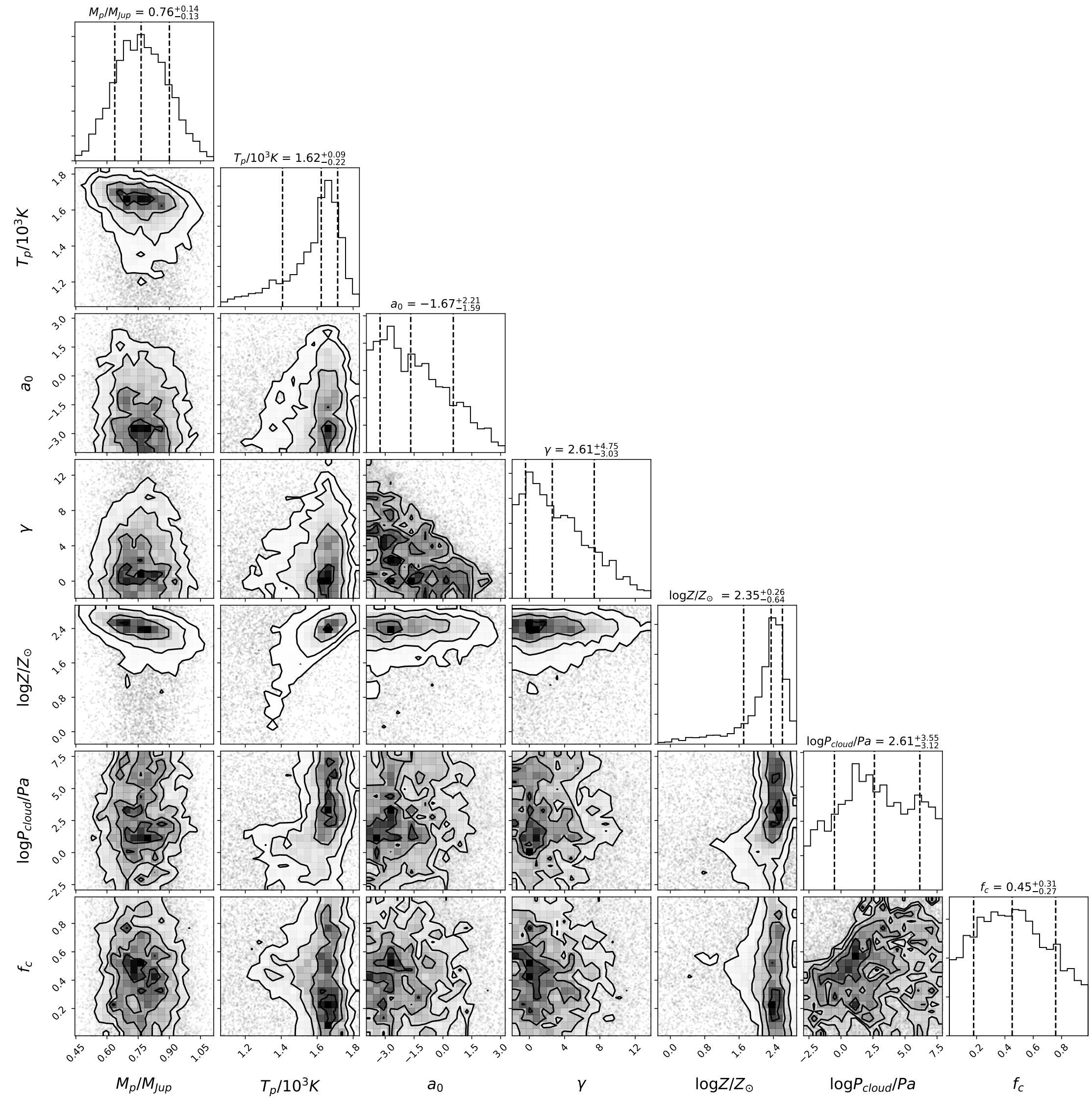


Figure 3.7: Corner plot illustrating the posterior probability distribution from PLATON for the partial cloud and parametric scattering case. R_p, R_s, and C/O are not shown for clarity, since their marginalized posterior distributions are the same as in the fiducial case. The 1-σ metallicity range is shifted down to $\log_{10} Z/Z_\odot = 2.34^{+0.27}_{-0.64}$. Note that at $f_c = 0$ or $\log P_{cloud} > 10^{2.5}$, I recover the fiducial marginalized posteriors.

3.7.2.3 Instrumental Offsets

I model an instrumental offset as a constant value added to the forward model's binned transit depth in the wavelength range of the instrument of interest. For the physically motivated scenario (Scenario 1 from Section 3.6.1.2) , I set the priors for STIS G430L, STIS 750L, and WFC3 to be Gaussians centered on zero ppm with widths set to the uncertainty on the transit depth from their white light curves (105, 85 and 80 ppm, respectively; see Table 3.10).

The retrieved median WFC3 offset is non-trivial, with a median of about $1.5\times$ the white light uncertainty (130 ± 50 ppm). The offsets in the STIS G430L and G750L are less significant, at 58 ± 58 ppm and -65 ± 55 ppm, both well within their white light uncertainties. However, there is a significant median offset of ~ 120 ppm between the two instruments. This is driven primarily by the retrieval attempting to align transit depths in the overlapping wavelength region between the instruments.

The *Spitzer* 3.6 μm point drives the WFC3 offset: shifting the WFC3 depths down necessitates a smaller radius ratio, which better captures the relatively low transit depth at 3.6 μm. The ability to better capture the Spitzer 3.6 μm results in a higher evidence, indicating that this model is strongly preferred over the fiducial model (for a more detailed discussion, see Section 3.7.3).

When combined with partial clouds, the instrumental offsets cause a small decrease in the retrieved median metallicity ($\log_{10} Z/Z_{\odot} = 2.33^{+0.23}_{-0.25}$). This is for similar reasons as explained in the parametric scattering section; whereas parametric scattering justified the absence of an optical scattering slope in the low-metallicity solution by effectively removing Rayleigh scattering opacity, the instrumental offset model can decrease the WFC3 depths relative to the STIS depths to artificially allow for it.

Note that increasing STIS depths (instead of decreasing WFC3 depths) has the same affect on Rayleigh opacity and thus metallicity. However, it is not a viable solution since — unlike decreasing WFC3 depths — it does not improve the forward model's ability to capture the low Spitzer 3.6 μm point.

Allowing Gaussian-prior instrumental offsets had no significant effect on the results. However, it is possible that there is some unknown wavelength-independent systematic that biases the absolute transit depths of the instruments relative to one another. Though unlikely, to explore this I allowed offsets in the STIS G430L, STIS G750L, and WFC3 data to vary by about 5% (500 ppm) in either direction (Scenario 2 from Section 3.6.1.2). Due to the model preferring a lower radius ratio to best explain the Spitzer 3.6 μm point, the median WFC3 offset is a 250 ppm decrease, about 3× the white light uncertainty. Surprisingly, this large offset does not significantly change the 1-σ ranges for metallicity ($\log_{10} Z/Z_\odot = 2.26^{+0.24}_{-0.40}$). The size of the molecular features and the large differential between *Spitzer* photometric points drive the supersolar metallicity in this case. Regardless of the magnitude of the offset, I retrieve a high metallicity. The two large, uniform offset case, where both STIS instruments are offset by the same amount (Scenario 3 from Section 3.6.1.2), retrieves effectively identical posterior distributions as Scenario 2.

While it is worthwhile to understand the effect on the retrieval, there is no reason to expect such large instrumental offsets for HAT-P-41b. A transit depth offset can be caused by the necessity of analyzing each instrument differently. For example, not handling limb darkening consistently and not using consistent orbital parameters (i.e., inclination) for each analysis might cause an offset, but this is easily fixed and is not an issue for my dataset. Since the instruments' observations are from different dates, it is also possible that stellar variability could cause an offset. However, I have long-term photometry (Sec 3.3.3) that shows no such variability. Additionally, the STIS depths are in good agreement with HST UVIS observations (Wakeford et al., 2020). There is no indication that this particular observation is biased in any way, and unresolved companions are confidently ruled out (Evans et al., 2018). The most plausible source is unaccounted for uncertainties or bias from the spectral analysis, as the WFC3 spectrum derived in this chapter is shifted up ~ 90 ppm relative to the literature spectrum (Tsiaras et al., 2018), as noted in Section 3.5.2.2. However, this is still well below the 250 ppm value preferred by the large, uniform offsets models. I determine that offsets beyond the physically motivated values are unlikely.

3.7.2.4 Stellar Activity

Section 3.4.1 demonstrated that HAT-P-41 is consistent with a quiet star and stellar activity is not expected to impact the transit spectrum. However, to be conservative, I investigated if allowing for greater stellar variability impacted my conclusions.

The typical signature of un-occulted, cool starspots is to mimic a haze-like slope in the transit spectrum, and such a signature is clearly absent in the derived transit spectrum of HAT-P-41b. On the other hand, the signature of hot faculae is a steep optical drop-off towards shorter wavelengths (Rackham et al., 2019). Given that I see a drop in transit depths in the optical, the retrieval could plausibly be affected if faculae dominate over star spots, and so I focus on a faculae overabundance.

I assumed that the temperature of the stellar photosphere equals the stellar effective temperature. I modeled the faculae following the prescription from Rackham et al. (2019) and, accordingly, fixed the faculae temperature to $T_{phot} + 100$K. PLATON weights the contributions from the different temperature regimes via the fractional coverage parameter, which represents the overabundance of faculae in the unocculted regions. Rackham et al. (2019) states that moderately active F5V-dwarfs will have around 1% faculae coverage, and up to about 7% on the more active end. This is the faculae fraction, which is likely much higher than the faculae overabundance. However, I set a conservative uniform prior on the fractional coverage of 0-10% in order to determine if high activity would significant alter my conclusions.

I find that including stellar activity has no effect on the posterior probability distribution. It may seem that the STIS data could be explained by a featureless flat line and stellar activity instead of a TiO feature. However, the overabundance of faculae necessary to explain the drop in bluest six points (0.32–0.42 μm) produces a poor fit to the rest of the STIS data. Therefore, even when including stellar variability, TiO is necessary to explain the STIS depths. Allowing a wider range of faculae temperatures also had no effect.

In summary: there is no evidence of stellar variability from prior observations, and allowing for activity does not affect the retrieval.

3.7.2.5 Mie Scattering

Benneke et al. (2019a) recently invoked Mie scattering to explain anomalously low *Spitzer* transit depths. Given the relatively low value of HAT-P-41b's Spitzer 3.6 μm depth relative to the rest of the spectrum — the fiducial model's predicted depth at 3.6 μm is about 3.3-σ away from the observed depth — I included Mie scattering in this analysis. In PLATON, Mie scattering can be used in lieu of parametric Rayleigh scattering.

Each condensate is described by a wavelength dependent complex refractive index, n-ik, where n is the real part and k is the imaginary part of the index. This index explains how that particular condensate interacts with light with wavelengths similar to the particle size. PLATON assumes log-normal distribution in particle size with geometric standard deviation 0.5 to determine the abundance of different radii condensates for a given mean particle size (Zhang et al., 2019). The other relevant factors are cloud height (condensates are only relevant above that pressure; below it gray cloud opacity dominates), particle density at the cloud top pressure, and condensate scale height. The condensate scale height is parameterized as a fraction of the gas scale height, and it describes how the abundance of Mie scattering particles decreases with height. The refractive index is fixed for a given condensate, and the other four parameters are fit for in the retrieval (see 3.10).

PLATON tests one Mie scattering species at a time. Only a few species expected to form clouds in hot Jupiter atmospheres have condensation temperatures above HAT-P-41b's limb temperature ($T \sim 1600$K) (Wakeford & Sing, 2015). Those five (SiO_2, Al_2O_3, $CaTiO_3$, FeO, and Fe_2O_3) fall into two phenotypes: "low-n" with real refractive index $n \approx 1.5$ and "high-n" with $n \approx 2.5$. Though the k values vary more significantly, I find that they do not have a significant impact on the absorption cross section of the condensates. I use n and k values from Kitzmann & Heng (2018), which I average over the relevant wavelength range (0.3–5 μm). The n values are flat over this range, and so the average is an excellent approximation. I tested retrievals with both the low-n (corundum; Al_2O_3) and high-n (hematite; Fe_2O_3) phenotypes.

The priors for the fittable parameters are shown in Table 3.10. The prior for cloudtop pressure is the same as the fiducial model. Since the condensate radii must be such that they cause a relative drop in opacity around 4 μm (i.e., increase opacity more in the near-UV by more than around 4 μm), I can constrain the mean particle size reasonably well. I set the prior to be log-uniform with a range that contains all plausible values. The number density is not known ahead of time, so I set an uninformative log-uniform prior; widening the prior further did not affect the retrieval. Finally, it is unclear what physical constraints there are on condensate scale height. Fortney (2005) finds that condensate scale heights can be one third of the gaseous scale height for hot Jupiters, and Benneke et al. (2019a) found $H_{part}/H_{gas} \approx 3$ for a sub-Neptune. Using these values as guides, I set a conservative uniform prior on the fractional scale height and constrain it to be in the range 0.1–10.

Including Mie scattering opacity does not noticeably affect the results of the retrieval, and the Mie scattering parameters are not constrained by the retrieval. The inferred small gaseous scale height — which dampens features and is necessary to explain STIS and WFC3 feature sizes — makes it difficult to explain the large variations in the radius ratio. Combining Mie scattering with partial clouds — physically, an atmosphere with patchy clouds and Mie scattering particles distributed only above those clouds — alleviates the issue of explaining the large transit depth variation. Since partial clouds allow for higher scale heights, Mie scattering could then cause a larger drop in transit depth near *Spitzer* without needing to invoke an unreasonably high fractional scale height.

Figure 3.8 shows the corner plot for this model. Though the Mie scattering parameters are not constrained, at number densities above $\sim 10^8\,\mathrm{m}^{-3}$, particle radii around $0.15\,\mu\mathrm{m}$, and condensate scale heights greater than the gaseous scale height, lower metallicity and temperature values are possible. This is because added Mie opacity tends to mute features near its peak opacity. This provides a physical reason to expect smaller spectral features, and so a less small scale height is necessary to fit the features. The net impact is a decreased — but still supersolar — median metallicity of $\log_{10} Z/Z_{\odot} = 2.27^{+0.30}_{-0.55}$

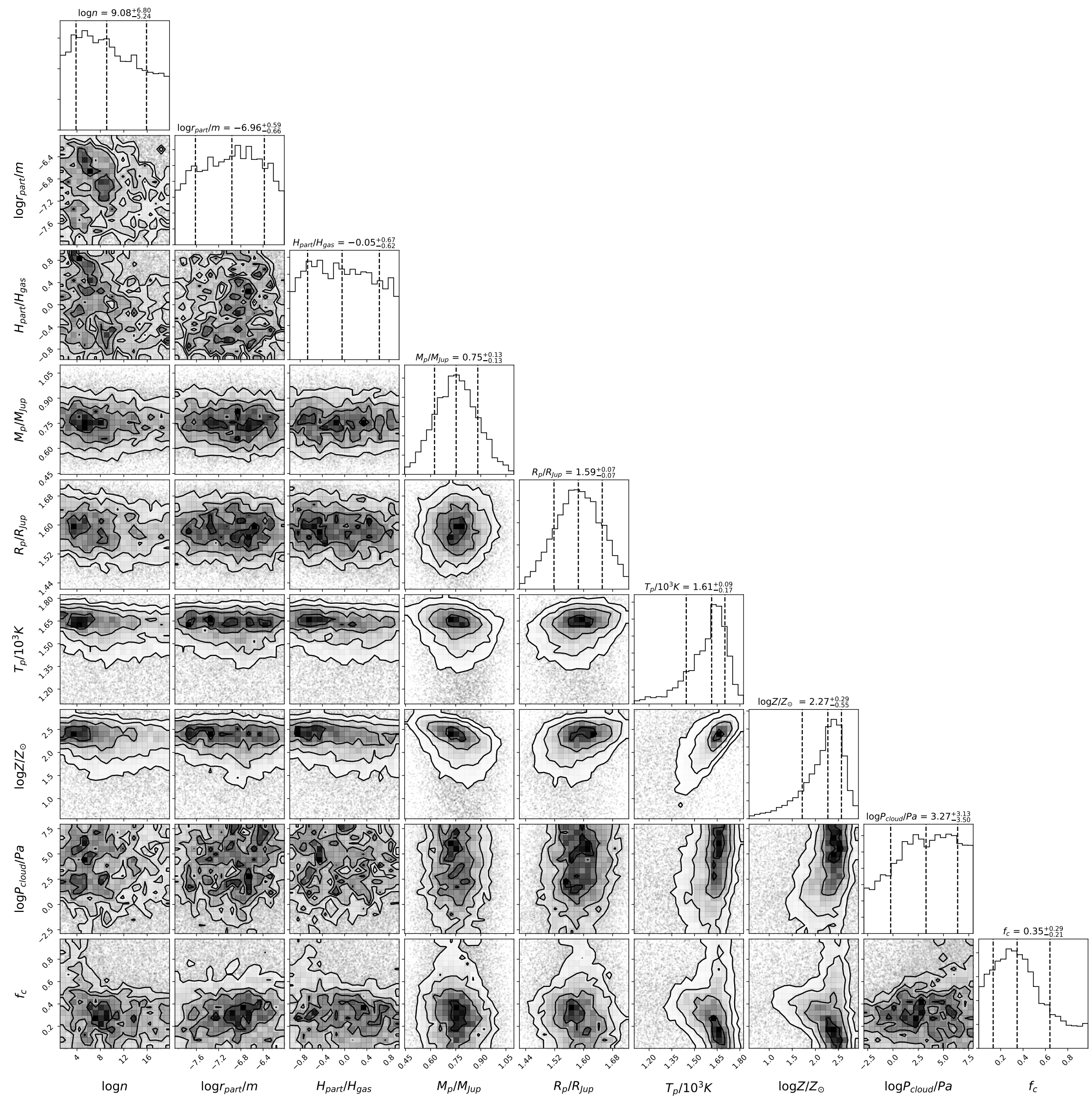


Figure 3.8: Corner plot illustrating posterior probability distribution from PLATON for the fiducial plus Mie scattering and partial clouds model. R_s, and C/O are not shown for clarity, since their marginalized posterior distributions are the same as in the fiducial case. The 1σ metallicity range is shifted down a bit to $\log_{10} Z/Z_{\odot} = 2.27^{+0.30}_{-0.55}$. Note that I recover the fiducial marginalized posteriors when any of the mean particle size, the particle number density, or the condensate scale height are too small.

3.7.3 Model Selection

Section 3.7.2 stepped through the PLATON retrieval for increasingly complex models, examining both how each additional parameter affected the posterior and why it affected it in that way. While knowing the effect of each model assumption is useful, it is important to determine a preferred model in order to effectively convey the results. In this section, I use Bayesian model comparison to select the best model.

Model selection is as important as parameter estimation in atmospheric retrievals. I determine the preferred model by a combination of physical arguments and Bayesian statistics. Specifically, I check if it is necessary to consider more complicated physics using the odds ratio, which is the Bayes factor between models (defined as the ratio of their evidences) multiplied by their prior probability ratio. The prior probability ratio is typically assumed to be one (i.e., the models are assumed to be equally likely). The odds ratio determines if one model should be preferred over another by intrinsically rewarding better fits while punishing overcomplicated structure (Trotta, 2008). This is entirely data-and-model defined, assuming appropriately uninformative priors are used. I compare the Bayesian evidences of each model in order to determine which should be favored.

The Bayesian evidences and 1-σ metallicity ranges for every model discussed in Section 3.7 are shown in Table 3.11. The 1-σ range is represented by both the median metallicity with quantiles (i.e., the central 68% of metallicity values) The 1-σ metallicity ranges are included to illustrate the uncertainty caused by model choice. The retrieved atmospheric metallicities are remarkably consistent across the models, and a supersolar metallicity is ubiquitous. This demonstrates that under PLATON's assumptions, supersolar metallicity is a robust conclusion.

Figure 3.9 emphasizes the insensitivity of the atmospheric parameters to model assumptions. This shows the one dimensional marginalized posterior distributions for metallicity, temperature, and C/O for five of the models I examined. These specific models are shown because they are "interesting" in that they differ from the fiducial model's posteriors the most. I emphasize that these are these are the models that most differ from the fiducial case. While including instrumental offsets tends to flatten the distributions, the peaks of all of the models are consistent.

I define the model selection-relevant columns here:

- $\ln \mathcal{Z}$ is the natural log of the Bayesian evidence. A higher value indicates the model is better able to describe the data without overfitting.

- $\mathcal{O}$ is the odds ratio in favor of a model over the fiducial model. It is the product of their Bayes factor and their prior probability ratio. The prior probability ratio is often assumed to be one, as is the case here. This can be directly interpreted: an odds ratio of 100 indicates 100:1 odds in favor of the more complex model. Values less than one indicate evidence against the corresponding model.

- **Interpretation** This is the empirically derived interpretation of odds ratio based on the Jeffreys' scale (Trotta, 2008).

Table 3.11 contains every notable model I considered. I did not do an iterative

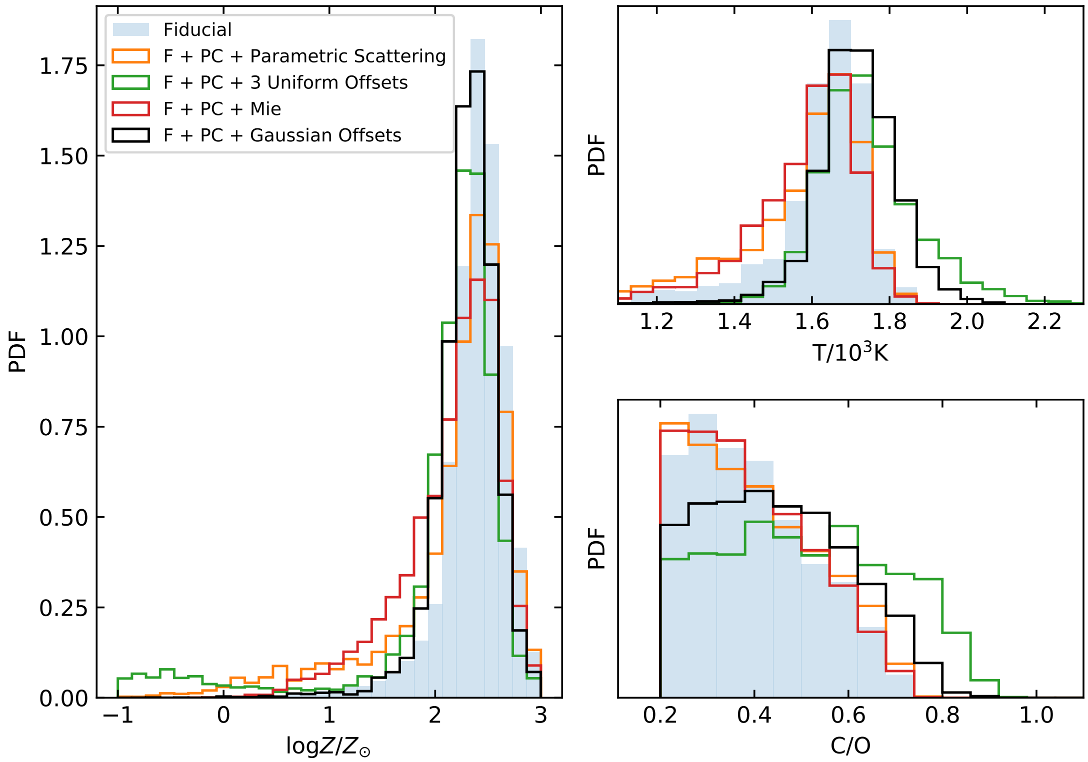


Figure 3.9: Marginalized posterior distributions for metallicity, temperature, and C/O from select models compared to solar values. Note that stellar C/O = 0.19 (Table 3.4). Though some models have low metallicity tails, the 68% credible interval for metallicity is robust (Table 3.11). Offsets allow for higher temperatures and C/O ratios, while both parametric and Mie scattering allow for lower temperatures.

combination of every model scenario, for two primary reasons. Most importantly, I am weary of overfitting the data. The fiducal model is already an excellent fit to the data ($\chi^2_\nu = 1.09$), so I must be careful about adding complications. Layering multiple parameter physical processes, such as Mie scattering and stellar activity, involves an extra five parameters and significantly overcomplicates the fit. Instead, I only combine complications when there is a physically motivated reason to do so, e.g. partial clouds. The second reason is computational difficulty. Some model combinations have enough free parameters to describe the data with many different

Table 3.11: Evidence and metallicity ranges for each plausible model. Only the models including instrumental offsets are preferred over the fiducial model. No model assumption changes the conclusion of a supersolar atmospheric metallicity.

Model	$\log_{10}Z$[a]	Z[b]	$\ln \mathcal{Z}$[c]	$\mathcal{O}$[d]	Interpretation[e]
Fiducial (F)	$2.41^{+0.23}_{-0.25}$	145–437	551.6	Ref	Default Model
F + Partial Clouds (PC)	$2.38^{+0.22}_{-0.37}$	102–398	551.7	1.1	Inconclusive
F + PC + Parametric Scattering	$2.34^{+0.27}_{-0.64}$	50–407	550.7	0.4	Inconclusive
F + PC + Stellar Activity	$2.41^{+0.27}_{-0.41}$	100–479	551.8	1.3	Inconclusive
F + Mie Scattering	$2.41^{+0.24}_{-0.29}$	132–447	551.0	0.6	Inconclusive
F + PC + Mie	$2.27^{+0.3}_{-0.55}$	52–372	551.8	1.3	Inconclusive
F + PC + 3 Gaussian Offsets	$2.33^{+0.23}_{-0.25}$	120–363	556.4	122	Strongly preferred
F + PC + 3 Uniform Offsets	$2.26^{+0.24}_{-0.4}$	72–316	556.9	213	Strongly preferred
F + PC + 2 Uniform Offsets	$2.30^{+0.26}_{-0.39}$	81–363	554.9	29	Moderately preferred

[a] Median log metallicity with 16% and 84% quantiles, in units of log solar metallicity
[b] 68% credible interval for metallicity, in units of solar metallicity
[c] Natural log of Bayesian Evidence
[d] Odds ratio between model and the fiducial model
[e] According to Jeffreys' scale (Trotta, 2008)

combinations, and so the retrieval does not converge on the timescale of weeks. Mie scattering combined with offsets falls in this category. However, given the overfitting concerns and the lack of evidence for just Mie scattering, I do not think this is a worrying omission.

It is generally best practice to assume the simplest model unless there is evidence in favor of extra parameters. That is why I list the fiducial model as the reference, and determine the evidence of the more complicated models. If the evidence of the model with extra parameters is not significantly greater, it means that the ability to explain to data was not improved enough to justify the added complexity. This essentially quantifies Occam's razor.

Following this logic, I determine the "fiducial + partial clouds + 3 Gaussian offsets" model to be the best model. Only the Gaussian offset and uniform offset models are preferred over the fiducial model. While the three uniform offsets model has the highest evidence/weight, the odds ratio between that and the Gaussian offsets model is 1.75. This is inconclusive on the Jeffrey's scale, meaning I am unable to distinguish between the models by evidence alone. Instead, I favor the Gaussian offsets model as more plausible, since its Gaussian priors are physically motivated by common-mode corrections.

The evidence for partial clouds is inconclusive, however partial clouds are more plausible than assuming 100% cloud coverage, as argued by MacDonald & Madhusudhan (2017); Welbanks & Madhusudhan (2019). Therefore, to be conservative, I choose the model which account for partial clouds as the "best" model.

The odds ratio only works in a direct comparison of two models and is not a statement on the absolute goodness-of-fit. The reduced chi-squared test statistic is a useful sanity check to ensure that the model is able to explain the variance in the data. The value for the best model is an ideal $\chi^2_\nu = 1.0$. The results section (Section 3.8) — and the abstract values — are based on parameter estimation from the "fiducial + partial cloud + Gaussian offsets" model.

3.7.3.1 Bayesian Model Averaging

Instead of model selection, it is possible to take a weighted average of the results from each model and therefore automatically take their respective evidences into account (Gibson, 2014b; Wakeford et al., 2016a, 2018). The benefit of Bayesian model averaging is the ability to quantify uncertainty in model selection, as well as avoiding having to arbitrarily choose between models with slightly different evidences. However, it requires a few assumptions: it is only valid if the set of models comprises the full model space, i.e, at least one model is a good description of the data. The weight-averaged uncertainties assume Gaussian-distributed posteriors, which is not strictly correct. However, it is useful in combining information from every model.

Here, I show the assumptions I make to use Bayesian model averaging. The χ^2_ν values for the models I tested are clustered around one, so it is fair to assume that a "correct" model is contained in the set. Figure 3.9 shows that although the posteriors are not perfectly Gaussian, they have sharp, unimodal peaks, and so the uncertainty derived from marginalization is informative.

The model weights are defined by Eq 3.2 (adapted from Gibson (2014b)). W_q is the weight assigned to model q, $P(M_q|D)$ is the the likelihood of model q given the data, and $P(D|M_q)$ is the likelihood of the data given model q, which is equivalent to the Bayesian evidence of model q, E_q. The denominator is a normalization term, summed over N models. I assume a conservative prior that each model is equally likely ($P(M_i) = 1$ for all i).

$$W_q = \frac{P(M_q|D)}{\sum_{i=1}^{N} P(M_i|D)} = \frac{P(D|M_q)P(M_q)}{\sum_{i=1}^{N} P(D|M_i)P(M_i)} = \frac{E_q P(M_q)}{\sum_{i=1}^{N} E_i P(M_i)} \tag{3.2}$$

The marginalized log metallicity with 1-σ uncertainties is calculated from equations 15 and 16 from Wakeford et al. (2016a). The result is $\log_{10} Z/Z_\odot = 2.29^{+0.24}_{-0.36}$ ($194^{+144}_{-109}\times$ Z$_\odot$). As expected, the highest weighted models are the offset models.

Bayesian model averaging demonstrates that in PLATON's chemical equilibrium framework, a supersolar metallicity is the most likely result even after accounting for uncertainty in model selection.

The marginalized metallicity is useful as a reference, but it is valuable to give the metallicity distribution for each specific model assumption. Marginalization is most appropriate when the specific model parameters are unimportant, however I am interested in the impact that modeling assumptions have on the atmospheric parameters. I emphasize that even for apparently "data-defined" methods, many assumptions have to be made and those should be explicitly stated for an appropriate interpretation.

3.8 Results for the Favored PLATON Model

In Section 3.7.3 I argued that the best PLATON model scenario is the fiducial model with partial clouds and physically motivated, Gaussian prior instrumental offsets added. The retrieved median spectrum with uncertainty contours is shown in Figure 3.10. It is an excellent fit to the data, with $\chi^2_\nu = 1.0$. In this section I discuss the details of the retrieved atmospheric parameter values.

3.8.1 Summary of Retrieved Parameters

The posterior distribution corner plot is shown in Figure 3.11. Both atmospheric metallicity and temperature are well constrained, and the C/O ratio, though relatively flat, has a strict upper limit. The median retrieved metallicity is super-

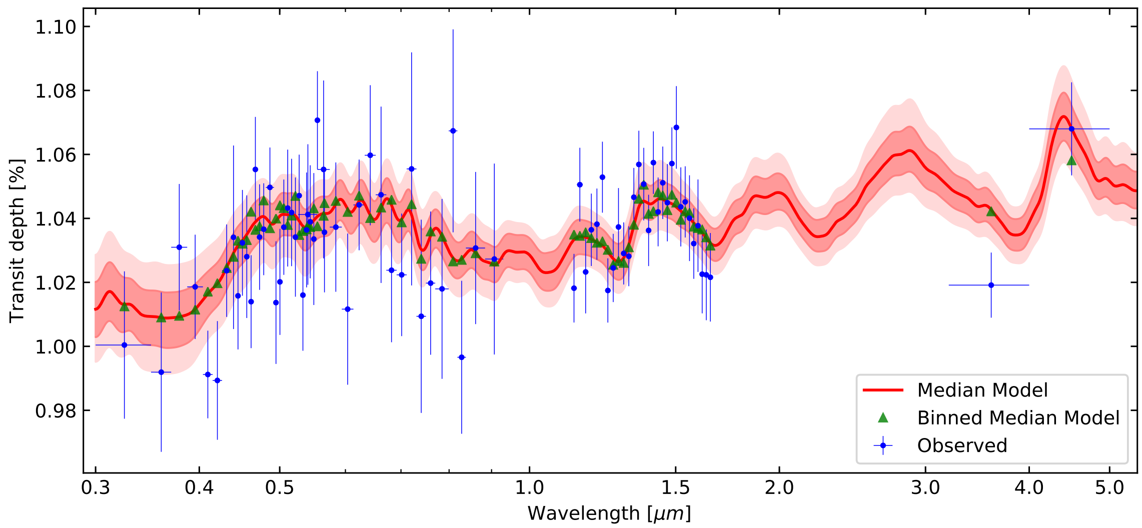


Figure 3.10: Median retrieved model with 1-σ and 2-σ uncertainty contours for the favored PLATON model (fiducial model with partial clouds and instrumental shifts with physically-motivated Gaussian priors).

solar ($\log_{10} Z/Z_{\odot} = 2.33^{+0.23}_{-0.25}$), and solar metallicity is inconsistent to 3-σ (lower limit 4.8× $Z_{\odot}$). As noted in Section 3.6.3, PLATON's derived metallicity is a proxy for [O/H], enabling comparison with its host star's oxygen abundance ([O/H]=0.37; Table 3.4). PLATON determines HAT-P-41b to be metal-enriched relative to its host star (($\log_{10} Z/Z_{star} = 1.97^{+0.23}_{-0.25}$), and it is inconsistent with the stellar metallicity to to 3-σ (lower limit 2.1× Z_{star}). The planetary C/O ($0.44^{+0.18}_{-0.15}$) has a 3-σ upper limit of 0.83. Though the planetary C/O is technically inconsistent with the stellar C/O to 1-σ (0.19; Table 3.4), the comparison is not valid as the planetary C/O prior had a computational lower limit of 0.20, and the posterior has significant likelihood at that limit. This "piling" at the prior boundary implies that the planetary C/O is consistent with the stellar C/O. The median isothermal limb temperature ($T_{\rm limb} = 1710^{+100}_{-80}$ K) is close to the equilibrium temperature of the planet

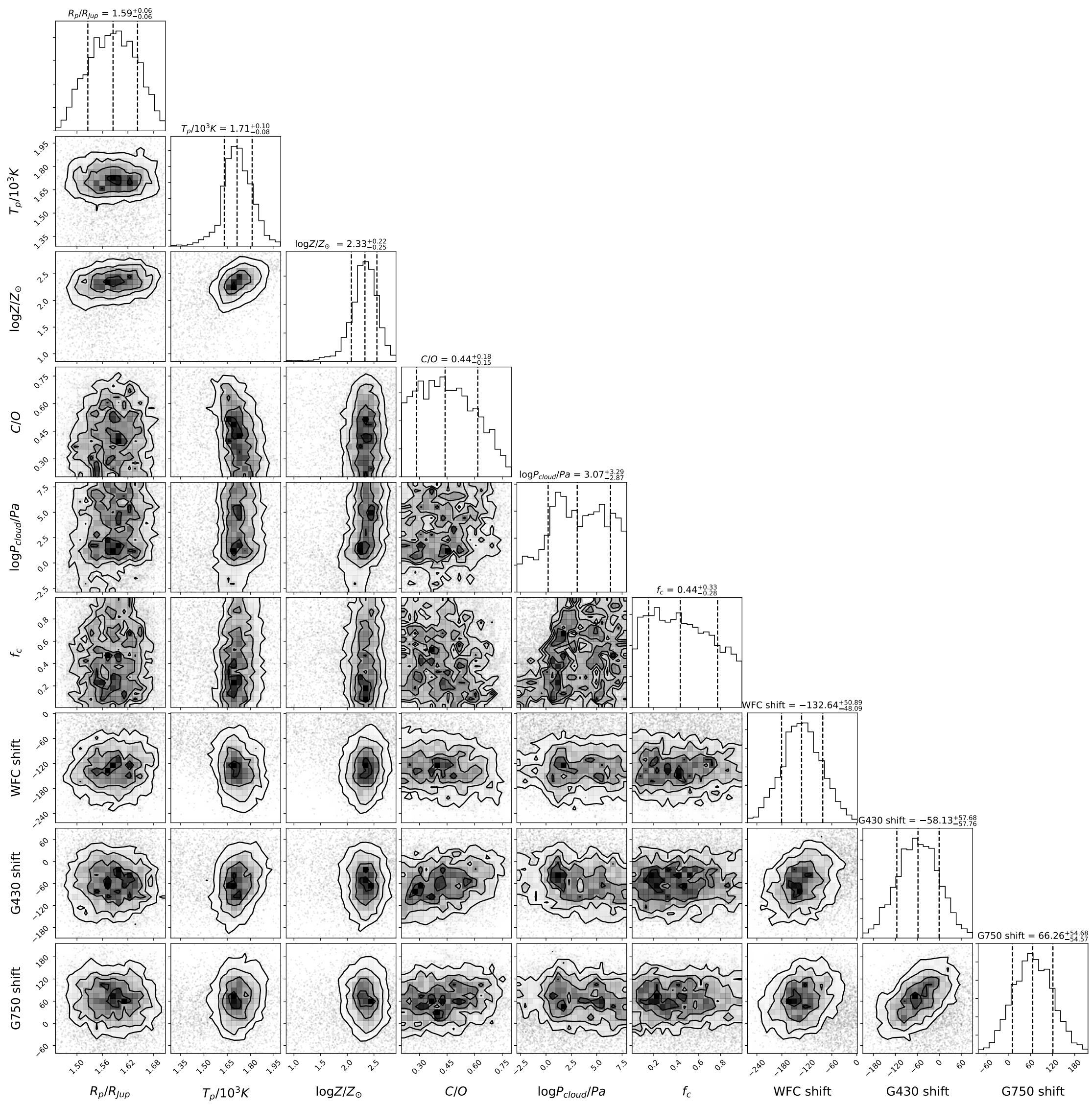


Figure 3.11: Corner plot for the best PLATON model (fiducial with partial clouds and Gaussian offsets). R_s and M_p are prior dominated and are excluded for clarity. The offsets are given in parts per million; for example, the median WFC3 offset indicates the retrieval favors shifting the WFC3 depths down by $\sim 132\,\text{ppm}$.

($T_{eq} = 1960$ K), which implies an efficient heat recirculation. These parameters lead to a high mean molecular weight ($\mu \sim 5.5$ AMU) atmosphere with a scale height of about 320 km.

The retrieved results are consistent with a clear atmosphere. Though cloud top pressure and cloud fraction are unconstrained, their joint marginalized posterior is constrained. A uniform grey cloud is only allowed deeper than $\sim$ 10 Pa (0.1 mBar), and clouds above that pressure are only possible if they cover less than about 40% of the limb. Hazes are dispreferred by model selection, and the median scattering opacity was 50× weaker than Rayleigh scattering in the model which allowed parametric scattering.

The retrieved relative shift between the STIS G430L and G750L instruments is 120ppm, due in part to the model attempting to align their overlapping regions. A downshift for the WFC3 data is preferred (WFC3 offset $= -132 \pm 50$ ppm). The stellar radius, the planetary mass, and the planetary radius are consistent with the prior values. The planetary mass and stellar radius are, as expected, dominated by their priors. The planetary radius ($R_p = 1.59 \pm 0.06$) is at the reference pressure of 1 Bar, and when calculated at the planet's photosphere it is consistent with the planetary radius derived based on stellar parameters from TIC-8.

3.8.2 Evidence of Water and Optical-Wavelength Absorbers

While the spectral features in STIS, WFC3, and *Spitzer* are attributed by PLATON to TiO, H_2O, and CO_2, respectively, the retrieval only robustly detects

Table 3.12: PLATON species detection evidences

Species	$\mathcal{O}$[a]	Detection Significance[b]
H_2O	46630	5.0σ
TiO	2.1	1.9σ
VO	2.3	1.9σ
TiO/VO	9.4	2.7σ
Na	1.1	1.2σ
CO_2	3.3	2.1σ
CO	0.4	N/A

[a] Odds ratio between model and the preferred PLATON model (fiducial model with partial clouds and Gaussian-prior instrumental shifts included)
[b] Benneke & Seager (2013)

H_2O - the H_2O abundance is constrained by observations, while the abundances of other species are primarily constrained by the assumption of chemical equilibrium. I note that while CO is more abundant than CO_2, CO_2 has a much larger cross section at $4.5\,\mu$m, such that even with a smaller abundance its opacity dominates over that of CO at the temperatures and C/O ratios inferred by the retrieval.

I determine if a species is detected by finding the odds ratio between the best model with and without opacity from a particular species. This breaks the assumption of chemical equilibrium, so it is not strictly correct, but it is a useful heuristic nonetheless. A species is considered detected only when the odds ratio significantly favors the model with the species' opacity. Table 3.12 shows the odds ratios — and their more familiar frequentist analog, the detection significances (Benneke & Seager, 2013) — for several relevant spectroscopically active species.

The odds ratio in favor of H_2O is ~ 46630, indicating that the model with water is $46630\times$ more likely than the model without water opacity. This is equivalent to a 5.0-σ detection in frequentist terms. The odds ratio in favor of CO_2 is 3.3, which is barely enough evidence to claim a weak detection. PLATON finds no evidence of Na, and CO is dispreferred. The odds ratios for TiO and VO are 2.1 and 2.3, respectively, and these are not favored enough to claim detections (less than 2-σ). However, TiO and VO are only seen as non-detections because they have similar cross-sections. When TiO opacity is ignored, the retrieval can compensate because VO opacity is able to describe the STIS feature just as well as TiO. If I ignore both VO and TiO then the model cannot describe the STIS data as well, and so the odds ratio in favor of TiO/VO is 9.4 (2.7-σ). Therefore, I find suggestive evidence of metallic oxide opacity, but I am unable to discern if it is due to TiO or VO. Based on the assumption of chemical equilibrium at the retrieved temperatures, PLATON attributes the STIS feature to TiO because it is more abundant and opaque in the spectrscopically active region for a solar Ti/V ratio.

3.9 AURA Retrieval Analysis and Results

A collaborator performed a second, complementary atmospheric retrieval analysis: a series of free-chemistry retrievals on HAT-P-41b using AURA (3.6.2) to constrain the atmospheric properties at the day-night terminator of the planet while allowing for deviations from chemical equilibrium. First, we consider the presence of different chemical species in the atmosphere of HAT-P-41b using its full broadband spectrum. Then, we consider the presence of possible transit depth offsets between data sets and their possible impact in the derived chemical abundances and associated metallicities.

3.9.1 Evidence of Water and Optical-Wavelength Absorbers

We perform a full retrieval on the broadband spectrum of HAT-P-41b and present the observations and retrieved median spectrum in Figure 3.12. The full retrieval provides constraints on the presence of H_2O, and provides indications for the presence of Na and/or AlO in the optical. The full retrieval finds $\log_{10}(X_{H_2O}) = -1.65^{+0.39}_{-0.55}$, $\log_{10}(X_{Na}) = -3.09^{+1.03}_{-1.83}$ and $\log_{10}(X_{AlO}) = -6.44^{+0.66}_{-0.91}$. While the retrieval with PLATON prefers TiO/VO to explain the STIS observations, the retrieval with AURA does not, and instead prefers a combination of Na and AlO. The retrieved TiO abundance is low and unconstrained ($\log_{10}(X_{TiO}) = -9.58^{+1.37}_{-1.50}$). Neither the CO nor CO_2 abundances are constrained by the retrieval. While the cloud/haze parameters are not tightly constrained, the retrieval indicates a coverage fraction of $\bar{\phi} = 0.25^{+0.26}_{-0.16}$ consistent with a mostly clear atmosphere. The temperature profile of the atmosphere is mostly unconstrained. We infer the temperature near the photosphere, at 100mbar, to be $T = 1345^{+349}_{-206}$ K. The posterior distributions for the relevant parameters are shown in Figure 3.13.

I utilise this full retrieval as a reference model to perform a Bayesian analysis and assess the impact of not considering some of these parameters in the models. This change in model evidence is then converted to its more familiar frequentist counterpart, a detection significance (DS) following Benneke & Seager (2013). Table 3.13 shows the different models considered, their model evidence, DS, and $\bar{\chi}^2$.

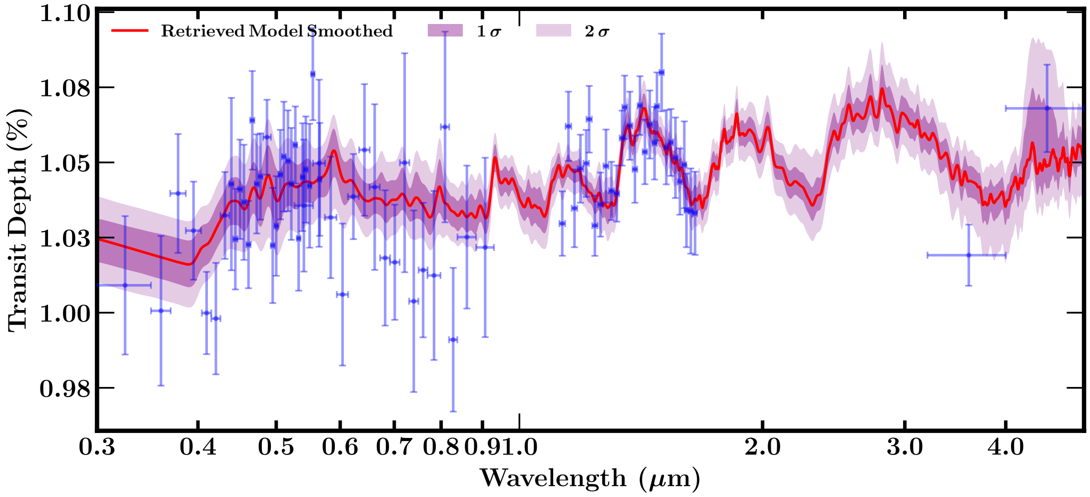


Figure 3.12: Retrieved spectrum of HAT-P-41b using STIS, WFC3 and *Spitzer* data. Observations are shown using blue markers. The retrieved median spectrum is shown in red while the 1-σ and 2-σ regions are shown using the shaded purple areas.

We find a robust detection of H_2O at a 4.89-σ confidence. There is suggestive evidence of Na and/or AlO with confidence levels of 2.09-σ and 2.58-σ, respectively. The removal of TiO from the models results in an increase in the model evidence, indicating a disfavor for this molecule to be present in these models. VO is similarly undetected. However, removing opacity from the three primary metal oxides (TiO, VO, and AlO), finds a moderate-to-strong "detection", with 3.59-σ confidence. This is similar to PLATON, which did not find evidence of TiO or VO individually, but found weak-to-moderate evidence of their combined presence (Sec. 3.8.2). This can be interpreted as follows: AURA is confident (to 3.6-σ) that the sharp dip in the blue STIS data (0.4–0.5 μm) is a real molecular feature due to a metallic oxide. The retrieval finds that the most likely candidate for the metallic oxide is AlO, as shown by it's 2.6-σ preference, whereas TiO and VO are individually dispreferred.

Table 3.13: Impact of AURA model assumptions on retrieved abundances.

Model	$\log_{10}(X_{H_2O})$	$\log_{10}(X_{Na})$	$\log_{10}(X_{AlO})$	$\log_{10} Z/Z_\odot$ [a]	$\ln(\mathcal{Z})$	$\bar{\chi}^2$	DS
Full model	$-1.65^{+0.39}_{-0.55}$	$-3.09^{+1.03}_{-1.83}$	$-6.44^{+0.66}_{-0.91}$	$1.72^{+0.39}_{-0.55}$	559.1	0.93	Ref.
No H_2O	N/A	$-2.41^{+0.99}_{-2.99}$	$-5.71^{+0.99}_{-1.39}$	N/A	548.9	1.37	4.89
No Na	$-1.62^{+0.42}_{-0.67}$	N/A	$-6.90^{+0.84}_{-1.05}$	N/A	558.0	0.95	2.09
No AlO	$-1.49^{+0.35}_{-0.70}$	$-4.32^{+1.88}_{-4.31}$	N/A	N/A	557.0	1.03	2.58
No TiO	$-1.70^{+0.41}_{-0.56}$	$-2.97^{+0.95}_{-1.25}$	$-6.39^{+0.66}_{-0.88}$	N/A	559.7	0.92	N/A
No Metal Oxides	$-1.52^{+0.38}_{-0.91}$	$-3.59^{+1.28}_{-1.47}$	N/A	N/A	554.2	1.21	3.59
Simpler model	$-1.65^{+0.40}_{-0.63}$	$-2.60^{+0.94}_{-1.10}$	$-5.81^{+0.51}_{-0.66}$	$1.72^{+0.40}_{-0.63}$	560.0	0.89	N/A
Gaussian shifts[c]	$-1.91^{+0.53}_{-0.68}$	$-2.38^{+0.81}_{-1.33}$	$-6.64^{+0.70}_{-0.96}$	$1.46^{+0.53}_{-0.68}$	562.0	0.90	N/A
Uniform shifts (3)[c]	$-2.96^{+0.98}_{-0.88}$	$-2.43^{+0.84}_{-1.34}$	$-7.05^{+0.75}_{-0.94}$	$0.40^{+0.98}_{-0.88}$	561.8	0.88	N/A
Uniform shifts (2)[d]	$-3.34^{+1.00}_{-0.86}$	$-3.43^{+1.35}_{-2.19}$	$-6.98^{+0.77}_{-0.78}$	$0.03^{+1.00}_{-0.86}$	560.7	0.89	N/A

[a]The metallicity is approximated from water abundance (see Section 3.9.1 for details).
[b]Detection significance (DS) of excluded species as compared to full model. Only valid if evidence of model is less than that of the full model.
[c]Shift applied to each of the three gratings (WFC3 G141, STIS G430L, STIS G750L; see Sec 3.9.2).
[d]Shift applied to each spectroscopic instrument (STIS, WFC3)

We assess the retrieved H_2O abundance relative to expectations from thermochemical equilibrium for solar elemental compositions (Asplund et al., 2009). Assuming a solar composition and 50% of the available oxygen in H_2O, the retrieved H_2O abundance corresponds to a log metallicity ([O/H]) of $\log_{10} Z/Z_\odot = 1.72^{+0.39}_{-0.55}$ (metallicity of $53^{+82}_{-38}\times$ $Z_\odot$). I also compare the retrieved H_2O abundance to the stellar metallicity of the host star ([O/H]=0.37, Table 3.4) and obtain a value of $\log_{10} Z/Z_{star} = 1.35^{+0.39}_{-0.55}$ (metallicity $23^{+33}_{-17}\times$ Z_{star})

We consider the possibility of fitting the data using a simpler model consisting mainly of the parameters that are reasonably constrained by the full model. The simpler model considers the chemical abundances of H_2O, Na, CO, AlO, an isothermal pressure-temperature profile, and a clear atmosphere. The retrieved median fit and confidence contours are shown in Figure 3.14. The simplified model retrieves values

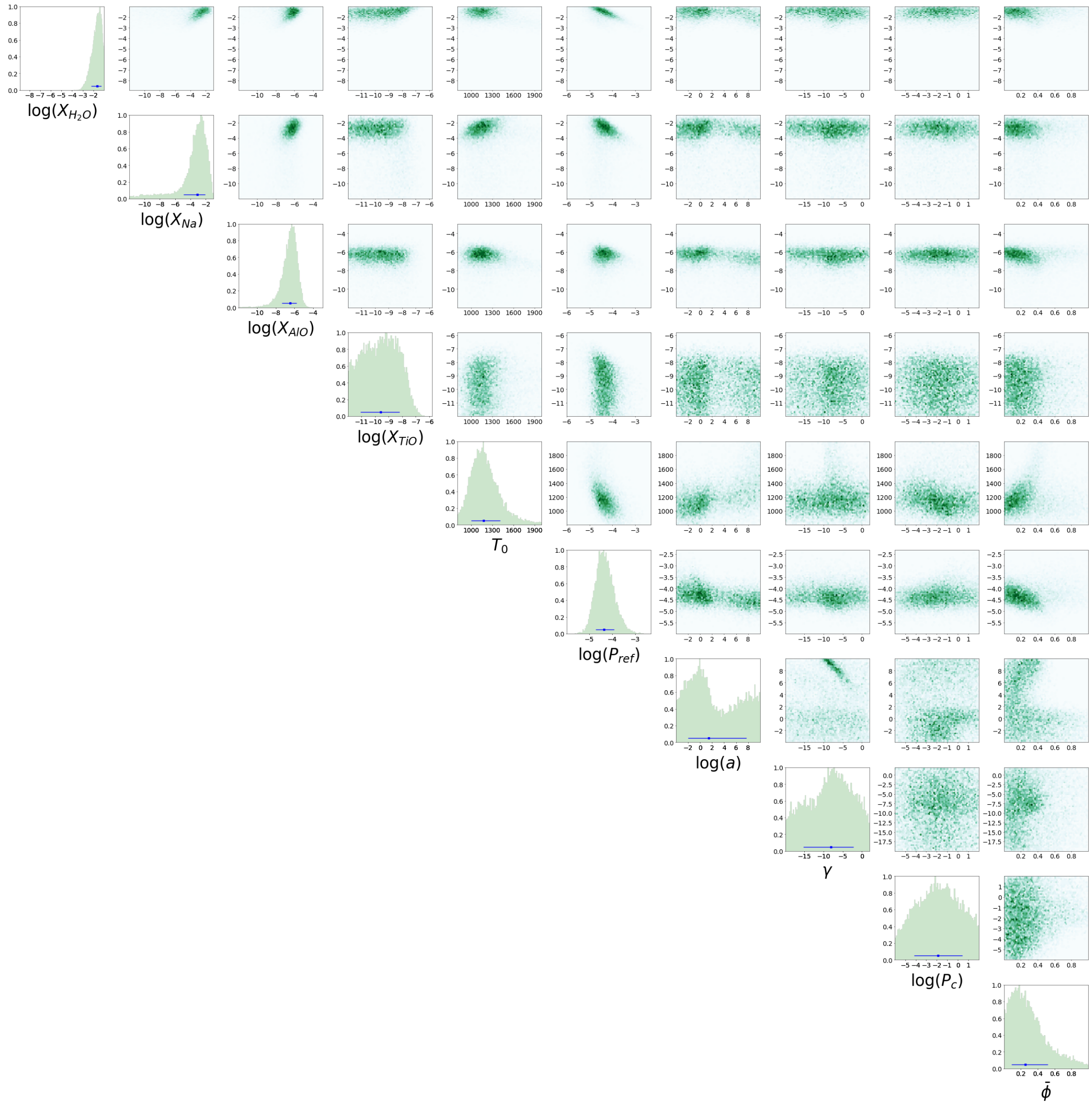


Figure 3.13: Posterior distributions of the relevant parameters for the full retrieval (Model 1 in Table 3.13) using STIS, WFC3 and *Spitzer* data. The abundances of H_2O, Na and AlO are constrained, while the cloud and haze parameters are not constrained. The parameter T_0, the temperature at the top of the atmosphere (10^{-6}bar) is shown as a subset of the P-T parameters used in the model.

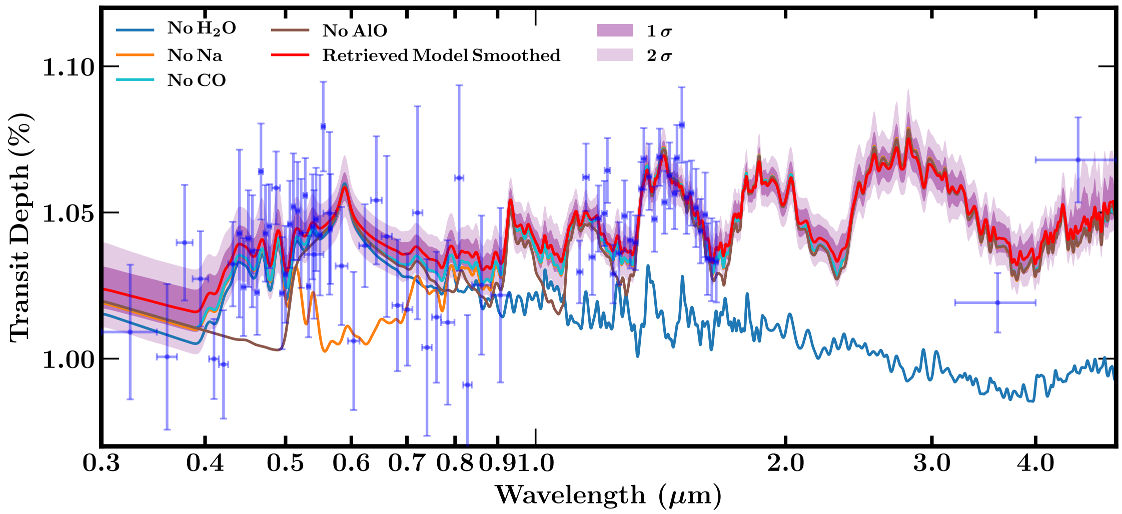


Figure 3.14: Retrieval of HAT-P-41b using a simplified model compared with the fiducial parameter set (see Section 3.9.1). Observations are shown using blue markers. The retrieved median spectrum is shown in red while the 1-σ and 2-σ regions are shown using the shaded purple areas. Forward models using the retrieved median parameters show the contributions to the spectra due to individual chemical species. The forward models shown exclude absorption due to H_2O (blue), Na (orange), CO (cyan), and AlO (brown).

consistent with the full model. The retrieved values are $\log_{10}(X_{H_2O}) = -1.65^{+0.40}_{-0.63}$, $\log_{10}(X_{Na}) = -2.60^{+0.94}_{-1.10}$, $\log_{10}(X_{AlO}) = -5.81^{+0.51}_{-0.66}$, and $\log_{10} Z/Z_\odot = 1.72^{+0.40}_{-0.63}$. The retrieved isothermal temperature is T$= 1120^{+170}_{-140}$ and consistent with the inferred temperature at 100 mbar from the full retrieval. The posterior distribution for the retrieved parameters is shown in Figure 3.15.

We use these retrieved parameters to generate a set of forward models to assess the spectroscopic contribution from each chemical species. Figure 3.14 shows that the WFC3 observations are better explained by the H_2O absorption feature at $\sim$1.4 μm driving its strong detection in the spectrum of HAT-P-41b. On the other hand, a series of chemical species in the optical can provide some degree of fit to the STIS observations. In the optical, between $\sim$0.5–0.7 μm, the broadened wings of Na

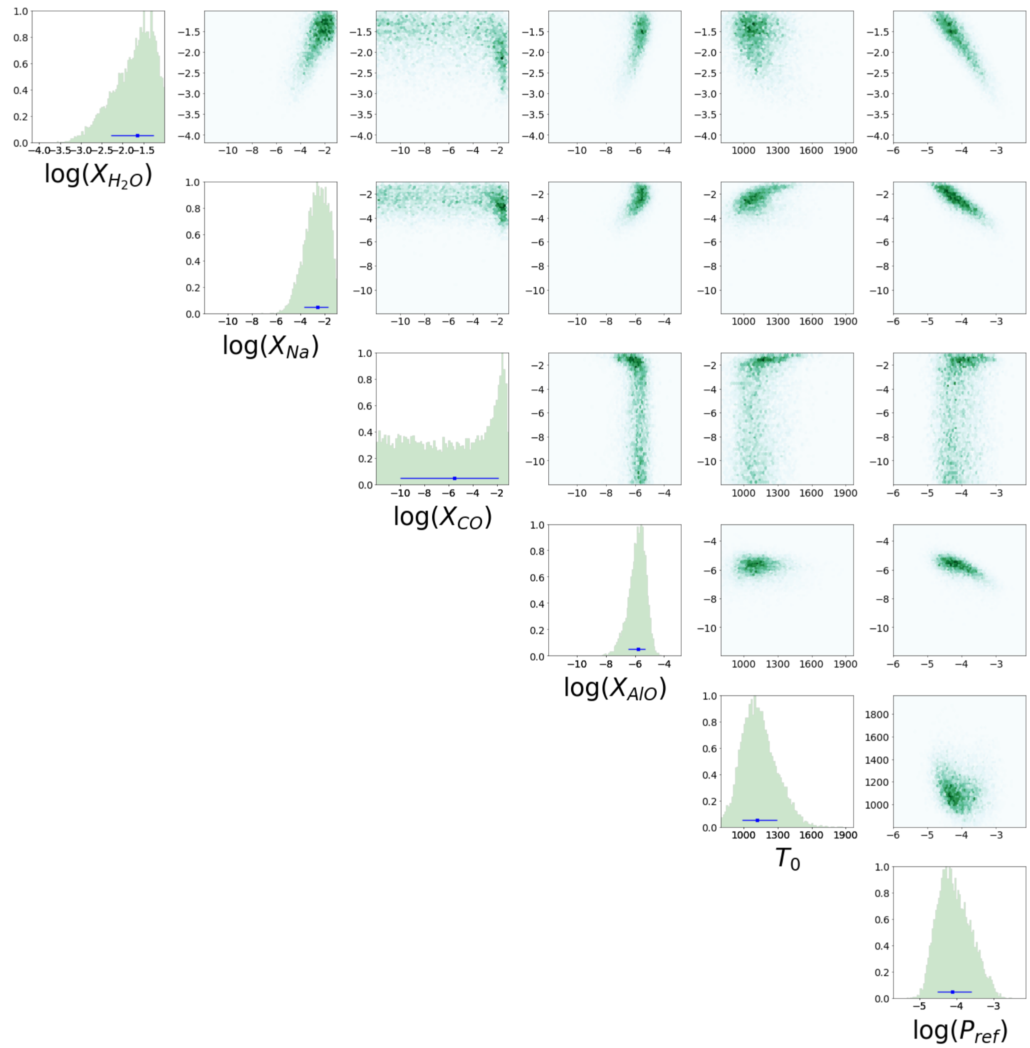


Figure 3.15: Full posterior distributions for the simple AURA model.

along with its absorption peak provide a fit to observations. AlO provides some fit to the substructure present in the STIS observations, particularly the increased transit depth between 0.4–0.5 μm. Lastly the abundance of CO is not constrained and its contribution to the spectrum is minimal. CO is responsible for small changes in the optical and infrared that are well within the error bar of the observations.

The AURA retrieval analysis of the broadband transmission spectrum of HAT-P-41b provides excellent fits to the data; using its fiducial model (Model 1) we obtain a best-fit $\bar{\chi}^2$ of 0.93 and $\ln(\mathcal{Z}) = 559.1$. I note that we do not require additional continuum opacity sources (e.g., H^-) in order to explain the data, as recently claimed by Lewis et al. (2020).

3.9.2 Possible offsets in the data

Lastly, we consider the presence of offsets in the data and their effect on the retrieved atmospheric properties. We consider the three scenarios from Section 3.6.1.2. We note that these retrieved offsets are relative to the atmospheric model and that in all scenarios the *Spitzer* observations remain unchanged. We consider both Gaussian and uniform priors, as seen in Table 3.8.

We present the results of considering the presence of three offsets with Gaussian priors informed by the analysis of the white light transit curves (Model 8; Scenario 1 from Section 3.6.1.2). These priors are shown in Table 3.8. The retrieved shifts are -52^{+61}_{-63} ppm for G430L, 80^{+59}_{-56} ppm for G750L and -91^{+48}_{-50} ppm for WFC3. Similar to PLATON, the retrieval generally prefers to increase the

G750L depths, primarily motivated by aligning the transit depths in the overlapping wavelength region between G430L and G750L. The retrieval also prefers to decrease WFC3 depths in order to better capture the *Spitzer* 3.6 μm depth. The retrieved abundances are $\log_{10}(X_{H_2O}) = -1.91^{+0.53}_{-0.68}$, $\log_{10}(X_{Na}) = -2.38^{+0.81}_{-1.33}$, and $\log_{10}(X_{AlO}) = -6.64^{+0.70}_{-0.96}$. Although the retrieved H_2O abundance corresponds to a lower metallicity estimate, the derived range $\log_{10} Z/Z_{\odot} = 1.46^{+0.53}_{-0.68}$ is consistent with the fiducial model and describes a metal-rich atmosphere. The median metallicity is superstellar ($\log_{10} Z/Z_{star} = 1.09^{+0.40}_{-0.63}$), though it is consistent with stellar metallicity to within 2-σ.

Second, we present the results for the case with three uniform shifts between HST-STIS G430L, HST-STIS G750L, and HST-WFC3 observations (Scenario 2 from Section 3.6.1.2). The retrieved G430L shift is consistent with 0 (1^{+144}_{-156} ppm), but the retrieval prefers a large positive G750L offset (176^{+151}_{-160} ppm) and a large negative WFC3 offset (-189^{+91}_{-94} ppm). In addition to aligning the overlapping G430L-G750L region, it is possible that this large G750L shift is due to the model forcing the data to match features it finds easier to explain. This uncertainty is the danger in using uniform prior offsets, especially in an already-flexible free-chemistry retrieval. The retrieved abundances are shown in Table 3.13 as Model 9 and are $\log_{10}(X_{H_2O}) = -2.96^{+0.98}_{-0.88}$, $\log_{10}(X_{Na}) = -2.43^{+0.84}_{-1.34}$, and $\log_{10}(X_{AlO}) = -7.05^{+0.75}_{-0.94}$. While the retrieved abundances for these three species are consistent within 1-σ with the full unshifted model, the retrieved H_2O abundance corresponds to a lower metallicity estimate consistent with solar, sub-solar, and sub-stellar values $\log_{10} Z/Z_{\odot} = 0.40^{+0.98}_{-0.88}$.

Lastly, we present the results accounting for offsets in the STIS and WFC3 observations using a uniform prior, while keeping the *Spitzer* observations unshifted (Scenario 3 from Section 3.6.1.2). The retrieval results in a shift in the STIS data of 90^{+167}_{-157} ppm and a shift in the WFC3 data of -204^{+97}_{-98} ppm. While the retrieved value for the STIS observations is consistent with no shift, the WFC3 observations preferentially retrieve a negative offset. The derived abundances, shown as Model 10 in Table 3.13, are $\log_{10}(\mathrm{X_{H_2O}}) = -3.34^{+1.00}_{-0.86}$, $\log_{10}(\mathrm{X_{Na}}) = -3.43^{+1.35}_{-2.19}$, $\log_{10}(\mathrm{X_{AlO}}) = -6.98^{+0.77}_{-0.78}$. The H_2O abundance, like Model 9, corresponds to a metallicity consistent with solar and sub-solar values: $\log_{10} Z/Z_{\odot} = 0.03^{+1.00}_{-0.86}$.

Figure 3.16 shows the retrieved median models and confidence contours along with their respectively shifted observations for the cases described in this Section (Models 8, 9, and 10).

The models considering instrumental shifts are all preferred over the fiducial model at above the 2-σ level. The model with Gaussian priors has a preference at the 2.9-σ level, followed by the model with three uniform shifts at a 2.8-σ level. The model with two uniform shifts is preferred over the fiducial model at 2.3-σ. We note that while both models with three offsets are similarly preferred over the fiducial model, the associated metallicity ranges are different. The model with three uniform shifts retrieves an H_2O abundance corresponding to a metallicity estimate consistent with substellar and stellar values. On the other hand, the model with Gaussian priors retrieves an associated metallicity range mostly superstellar and in agreement with the fiducial model. These results highlight the sensitivity of the inferred metallicity ranges to possible large offsets between instruments. Model

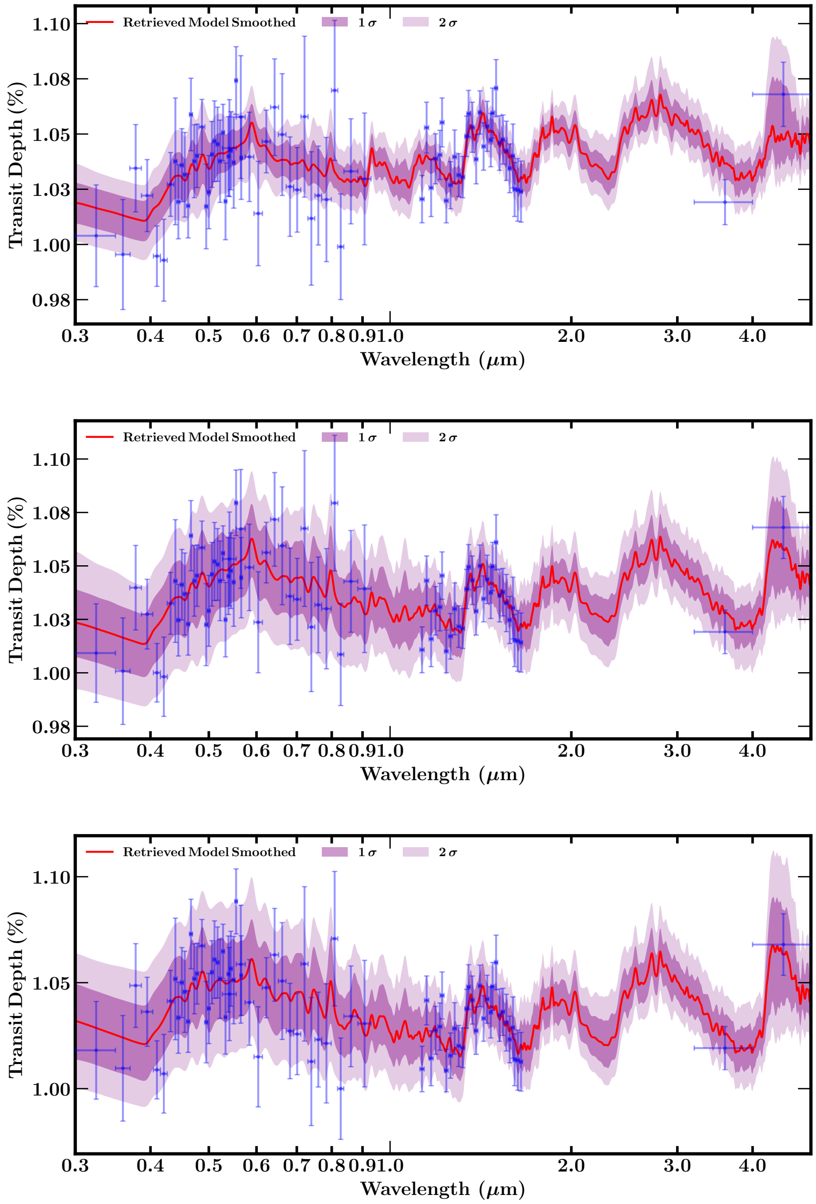


Figure 3.16: Retrieved spectrum of HAT-P-41b allowing for offsets in the STIS and WFC3 data sets. Observations are shown using blue markers and are shifted according to the models' retrieved median shifts. The retrieved median spectrum is shown in red while the 1-σ and 2-σ regions are shown using the shaded purple areas. *Top:* Three shifts with Gaussian priors (Model 8) and retrieved median offsets of ~ -50 ppm for STIS G430L,~ 80 ppm for G750L, and ~ -90 for WFC3. *Middle:* Three shifts with uniform priors (Model 9) and retrieved median offsets of ~ 0 ppm for STIS G430L,~ 180 ppm for G750L, and ~ -190 for WFC3. *Bottom:* Two shifts with uniform priors (Model 10) and retrieved median offsets of ~ 90 ppm for STIS and ~ -200 ppm for WFC3.

comparisons suggest a preference for the models considering offsets, though it is inconclusive between these models. I favor the more physically plausible Gaussian prior model (i.e., Model 8) as the reference for my discussion (Section 3.10).

3.10 Discussion

3.10.1 Comparison Between Retrieval Methods

3.10.1.1 Results Comparison

In this Section, I compare the results from the preferred PLATON and AURA models. These include the fiducial model with partial clouds and Gaussian instrumental offsets for PLATON (Section 3.7.2.3) and the Gaussian instrumental offset model for AURA (Model 8; Section 3.9.2).

The similarities reveal the most robust conclusions of my analysis, since they are retrieved despite the many different assumptions that went into each method. Notably, both retrievals robustly find a metal-rich atmosphere with metallicity (defined as O/H) inconsistent with the solar metallicity at >2-σ. Both methods find a decisive (>4.8-σ) water vapor detection, and at least a moderate detection (>2.7-σ) of a non-haze gas absorption feature in the optical. Further, both PLATON and AURA retrievals are consistent with a mostly clear atmosphere, with neither finding strong evidence of haze or uniform, high-altitude grey clouds.

Though the atmospheric properties derived from PLATON and AURA are similar, there are noteworthy differences. AURA infers a cooler limb temperature at

100 mbar (1320^{+270}_{-200} K compared to 1710^{+100}_{-80} K for PLATON) as well as a lower metallicity of $\log_{10} Z/Z_{\odot} = 1.46^{+0.53}_{-0.68}$ compared to $\log_{10} Z/Z_{\odot} = 2.33^{+0.23}_{-0.25}$ for PLATON, a difference of 1.3-σ. This translates to $29^{+69}_{-23}\times \mathrm{Z}_{\odot}$ for AURA and $214^{+149}_{-88}\times \mathrm{Z}_{\odot}$. The optical absorber also differs: AURA determines the best description of the STIS feature to be absorption from sodium and AlO, whereas PLATON prefers some combination of TiO and VO absorption. Finally, AURA makes no claim on the C/O ratio as it is a free retrieval framework and no C-bearing species are detected nor meaningfully constrained. On the other hand, the chemical equilibrium assumption allows PLATON to find a 3-σ upper limit on the C/O ratio of C/O< 0.83.

To further contextualize the results, I added the functionality to retrieve the abundance profiles of relevant molecules in PLATON. I show abundance profiles for six spectroscopically relevant species from AURA which are also included in PLATON — H_2O, CO, CO_2, Na, TiO, and VO — in Figure 3.17. I emphasize the enforcing chemical equilibrium narrows the abundance constraints, and I am not reporting these abundances. Instead, they should be interepreted as the expected abundance profiles under the conditions of stable chemical equilibrium for the reported temperature, metallicity, and C/O ratio. As an example, I find no observational constraint on CO, but its abundance is well defined under chemical equilibrium for the temperatures and metallicities that I do observationally constrain via the water feature. Still, these profiles provide a useful baseline for comparison to free-chemistry retrieval abundances.

The abundance profiles for the optical-wavelength absorbers reflect the disagreement on the primary gas absorber: AURA prefers Na, and so it retrieves

$\sim 10\times$ more Na than PLATON and significantly less TiO and VO. Note that the decreasing TiO and VO abundance with increasing pressure for PLATON is due to those molecules condensing out of the atmosphere. PLATON's inferred water abundance is typically a few times greater than AURA's, reflecting the difference in inferred metallicities. CO_2 and CO are unconstrained by AURA, while PLATON finds a high abundance of CO_2 is consistent with the *Spitzer* observations. This difference is expected, given that PLATON finds weak evidence of CO_2 while AURA found none. Interestingly, this may relate to AURA's only chemical constraint, which is that CO2 must be less abundant than CO and H_2O due to the inferred temperatures (Section 3.6.2).

In total, AURA finds a cooler atmosphere with less oxygen but a large sodium enrichment to explain the optical absorption, while PLATON finds a hotter, higher oxygen-abundance atmosphere with TiO/VO absorption in the optical.

3.10.1.2 Impact of Retrieval Model Assumptions

The differences between a free-chemistry retrieval (AURA) and one constrained by chemical equilibrium (PLATON) are the natural result of the different assumptions made by each method. I therefore consider the PLATON and AURA retrievals to be two orthogonal analyses. I examine the impact of the differences by first explicitly listing the notable assumptions in each method, and then by providing the rationale for which assumptions are driving the differences.

The relevant methodological differences for PLATON as compared to AURA

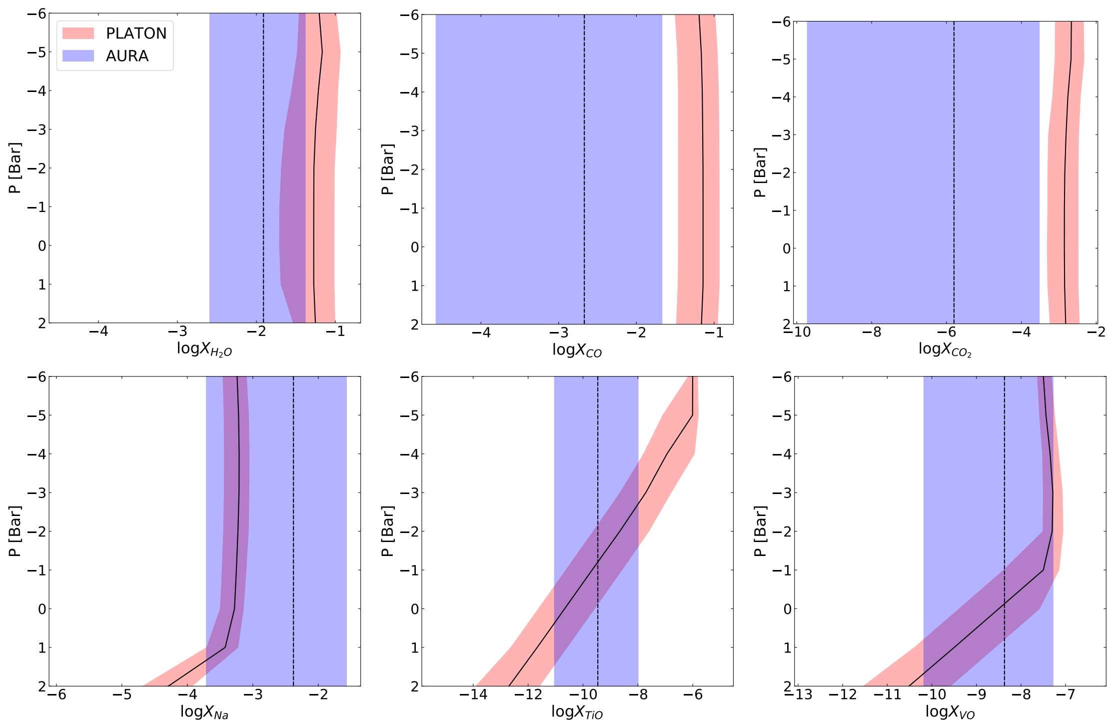


Figure 3.17: Abundance profiles for PLATON (red) and AURA (blue) for six relevant gaseous species. PLATON abundance profile distributions are derived by sampling the posterior 200 times, calculating the abundance profiles for each species for each sample, and finding the median value (solid black line) with 1-σ uncertainties at each pressure layer. AURA assumes abundances to be constant with pressure. The median retrieved value (dashed black line) and 1-σ uncertainty range are shown.

are 1) the assumption of chemical equilibrium, 2) fixing the elemental ratios between all metals other than carbon to their solar values, 3) assuming an isothermal profile for the atmosphere, 4) not including opacity from AlO, and 5) not including the Allard et al. (2019) H_2-broadened Na line profile.

I find that the Allard et al. (2019) H_2-broadened Na line profile is the key driver in the differences between the retrievals, and flexible element abundances and chemical equilibrium also play roles. AURA is the more flexible retrieval, so I first describe its solution before addressing why PLATON differs.

AURA's lower temperature solution is preferred for being able to explain the

H_2O feature in the WFC3 spectrum, while also explaining the STIS data with H_2-broadened Na absorption and capturing the *Spitzer* data. AURA is able to provide a fit to the STIS data by independently increasing the Na abundance and by also invoking AlO at relatively low temperatures. At this lower temperature ($T \sim 1300\,\mathrm{K}$), the amount of oxygen necessary for the water abundance and scale height to explain the observed water feature is about $29\times\,\mathrm{Z}_\odot$, with a mean molecular weight of about 2.7 AMU and a scale height of about 440 km.

Since PLATON has not yet incorporated the H_2-broadened Na line profile, the low temperature solution is a relatively poor fit to the STIS data. Instead, TiO/VO are needed to explain the STIS absorption feature, and these are only abundant enough in chemical equilibirum (with fixed metal ratios) at around 1650 K. At this higher temperature, a higher mean molecular weight is required for the same scale height, which must be small enough to explain the molecular feature sizes as well as the dominance of TiO/VO absorption over Rayleigh scattering. The atmospheric metallicity necessary to achieve the higher mean molecular weight is the much higher $\sim 200\times\,\mathrm{Z}_\odot$. Therefore, the differences make sense in light of the stricter assumptions.

To provide more support to this idea, I compare results with a those of a third retrieval method, ATMO (Amundsen et al., 2014; Tremblin et al., 2015, 2016, 2017; Sing et al., 2016), which acts as a middle ground between PLATON and AURA. ATMO's spectral retrievals can further help to gain insight into the effect of retrieval assumptions as it includes the Allard et al. (2007) pressure-broadened sodium line but also has the added flexibility of performing a free-element equilibrium-chemistry retrieval. With this assumption for the chemistry, the elemental abundances for each model are freely fit and calculated in equilibrium on the fly. Four elements were selected to vary independently, as they are major species which are also likely to be sensitive to spectral features in the data, while the rest were varied by a trace metallicity parameter ($[\mathrm{Z_{trace}/Z_{\odot}}]$). By separately varying the carbon, oxygen, sodium and vanadium elemental abundances ($[\mathrm{C/C_{\odot}}]$, $[\mathrm{O/O_{\odot}}]$, $[\mathrm{Na/Na_{\odot}}]$, $[\mathrm{V/V_{\odot}}]$) it allows for non-solar compositions but with chemical equilibrium imposed such that each model fit has a chemically-plausible mix of molecules given the retrieved temperatures, pressures and underlying elemental abundances.

The resulting retrieved atmospheric parameters describe an atmosphere most consistent with the one described by AURA. ATMO prefers a temperature of 1190^{+170}_{-120} K and a metallicity (as defined by the oxygen abundance) of $\log_{10} O/O\odot = \log_{10} Z/Z_{\odot} = 1.53^{+0.55}_{-0.67}$, in excellent agreement with AURA's values, and consistent with PLATON's metallicity to 1.3-σ, though the retrieved temperatures differ significantly. Like AURA, ATMO finds an enhanced sodium abundance, though uncertainties are large ($\log_{10} Na/Na\odot = 1.40^{+0.75}_{-1.80}$). This supports the idea that the inclusion of H_2-broadened sodium line profiles and the flexibility of non-solar metal ratios — and not necessarily the equilibirum chemistry constraint — allow for the low-temperature, lower oxygen abundance solution found by AURA. The metallicities on all three retrievals indicate a metal-rich atmosphere and agree at the $\sim$1.3-σ level.

Like PLATON (Section 3.8), ATMO also finds a subsolar C/O ratio (C/O = $0.17^{+0.53}_{-0.16}$ consistent with stellar (C/O = 0.19), though carbon is not well constrained so the uncertainties are large. The 3-σ upper limit of 0.94 is in good agreement with

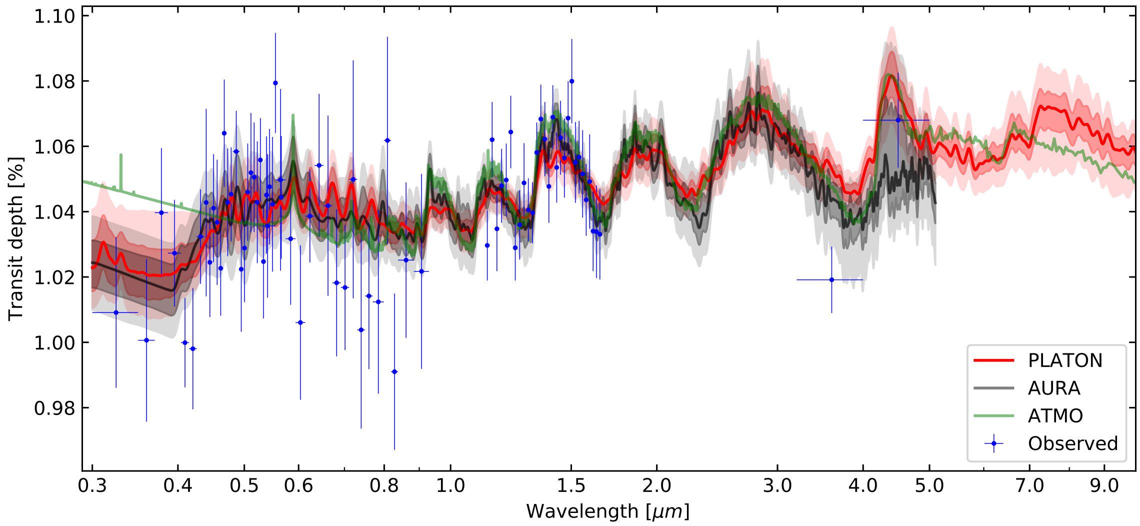


Figure 3.18: Comparison of the median retrieved model for each retrieval method's fiducial model. 1-σ and 2-σ uncertainty contours are included for PLATON and AURA. PLATON and AURA are smoothed with a Gaussian filter with σ = 15 for clarity. The chemical equilibrium assumption used by PLATON and ATMO allows for meaningful predictions at unobserved wavelengths, and so those models are shown out to 10 μm.

PLATON's 0.83 upper limit. However, unlike PLATON or AURA, ATMO finds no evidence of optical absorbers beyond Na, and instead prefers a haze and Na to explain the STIS optical data.

Figure 3.18 elucidates the differences in retrievals by showing the median retrieved fiducial model for PLATON (red), AURA (black), and ATMO (green) from 0.3–10 μm. The 1- and 2-σ uncertainty contours are shown for PLATON and AURA, both of which are smoothed with a Gaussian filter with σ = 15 for clarity. The AURA predictions are only shown up to 5 μm - as a free-chemistry retrieval, AURA retrieving on the 0.3–5 μm data does not place meaningful constraints on multiple molecules with significant opacity in the 5–10 μm range. Therefore, a prediction is not warranted.

While there are subtle differences, such as PLATON and ATMO's preference for CO_2 at $4.5\,\mu$m and sodium's prominence at $0.6\,\mu$m in the AURA and ATMO retrievals, the most obvious difference is below $0.5\,\mu$m, where ATMO prefers a haze instead of a metallic oxide feature. Though ATMO does not include AlO as an opacity source, this difference is likely due to different condensation schemes. PLATON uses GGchem's prescription (Woitke et al., 2018) such that species condense out when it is energetically favorable. AURA is a free-chemistry retrieval, so there are no restrictions on oxides being in the gaseous phase. ATMO, however, includes rainout chemistry (Goyal et al., 2019), such that if a species condenses at a higher pressure, that then depletes the element above that layer. It is plausible that although PLATON's condensation scheme allows TiO/VO to be in the gas phase around 1700K, ATMO's scheme does not, making the metallic oxide feature difficult to capture.

In total, I tentatively favor AURA's derived atmospheric parameters over PLATON's, for two main reasons. First, the inclusion of the most up-to-date sodium line profiles and AlO opacity impact the retrieval. Second, constraints from interior modeling (Section 3.10.2), though not necessarily decisive, are consistent with AURA and in tension with PLATON. Overall, this paints a picture of an atmosphere with a supersolar — but not necessarily superstellar — metallicity, sodium enrichment, possible disequilibrium metallic oxides (e.g., circulated from dayside, dredged up due to vertical mixing), and a planet with a well-mixed interior and a limb temperature lower than the equilibrium temperature.

3.10.2 Comparison to Interior Modeling Metallicity Constraints

Though they both describe metal-rich atmospheres, the 1-σ retrieved atmospheric metallicities ranges from AURA and PLATON are inconsistent ($\log_{10} Z/Z_{\odot}$ = 0.78–1.99 and $\log_{10} Z/Z_{\odot}$ = 2.08–2.56, respectively). Further, it is questionable whether such supersolar metallicities — especially those retrieved by PLATON – are physically reasonable. I check the viabilitiy of these values by comparing them to atmospheric metallicity constraints from interior structure models.

Thorngren & Fortney (2019) demonstrated how interior models can constrain atmospheric metallicity. Essentially, this is a three step process: 1) Determine what range of bulk metallicities are necessary for structure models to explain the observed radius, taking into account the planet's mass, age, heating efficiency, and parameter uncertainties, 2) set the maximum bulk metallicity to be the 3-σ upper limit of the derived posterior distribution, and 3) set the maximum atmospheric metallicity to be equal to the maximum bulk metallicity.

The third step assumes that the atmospheric metallicity cannot be greater than the core's metallicity for significant timescales due to convection or Rayleigh-Taylor instability. They define metallicity as the ratio of all metals to hydrogen compared to the ratio in the Sun's photosphere. This is a good proxy for O/H, and so it is a valid comparison to the retrieved atmospheric metallicities. For more details on the derivation, see Thorngren & Fortney (2019).

Using stellar parameters from Hartman et al. (2012) (Table 3.3), the interior structure model fit yields a bulk metal abundance ratio of $\mathrm{Z}/\mathrm{Z}_{\odot} = 33.7 \pm 9.1$, corresponding to a maximum atmospheric metallicity of 50$\times$ $Z_\odot$ (D. Thorngren, private communication). There is no significant uncertainty on this number, as it is the 3-σ upper limit of the distribution. This is consistent with the metallicity from the AURA retrieval, but it is in tension with PLATON's retrieved metallicity — 50$\times$ $Z_\odot$ falls outside PLATON's 1-σ range (but within 2-σ, as the metallicity distribution is asymmetric PLATON's 2-σ lower limit is 37$\times Z_\odot$). This could indicate that the "true" atmospheric metallicity falls in the lower range of PLATON's retrieved metallicity, or it could be interpreted as slight evidence in support of AURA over PLATON. Either way, the atmospheric metallicity approaching the bulk metallicity indicates a well-mixed interior. Such vertical mixing could allow for micron-sized particles to stay afloat in the atmosphere, potentially facilitating gaseous metal oxide survival and Mie scattering.

3.10.3 Implications for Planet Formation

The atmospheric metallicity I retrieve for HAT-P-41b provides important constraints on the formation and migration history of the planet. At the outset, the super-solar metallicity (O/H) of $\sim$30–200$\times$ $Z_\odot$ requires substantial accretion of solids, beyond several Earth masses of H_2O ice, during the planet's evolutionary history. It is unlikely that such a large amount of volatile accretion is possible at the planet's current orbit. Therefore, the planet is unlikely to have formed in-situ (Batygin et al., 2016) but instead formed far out beyond the H_2O snow line and migrated inward. The formation location and migration path of a giant planet can significantly affect its chemical composition. Beyond the H_2O snow line the gas in the protoplanetary disk is depleted in oxygen whereas the solids are enriched in oxygen (Öberg et al., 2011). Therefore, planets with high enrichment of oxygen require predominant accretion of H_2O-rich planetesimals while forming and migrating through the protoplanetary disk.

The high metallicity (specifically O/H) of HAT-P-41b, therefore, supports the migration of the planet through the disk via viscous torques (Madhusudhan et al., 2014b). This is in contrast to other hot Jupiters with low O/H abundances which have been suggested to be caused by insufficient solid accretion, e.g. via disk-free migration (Madhusudhan et al., 2014b) or formation via pebble accretions whereby the oxygen-rich solids are locked in the core (Madhusudhan et al., 2017). The fact that HAT-P-41b's orbit is moderately misaligned to the host star's rotation axis is also in tension with the disk migration hypothesis, since spin-orbit misalignments are considered to be evidence of disk-free migration and planet-planet interactions (Winn et al., 2010). In principle, instead of disk migration, super-solar elemental abundances could be caused by accreting gas whose metallicity has been enhanced due to pebble drift (Öberg & Bergin, 2016; Booth et al., 2017). But while pebble drift can cause metal enhancements up to $\sim 10\times$ $Z_\odot$, much larger enhancements as constrained in the present case are unlikely to be explained by this process. More importantly, such enhancements due to pebble drift are also expected to cause high C/O ratios ($\sim$1), which may be at odds with the high H_2O abundace and the low C/O ratio retrieved for the planet.

Overall, the most plausible explanation for the potentially high atmospheric metallicity inferred for HAT-P-41b is formation outside the H_2O snowline and migration inward while accreting substantial mass in planetesimals. If confirmed, this would be a departure from other hot Jupiters observed hitherto which have generally shown low H_2O abundances, indicative of the low accretion efficiency of H_2O-rich ices that is possible for disk-free migration mechanisms (Madhusudhan et al., 2014b; Pinhas et al., 2019; Welbanks et al., 2019). Such an abundance is also a substantial departure from expectations based on Solar System giant planets. The metallicity of Jupiter in multiple elements is $\sim$1–5$\times$ $Z_\odot$ Atreya et al. (2016); Li et al. (2020). With the mass of HAT-P-41 b being similar to that of Jupiter, its higher metallicity would indicate an even higher amount of solids accreted than that of Jupiter in the Solar System.

3.11 Summary

I have conducted a comprehensive, multi-pronged Bayesian retrieval analysis of the 0.3–5 μm transit spectrum of HAT-P-41b derived from HST STIS (previously unpublished; Section 3.5.1), HST WFC3 (re-analysis; Section 3.5.2), and *Spitzer* (independent analysis; Section 3.5.3) transit observations. I determined the host star has, at most, a low level of stellar activity ($\log L_\mathrm{X}/L_\mathrm{bol} < -5.2$) using both visible and X-ray photmetric monitoring observations (Section 3.4.1).

We performed two complementary retrieval analyses: a relatively strict PLATON analysis (Section 3.6.1, Section 3.8) assuming chemical equilibrium and solar metal ratios (except carbon), and a more flexible AURA free-chemistry retrieval

(Section 3.6.2, Section 3.9.1). Both methods' fiducial models are excellent fits to the entire transit spectrum. I further tested an array of more complicated models (Sections 3.7 and 3.9.2), including instrumental transit depth biases (offsets), parametric rayleigh scattering, partial cloud coverage, Mie scattering (PLATON only), and stellar activity (PLATON only). I find the conclusions to be insensitive to model choice within a paradigm.

Despite PLATON and AURA's differing model assumptions, priors, and even opacity sources, I find several shared conclusions between the two methods (Section 3.10.1). Both PLATON and AURA retrieve a high atmospheric metallicity (O/H) that is inconsistent with $Z_\odot$ to greater than 2-σ ($\log_{10} Z/Z_\odot = 1.46^{+0.53}_{-0.68}$ compared to $\log_{10} Z/Z_\odot = 2.33^{+0.23}_{-0.25}$, respectively). They also both are consistent with a haze-free and cloud-free atmosphere, and both find a decisive water vapor detection and at least suggestive evidence of an optical absorption feature. We further confirm the result by performing a middle-ground retrieval, ATMO, and find results generally consistent with AURA's (Section 3.10.1.2). I determine the inclusion of H_2-broadened sodium opacity impacts the retrieved metallicities. While I consider AURA to be more physically plausible due to its consistency with interior modeling constraints and inclusion of H_2-broadened sodium opacity, I present the results from both PLATON and AURA as assumption-dependent orthogonal analyses. Overall, this study emphasizes the importance of comparative retrievals with different forward modeling, prior, and model selection assumptions in order to best contextualize presented results.

3.12 Addendum

A separate group's analysis of HAT-P-41b was released concurrently with mine (Sheppard et al., 2021), releasing on arXiv simultaneously. Lewis et al. (2020) used UVIS observations (analyzed in the Part I paper Wakeford et al. (2020)) in conjunction with their own derived *Spitzer* and WFC3 transit depths to conduct transit spectroscopy on HAT-P-41b. Their UVIS observations and analysis are the first application of UVIS to exoplanet transit spectroscopy. Similar to AURA, PLATON, and ATMO retrievals, their retrievals needed to invoke something to explain the unexpectedly small size of the clear water feature in the WFC3 data. They perform several retrievals and fin their UVIS and WFC3 data is best described by an orders-of-magnitude overabundance of H- as compared to expectations in equilibrium chemistry, which they argue is plausible via photochemistry (Lavvas et al., 2014). Their median metallicity is superstellar, and interestingly find a water abundance consistent with the AURA retrieval but convert from H_2O to O/H differently causing the inferred metallicities to disagree by a factor of two. This type of disequlibrium process is not capturable by my PLATON retrieval, but even for a free-chemistry AURA retrieval with H- opacity we ran we did not detect H-. We emphasize that both PLATON and AURA fit the data with reduced χ^2 values consistent with one. Additionally, Espinoza & Jones (2021) applied a version of CHIMERA (assumes chemical equilibrium, Line et al., 2013) to the exact data in Sheppard et al. (2021) and retrieved a superstellar metallicity of 100× solar ($\log Z/Z_\odot \sim 2.00 \pm 0.30$, consistent with the PLATON retrievals. I conclude that the differences in conclusions are due to UVIS-STIS data differences, and not modeling differences.

Chapter 4: Constraining the Dayside Thermal Structure of Hot Jupiters from Secondary Eclipse Observations

4.1 Overview

In this chapter, I derive the HST WFC3 emission spectra for two highly irradiated hot Jupiters (WASP-18b and WASP-19b) and retrieve their atmospheric properties. Most notably, I find evidence for a strong thermal inversion in the dayside atmosphere of the highly irradiated hot Jupiter WASP-18b ($T_{eq} = 2411K$, $M = 10.3M_J$) based on emission spectroscopy from Hubble Space Telescope secondary eclipse observations and *Spitzer* eclipse photometry. I demonstrate a lack of water vapor in either absorption or emission at $1.4\,\mu$m. However, I infer emission at $4.5\,\mu$m and absorption at $1.6\,\mu$m that I attribute to CO, as well as a non-detection of all other relevant species (e.g., TiO, VO). The most probable free-chemistry atmospheric retrieval solution indicates a C/O ratio of 1 and a high metallicity (C/H=$283^{+395}_{-138}\times$ solar). However, water dissociation and H- opacity could explain the spectrum without necessitating a super-solar metallicity. The derived composition and T/P profile suggest that WASP-18b is the first example of a planet with a non-oxide driven thermal inversion. I find moderate evidence (2.8σ) of water absorption with non-depleted abundance (log X_{H_2O}=-3.64$^{+1.44}_{-0.72}$; consistent with stellar) and a likely sub-stellar C/O (C/O<0.63), consistent with transit analyses. I also retrieve a non-inverted T-P profile in WASP-19b, which, at an equilibrium temperature of 2120 K, is at the border of where both TiO-driven inversions and water dissociation in hot Jupiter atmospheres are expected to become important.

4.2 Introduction

Hot Jupiters have been vital in revealing the structural and atmospheric diversity of gas-rich planets (see recent reviews by Crossfield, 2015; Madhusudhan et al., 2016; Deming & Seager, 2017). Since they are exposed to extreme conditions and relatively easy to observe through transit and eclipse spectroscopy, hot Jupiters provide a window into a unique part of parameter space, allowing us to better understand both atmospheric physics and planetary structure.

An outstanding question that has emerged for highly irradiated planets is the presence and origin of stratospheric thermal inversions, which have been detected in several extremely irradiated hot Jupiters (Haynes et al., 2015; Evans et al., 2017). Hubeny et al. (2003) predicted that thermal inversions in highly-irradiated atmospheres would be caused by the presence of optical absorbers (e.g. TiO and VO) high in the atmosphere, but there may be other causes such as insufficient cooling (Mollière et al., 2015) or sulfur-based aerosols (Zahnle et al., 2009b).

It is also unclear what conditions are necessary for TiO or VO to exist in the gaseous form and drive inversions. There is reason to expect a correlation with planet temperature (cold traps, dissociation) and gravity (vertical cold traps), but observational evidence is limited and the exact temperatures and gravities where these processes dominate is unclear (Parmentier et al., 2013; Beatty et al., 2017; Parmentier et al., 2018). H- opacity and molecular dissociation is expected to impact the thermal structure of ultra-hot Jupiters and mask water features (Parmentier et al., 2018; Lothringer et al., 2018), though the magnitude and prevalence of this effect is uncertain.

Huitson et al. (2013) originally observed WASP-19b in transit with WFC3 and STIS (optical) and found a clear water detection, no evidence of TiO/VO, and likely a low C/O ratio. This was confirmed by Sing et al. (2016), who also found no evidence of CO or CO_2 in the *Spitzer* transit data. Additionally, Benneke (2015) reported a water abundance of 0.2–30x solar ($\log X_{H_2O}$=-$3^{+1.2}_{-1.0}$). However, Sedaghati et al. (2017) observed WASP-19b with VLT and claimed a 7σ detection of TiO in transit (and 7.5σ detection of water), in strong contrast with previous results. Espinoza et al. (2019) then challenged this claim, showing that observations over a similar optical range with the ground-based telescope Magellan/IMACS found no evidence of TiO. Finally, Sedaghati et al. (2021) followed up the planet with high-resolution cross-correlation spectra from VLT/ESPRESSO, finding a "barely significant" peak of TiO. Since TiO is predicted to drive thermal inversions, measuring the thermal profile of WASP-19b can provide insight into the likelihood of TiO in the atmosphere. Though photometric and ground-based secondary eclipse observations exist (e.g, Anderson et al., 2013), no high quality space-based spectroscopic eclipse observations have been analyzed.

Constraints on the structure and composition of exoplanetary atmospheres allow us to test, refine, and generalize planetary formation models. Volatile ices are expected to play an important role in planet formation; thus a constraint on the composition of a hot planet's atmosphere gives us insight on how and where it was formed (Öberg et al., 2011; Madhusudhan et al., 2014a). In our Solar System there is an inverse mass vs. atmospheric metallicity relationship, and whether or not it extends to exoplanets is informative to planetary formation and migration models. There is some evidence that the trend holds (Kreidberg et al., 2014), however that parameter space is not yet sufficiently populated to enable firm conclusions.

In this chapter I use Hubble Space Telescope (HST) spectroscopy and *Spitzer* IRAC photometry of secondary eclipses to explore the thermal structure and composition of the dayside atmosphere of two hot Jupiters. First I analyze WASP-18b, an extremely hot ($T_{eq} = 2411$K) and massive ($M = 10.3\mathrm{M_{Jup}}$) hot Jupiter orbiting an F-type star with an orbital period of less than one day (Hellier et al., 2009). Then I analyze WASP-19b ($T_{eq} = 2120$K, M=1.11$\mathrm{M_{Jup}}$), a hot Jupiter orbiting an active G8-star, also with a period of less than a day. WASP-19b is particularly interesting given that its gravity and temperature place it on the boundaries of the parameter space where processes such as water dissociation and TiO cold traps are expected to impact the chemistry and thermal structure of hot Jupiter atmospheres.

4.3 Observations

I used Wide Field Camera-3 (WFC3) observations of five secondary eclipses of WASP-18b from the HST Treasury survey by Bean et al. (Program ID 13467). WFC3 obtains low resolution slitless spectroscopy from 1.1 to 1.7 μm using the G141 grism (R=130), as well as an image for wavelength calibration using the F140W filter. Grism observations were taken in spatial scan mode (Deming et al., 2013) with a forward-reverse cadence (Kreidberg et al., 2014). The first three visits, taken between April-June 2014, are single eclipse events. Visit 4, taken in August 2014, contains two eclipses in an orbital phase curve, and I extract those eclipses and analyze them separately. I also extract a single eclipse of WASP-19b from the HST program GO-13431 phase curve observation (PI: C. Huitson), also taken in spatial scan mode with a uni-directional reverse cadence.

A collaborator re-analyzes two eclipse observations of WASP-18b taken in the 3.6 μm and 4.5 μm channels of the *Spitzer Space Telescope*'s IRAC instrument (Program ID 60185). The 3.6 μm observation was performed on 2010 January 23, while the 4.5 μm observation was taken 2010 August 23. Both observations were taken using an exposure time of 0.36s in subarray mode, and were first analyzed in Maxted et al. (2013). *Spitzer* eclipse depths for the WASP-19b are taken from Garhart et al. (2020). Depths from cold *Spitzer* (IRAC3 and IRAC4) are taken from Nymeyer et al. (2011, WASP-18b) and Anderson et al. (2013, WASP-19b).

4.4 *HST* Data Analysis

The data analysis is an earlier version of the `DEFLATE` analysis described in Section 2.3. I briefly summarize it here. My grism spectroscopy analysis utilized HST "ima" data files. I separated the data by scan direction, removed background flux, and corrected for cosmic rays and bad pixels. I removed background flux via the "difference frames" method outlined in the appendix of Deming et al. (2013), and set the aperture to maximize the amount of source photons in my analysis. All corrections are propagated to the flux errors, which are retrieved from the "ima" error extension and intrinsically account for read noise and bias. The end result is two reduced light curves - one forward scan and one reverse scan - for each eclipse, which I analyze separately.

The F140W photometric image determines the location of the zero-point, which I used to assign a wavelength to each column. I confirmed the wavelengths by fitting an appropriate ATLAS stellar spectrum (e.g, for WASP-18b, T=6400K, $\log g$=4.3, [Fe/H]=0.1) (Castelli & Kurucz, 2004), multiplied by the grism sensitivity curve, to an observed in-eclipse spectrum.

I note that the archival data on WASP-19b is technically from a "failed" phase curve observation. This is because the visit relied on gyro guiding instead of FGS due to an issue with the guide star, resulting a large horizontal shift over the course of the observation that is difficult to deconstruct. However, that is for the entire days-long phase curve. Over the course of a 6-hour eclipse, the shift is almost negligible, amounting to almost a pixel over the course of the observation. This sub-pixel shift

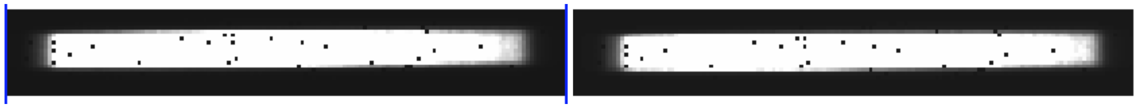


Figure 4.1: First and last exposures of WASP-19b eclipse observation. The colorscale is exaggerated to make the location of bad pixels (black), which are constant on the detector, more visible. The detector shift during eclipse is no more than a single pixel, which is not significant enough to impact the quality of the data.

is easily accounted for in systematic modeling. Further, each wavelength bin is 6 pixels, and in low-resolution spectroscopy trying to capture molecular features, such shifts have minimal impact. For a 6 pixel bin size, a "typical" water feature has a width on the order of ~42 pixels. Thus, these eclipses still contain useful information and our worth retrieving on.

4.4.1 Light Curve Analysis

The light curve analysis is the root of what would become `DEFLATE` (described in Section 2.4). I summarize it below on WASP-18b, which was a cominbation of four observations. I used the same process on WASP-19b.

Empirical methods are necessary to correct for non-astrophysical systematic effects in WFC3 spectroscopy (Berta et al., 2012; Haynes et al., 2015). Correction methodology is especially important in emission spectroscopy, where the magnitude of systematic effects can be greater than the eclipse depth (Kreidberg et al., 2014). I thereby combined two strategies: initial removal of systematic trends using parametric marginalization (Gibson, 2014a; Wakeford et al., 2016b), and further detrending by subtraction of scaled band-integrated residuals from wavelength bins (Mandell et al., 2013; Haynes et al., 2015). This method accounts for uncertainty in instrument model selection, and residuals from the band-integrated analysis allow us to utilize the normally excluded first orbit of each HST data set in the spectroscopic analysis.

Fitting a band-integrated light curve provides residuals that I use to remove unidentified systematics from the spectrally resolved light curves. I calculate the HST phase (parameter for ramp and HST breathing), planetary phase (parameter for visit-long slope), and a wavelength shift derived by cross correlating each spectrum with the last spectrum for the visit (parameter for jitter) for each exposure in a time series. The grid of systematic models comprises a combination of a linear planetary phase correction and up to four powers of HST phase and wavelength shift. These models are then multiplied by a Mandel & Agol (2002) eclipse model. I simultaneously fit for the eclipse depth, all systematic coefficients, and - for two light curves with ingress and egress points - the center of eclipse time. All other system parameters are fixed to literature values. For WASP-19b, which has good egress coverage, I also fit for a/R_{star}. All other system parameters are fixed to literature values (Hebb et al., 2010; Tregloan-Reed et al., 2013; Wong et al., 2016; Sedaghati et al., 2017).

I use a Levenberg-Markwardt (L-M) least squares minimization algorithm (Markwardt, 2009) to determine the parameter values. An example band-integrated light curve with systematic effects removed using the best-fitting model is shown in the leftmost panel of Figure 4.2. For WASP-18b, the scatter (RMS) of the residuals of the band-integrated curves ranges from 1.3-5.5× the photon noise, indicating that there is excess noise beyond the photon limit present. Excess noise in the band-integrated curves is also shown by comparisons of the cumulative distributions of residuals with those of a photon-limited Gaussian (see bottom-left panel of Figure 4.2). However, the structure of this excess noise does not change with wavelength, allowing for its removal from the corresponding spectral light curves.

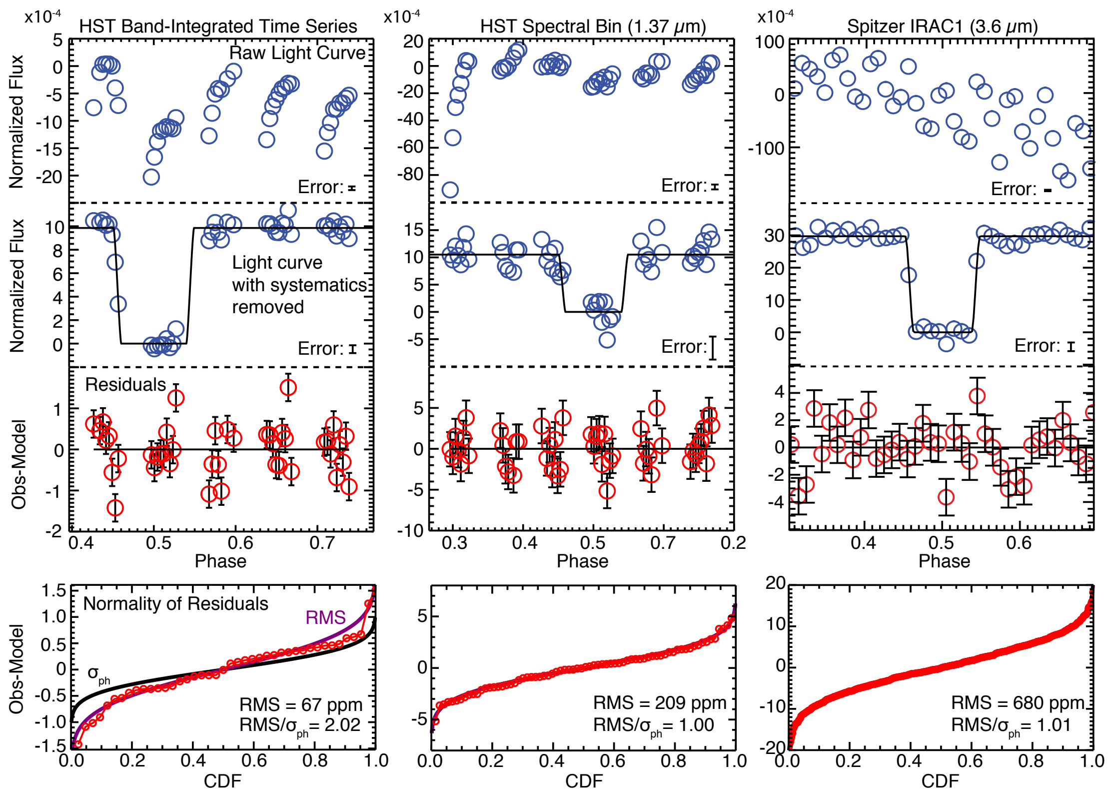


Figure 4.2: An example of the detrending process for an HST band-integrated light curve (left), a light curve for an HST spectral bin (middle), and a *Spitzer*/IRAC photometry light curve (binned for clarity). The HST band-integrated results fall within $1.3-5.5\times$ the photon noise limit, while both the HST spectral bins and the *Spitzer* data typically achieve close-to-photon-limited results. The bottom row compares the cumulative distribution function (CDF; red dots) of the residuals to that of a Gaussian with dispersion equal to the photon noise (black line). Good agreement is obtained for the HST spectral and Spitzer residuals, while excess scatter is observed for the HST band-integrated residuals. For the latter, the CDF of a Gaussian with dispersion equal to the residual RMS is also plotted for comparison.

To derive the emission spectrum, I bin the exposures in wavelength between the steep edges of the grism response and fit these spectrally resolved light curves. I remove wavelength-dependent systematics by fitting each spectral bin separately in a process that mimics the band-integrated process, with three exceptions. First, the eclipse mid-time is now fixed to the value determined by the band-integrated analysis. Second, it is possible that shifts on the detector are wavelength-dependent, so the jitter parameter is recalculated for each wavelength bin using only that portion of the spectrum in the cross-correlation procedure. Third, each systematic model now incorporates the residuals from the band-integrated fit of the same model as a decorrelation variable. The amplitude of the residuals is a free parameter, although the shape is assumed to be constant in wavelength. This removes any remaining wavelength-independent trends in the data. An example result of a reduced spectral bin light curve is shown in the central panel of Figure 4.2.

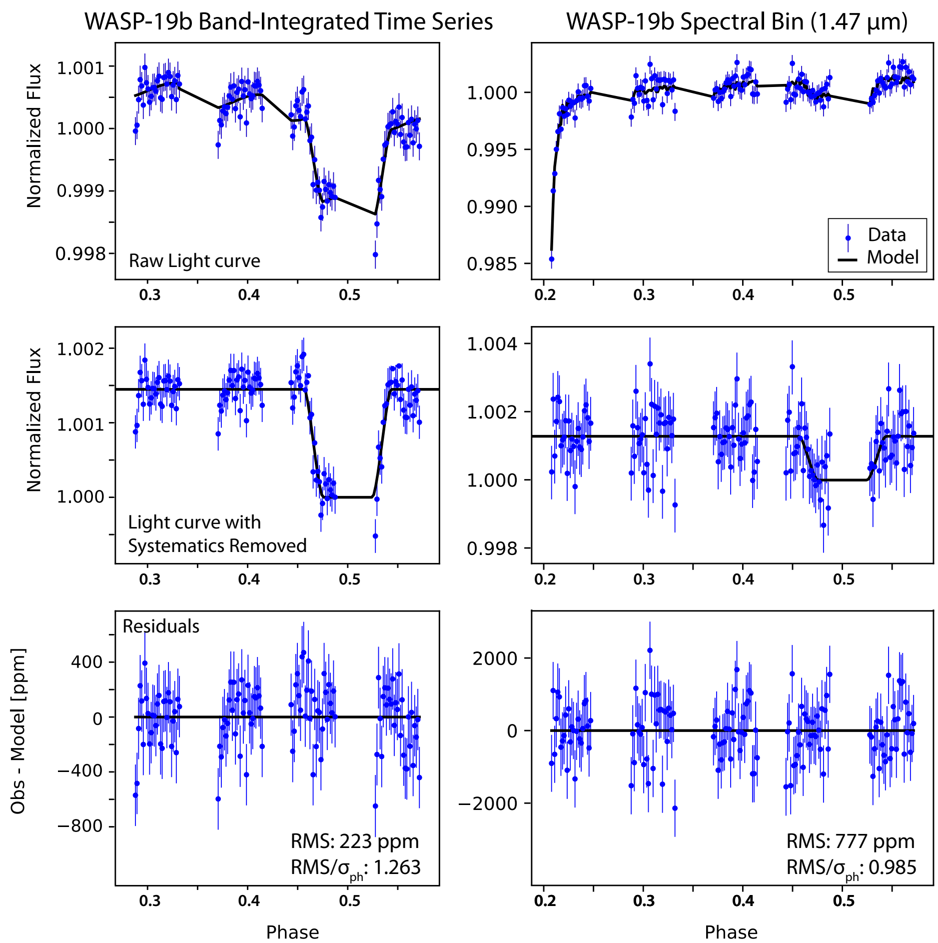


Figure 4.3: Band-integrated light curve (left) and a sample spectral bin light curve (right) for HST WFC3 observations of WASP-19b. The top panels show the raw data with the best-fit model overplotted, and the middle panels show the light curve and model after detrending systematic effects. The bottom panels show the difference between the observations and the best-fit model at each point in the time series and give the standard deviation of those residuals (RMS).

Finally, eclipse depths from the multiple visits are combined via an inverse-variance weighted mean, giving the emission spectrum for WASP-18b. The spectra for all visits are shown in Figure 4.4.

For WASP-18b, the average RMS of the systematic-reduced spectroscopic light curves is 1.04× the photon noise and the median RMS is 0.97× the photon noise,

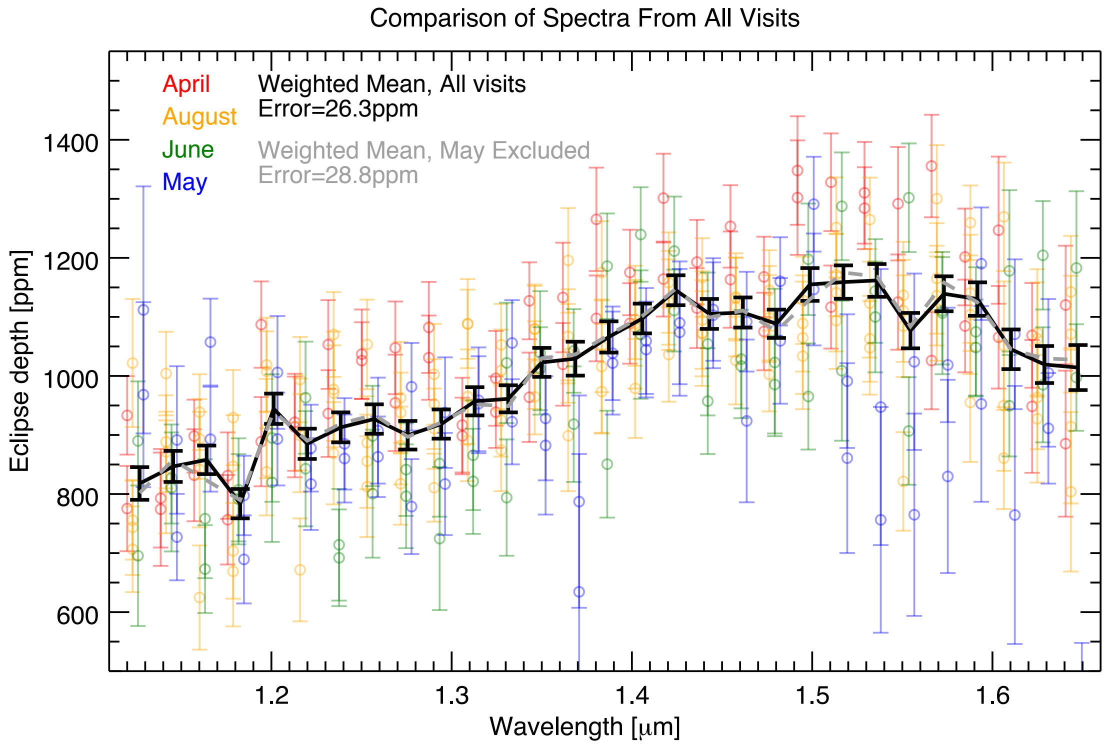


Figure 4.4: Spectra for all of the HST visits, horizontally offset for clarity, with the weighted mean overplotted. Depths from both the forward and reverse scan light curves are plotted for each eclipse. The May data receives a low weight due to the large uncertainties, and therefore does not impact the results beyond the individual uncertainties, as shown by the dashed grey line. Values for the individual data points are available from the authors upon request.

indicating that shot noise is the dominant error source. The close agreement between the cumulative distributions of residuals and those of a Gaussian with a width determined by the photon noise provides further evidence that the analysis achieved photon-limited results for the vast majority of spectral curves (see bottom-center panel of Figure 4.2).The remaining spectral curves have residuals with an RMS greater than 1.5× the photon limit, indicating that excess noise is present. These only constitute 6% of all spectral bins, and every one is from the single eclipse observation taken in May. I explored removing the May dataset due to this increased noise, but the exclusion of these data did not affect the variance-weighted spectrum, and I chose to include this visit in subsequent analyses. Figure 4.4 contains the emission spectra from every visit, demonstrating the consistency of the structure of the spectrum. My analysis routine finds that the outlier depths from the May visit have very high errors due to the presence of correlated noise, and so they contribute very little to the weighted spectrum.

For WASP-19b the exposures are separated into wavelength bins six pixels (0.028 μm) in size and fit individually to derive the emission spectrum. The RMS of the residuals of the band-integrated light curve is 1.26x the theoretical photon-noise limit, which indicates there is some excess noise present. Further, there is an indication of red noise in the white light curve. An underestimation of the ramp systematic seems to cause a majority of points in the third orbit (bottom left panel of Figure 4.3) to be positive residuals. However, the average and median RMS of the residuals of the spectroscopic light curves are 1.02x and 1.01x the theoretical limit, indicating that no additional noise is present in the binned data after subtracting the band-integrated residuals (see right half of Figure 4.3). An analysis of the RMS of the residuals as a function of binning in time further demonstrates that shot noise is the dominant error source (see Figure 4.5). This indicates that the structure that the model grid was unable to capture in the white light model was successfully used to remove the same structure in the spectral bins. The relative shape of the spectrum is then reliable, only the absolute depth is suspect. I followed up on this by deriving the spectrum independently with the physical charge-trap model from Zhou et al. (2017). Given that the ramp is the dominant systematic in this light curve, this model is an excellent fit to the white light curve, finding a depth about 75ppm deeper than marginalization. However, the derived spectra are in excellent agreement. I emphasize that the emission spectrum is not dependent on methodology.

To further check my methodology, I reanalyzed published emission spectra for WASP-43b (Kreidberg et al., 2014b), WASP-103b(Cartier et al., 2017), and WASP-121b (Evans et al., 2017). I find an agreement to the published spectra, with a mean point-by-point variation (difference / uncertainty) of 89%, 23%, and 50% for the three data sets, respectively, demonstrating the consistency of my analysis pipeline with those published by other authors.

4.5 *Spitzer* Re-analysis

Spitzer secondary eclipse measurements of WASP-18b were reported by Maxted et al. (2013), and a collaborator has re-analyzed key portions of those data. We confine our re-analysis to the 3.6 and 4.5 μm bands, because the instrumental systematic errors are greatest in those bands, and there are new methods to correct those systematics.

We use an updated version (Tamburo et al., 2017) of the Pixel-Level Decorrelation framework (Deming et al., 2015). Our photometry uses 11 different circular aperture sizes (with radii ranging from 1.6 to 3.5 pixels). We decorrelate the instrumental systematics while simultaneously fitting for the eclipse depth, using binned data, as advocated by Deming et al. (2015) and Kammer et al. (2015). The fitting code selects the optimal aperture and bin size, and obtains an initial estimate of the

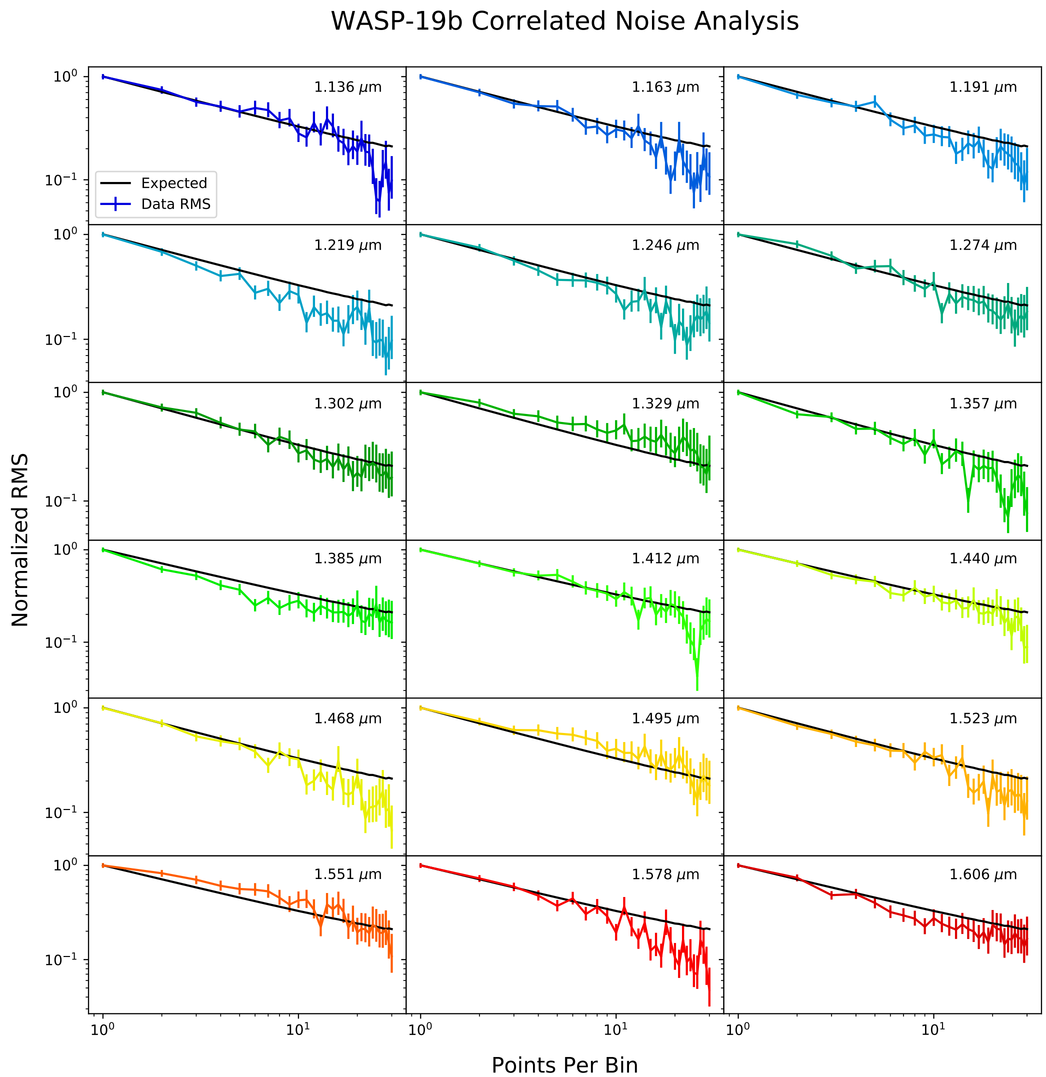


Figure 4.5: Analysis of the temporal correlated noise in each spectral bin for WASP-19b. The data is binned up in time and the RMS of the light curve residuals is calculated; the results are then normalized by the RMS of the light curve with minimal binning (i.e, one point per bin) and compared with the predicted trend assuming there is no correlation in time ($RMS_0/\sqrt{N}$), where N is the number of exposures per bin). The light curves show no evidence of correlated noise; the most extreme deviations (bin 1.329 μm) are consistent with white noise within 1-2σ.

eclipse depth and the pixel basis vector coefficients using linear regression. We then implement an MCMC procedure (Ford, 2005) to explore parameter space, refine the best-fit values, and determine the errors. At each step, we allow the central phase, orbital inclination, and eclipse depth to vary, but lock all other orbital parameters to the values used in the WFC3 analysis. We also vary the multiplicative coefficients of our basis pixels (see Deming et al., 2015) and visit-long quadratic temporal baseline coefficients at every step. Our best fits use aperture radii of 2.0 and 2.5 pixels, and bin sizes of 76 and 116 points at 3.6 and 4.5 μm, respectively. The scatter in the binned data, after removal of the best-fit eclipse, is 1.01 and 0.95$\times$ the photon noise at 3.6 and 4.5 μm, respectively, those ratios being statistically indistinguishable from unity.

We ran three chains of 500,000 steps for both bands, confirming their convergence through the Gelman-Rubin statistic (Gelman & Rubin, 1992). We combine all chains of eclipse depth into a unified posterior distribution for each band, and fit a Gaussian to this distribution to determine the error on eclipse depth. Our results are included in Table 4.1, and exhibit excellent agreement with Maxted et al. (2013), but with smaller errors.

4.6 WASP-18b

4.6.1 Atmospheric Retrieval

I use the WFC3 spectrum along with the Spitzer and ground-based Ks band photometry to constrain the composition and temperature structure of the dayside

Table 4.1: WASP-18b Thermal Emission Spectrum

Instrument	λ [μm]	Depth [ppm]	Instrument	λ [μm]	Depth [ppm]
WFC3 G141	1.118–1.136	818 ± 28		1.434–1.452	1105 ± 25
	1.136–1.155	847 ± 26		1.452–1.471	1107 ± 25
	1.155–1.173	858 ± 24		1.471–1.489	1088 ± 24
	1.173–1.192	784 ± 25		1.489–1.508	1155 ± 28
	1.192–1.211	944 ± 26		1.508–1.527	1159 ± 28
	1.211–1.229	885 ± 26		1.527–1.545	1162 ± 28
	1.229–1.248	913 ± 25		1.545–1.564	1077 ± 30
	1.248–1.266	927 ± 25		1.564–1.582	1139 ± 30
	1.266–1.285	900 ± 24		1.582–1.601	1130 ± 28
	1.285–1.304	919 ± 25		1.601–1.620	1045 ± 34
	1.304–1.322	957 ± 24		1.620–1.638	1019 ± 31
	1.322–1.341	961 ± 23		1.638–1.657	1014 ± 38
	1.341–1.359	1022 ± 25	IRIS2 K_s	2.0–2.3	1300 ± 300[a]
	1.359–1.378	1029 ± 29	*Spitzer* IRAC1	3.2–4.0	2973 ± 70
	1.378–1.396	1066 ± 26	*Spitzer* IRAC2	4.0–5.0	3858 ± 113
	1.396–1.415	1097 ± 25	*Spitzer* IRAC3	5.0–6.4	3700 ± 300[b]
	1.415–1.434	1145 ± 25	*Spitzer* IRAC4	6.4–9.6	4100 ± 200[b]

NOTE—WFC3 bin size = 0.0186μm

[a] Anglo-Australian Telescope (Zhou et al, 2015)

[b] Nymeyer et al, 2011

atmosphere of WASP-18b. We use the HyDRA retrieval code (Gandhi & Madhusudhan, 2018), which comprises a thermal emission model of an atmosphere coupled with a nested sampling algorithm for Bayesian inference and parameter estimation. The forward model, based on standard prescriptions for retrieval (Madhusudhan & Seager, 2009; Madhusudhan et al., 2011), computes line-by-line radiative transfer in a plane parallel atmosphere under the assumptions of hydrostatic equilibrium and local thermodynamic equilibrium. The pressure-temperature (P-T) profile and chemical compositions are free parameters in the model.

The model includes 14 free parameters. For the P-T profile, we use the parametrisation of (Madhusudhan & Seager, 2009) which involves six free parameters. The atmosphere comprises 100 layers equally spaced in log-pressure between 10^{-6} bar and 10^{2} bar. For the atmospheric composition we consider several species expected to be prevalent in very hot Jupiter atmospheres and with significant opacity in the observed spectral range (Madhusudhan, 2012; Moses et al., 2013; Venot & Agúndez, 2015). This includes H_2O, CO, CH_4, CO_2, HCN, C_2H_2, TiO, and VO. The uniform mixing ratio of each species are free parameters in the model. We assume an H_2/He rich atmosphere with a solar He/H_2 ratio of 0.17. We consider line absorption from each of these species and collision-induced opacity from H_2-H_2 and H_2-He. The sources of opacity data are described in Gandhi & Madhusudhan (2017); the molecular linelists are primarily from EXOMOL (Tennyson et al., 2016) and HITEMP (Rothman et al., 2010), and the CIA opacities are from Richard et al. (2012). The retrieval explores model parameter space with Bayesian nested sampling using the MultiNest code via the Python wrapper, PyMultiNest (Skilling,

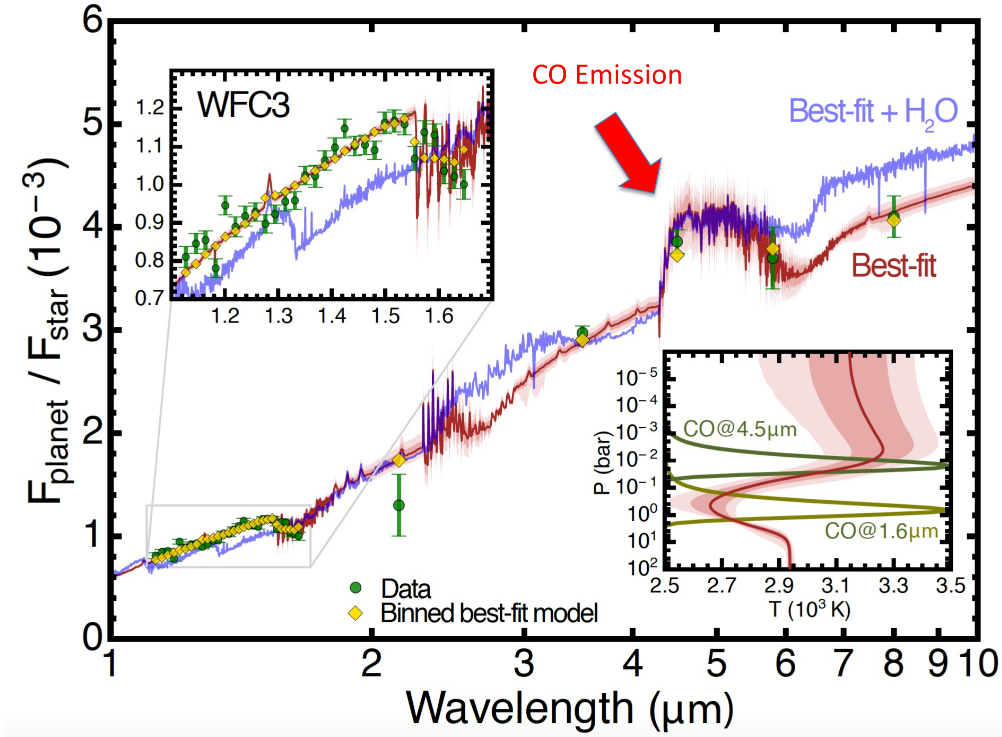


Figure 4.6: Observed spectrum and retrieved solutions. WFC3 and Spitzer data are shown in green. The median retrieved spectrum, with the uncertainty envelopes, is shown in red. The binned median model, in yellow, with $\chi^2_{red} = 3.67$ is an unambiguously better fit than a blackbody ($\chi^2_{red} = 15.2$). A fiducial model with solar-abundance H_2O absorption is shown in blue to demonstrate the lack of an H_2O feature in the data. The results favor a thermal inversion, and the only spectral features detected are those of CO at 1.6 and 4.5 μm. The retrieved P-T profile with error contours is shown in the lower-right inset along with normalized contributions functions at 1.6 and 4.5 μm.

2004; Feroz et al., 2013; Buchner et al., 2014). We sample the multi-dimensional parameter space using 4,000 live points for a total of more than one million model evaluations.

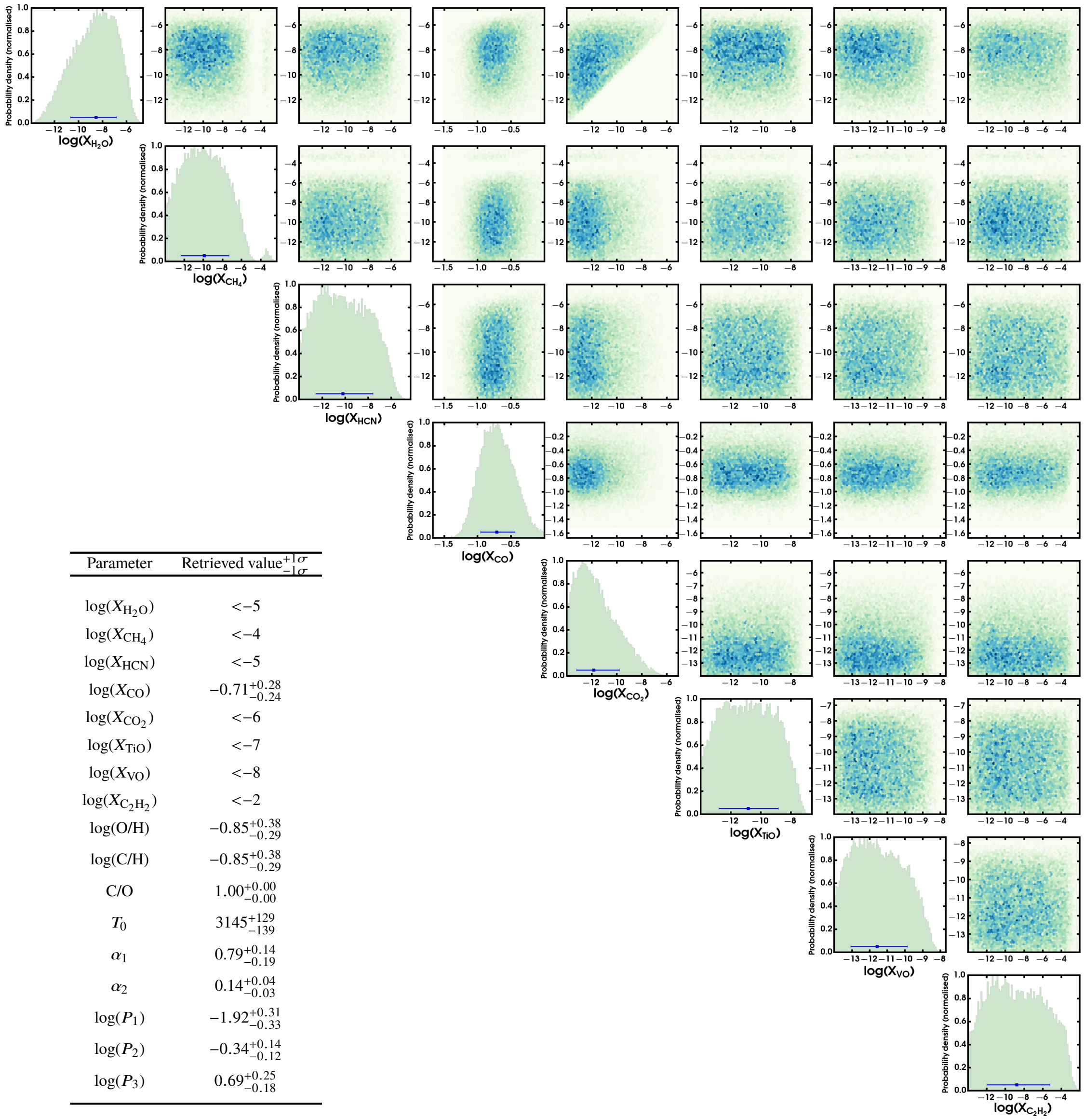
The best-fit retrieval requires a strong thermal inversion in the dayside atmosphere. The bottom inset of Figure 4.6 shows the retrieved P-T profile with confidence contours, indicating an upper atmospheric temperature increase. The requirement of a thermal inversion is guided by the strong emission inferred in the 4.5 μm Spitzer IRAC band, with a brightness temperature of 3100 ± 50 K, which is significantly higher than the rest of the data. This can be explained by the presence of a thermal inversion in the atmosphere along with the presence of either CO or CO_2, which both exhibit pronounced spectral features in the 4.5 μm band (Burrows et al., 2007; Fortney et al., 2008; Madhusudhan & Seager, 2010b). We break this degeneracy by requiring that CO_2 be less than H_2O as expected for hot Jupiter atmospheres (Madhusudhan, 2012; Heng & Lyons, 2016). Another subtlety is the apparent minor trough near $\sim 1.6\,\mu$m, which we attribute to CO absorption below the inversion layer ($\sim$ 1-10 bar), where temperature decreases outward. Emission in the 4.5 μm band is due to CO in the 0.001 - 0.1 bar range which contains the thermal inversion. As part of the nested sampling analysis, we compute the Bayesian evidence value for the retrieved spectrum. By comparing this value with that obtained for a model without a thermal inversion, we conclude that a thermal inversion is favored at the 6.3σ significance level. Similarly, comparison to a model lacking CO implies that the presence of CO is favored at the 6.1σ level. Interestingly, the transition point of the inversion occurs at 0.1 bar which is characteristic of all planets in the Solar System with inversions as well as models of hot Jupiters (Madhusudhan & Seager, 2009; Robinson & Catling, 2014).

Figure 4.7 shows the posterior probability distributions of all the model parameters. The data require a CO volume mixing ratio of $19^{+18}_{-8}\%$ in the atmosphere, which is $380^{+360}_{-160}\times$ the amount expected for a solar abundance atmosphere at this temperature in thermochemical equilibrium. The high CO abundance is primarily constrained by the emission required to explain the 4.5 μm IRAC point as well as the absorption trough in the WFC3 band at 1.6-1.7 μm. We detect no other chemical species (see Figure 4.7). In particular, the non-detection of H_2O at both 1.4 μm and 6 μm provides a robust 3σ upper-limit of 10^{-6} on the volume mixing ratio. The sum-total of constraints on the chemical species lead to a super-solar metallicity in the planet (C/H = O/H = $283^{+395}_{-138}\times$ solar O/H) and a C/O ratio of $\sim$1.

We also conducted free-chemistry retrievals with no priors on the CO_2 abundance and find the same key results. For both cases, the data require a strong thermal inversion, a C/O ratio of $\sim$1, and a super-solar metallicity.

4.6.2 Discussion

The constraints on the chemical abundances are consistent with expectations for a high C/O ratio atmosphere in the high temperature regime of WASP-18b (Madhusudhan, 2012; Moses et al., 2013) where chemical equilibrium are expected to be satisfied. At high temperatures, H_2O is expected to be the most dominant oxygen-bearing molecule for a solar-abundance elemental composition (e.g. with a C/O = 0.5) (Madhusudhan, 2012; Moses et al., 2013). In contrast, the low-abundance of H_2O observed is possible only if the overall metallicity and O abundance were low, or if the C/O ratio were high. Given the high abundance of CO we retrieve, the only plausible solution is both a high oxygen abundance and a high C/O ratio. The constraints on all the other species are also consistent with this scenario. While I cannot rule out a contribution from CO_2 emission in the 4.5 μm *Spitzer*



Parameter	Retrieved value$^{+1\sigma}_{-1\sigma}$
$\log(X_{H_2O})$	<-5
$\log(X_{CH_4})$	<-4
$\log(X_{HCN})$	<-5
$\log(X_{CO})$	$-0.71^{+0.28}_{-0.24}$
$\log(X_{CO_2})$	<-6
$\log(X_{TiO})$	<-7
$\log(X_{VO})$	<-8
$\log(X_{C_2H_2})$	<-2
log(O/H)	$-0.85^{+0.38}_{-0.29}$
log(C/H)	$-0.85^{+0.38}_{-0.29}$
C/O	$1.00^{+0.00}_{-0.00}$
T_0	3145^{+129}_{-139}
α_1	$0.79^{+0.14}_{-0.19}$
α_2	$0.14^{+0.04}_{-0.03}$
$\log(P_1)$	$-1.92^{+0.31}_{-0.33}$
$\log(P_2)$	$-0.34^{+0.14}_{-0.12}$
$\log(P_3)$	$0.69^{+0.25}_{-0.18}$

Figure 4.7: Posterior distributions from our spectral retrieval. The mixing ratios are quoted as common log values. H_2O and CO_2 provide only upper limits, but the high CO abundance implies a high metallicity and high C/O.

band, the high abundance of CO_2 needed would be chemically inconsistent with the non-detection of H_2O, and I therefore believe this scenario to be unlikely.

My inferences for this planet indicate an unusual atmosphere in several respects, calling for comment on the reliability of my conclusions. While the inference of a temperature inversion *per se* is no longer surprising for strongly irradiated planets (Evans et al., 2017; Haynes et al., 2015), both the very high metallicity and C/O ~ 1 have less precedent. Those aspects are forced upon us by the lack of observed water in the WFC3 and *Spitzer* bandpasses, by the slight decrease at the long end of the WFC3 band, and by the *Spitzer* photometry point at 4.5 μm. The non-detection of WFC3 water is certainly robust - several independent eclipses show no sign of the band head that should occur at 1.35 μm (Figure 2). I reiterate that the inference of a thermal inversion hinges critically on the single Spitzer photometric point at 4.5 μm. Previously, Nymeyer et al. (2011) postulated a temperature inversion for exactly that reason. Since our eclipse depth agrees with those from previous analyses (Nymeyer et al., 2011; Maxted et al., 2013), I consider this measurement robust with regard to analysis technique. This is further confirmed by a follow-up analysis by Garhart et al. (2020), which found a similarly high eclipse depth at 4.5μm. Nevertheless, the photometry does not reveal the resolved band structure of the 4.5 μm CO band in emission that would lead to an unequivocal detection of molecular emission. However, given the data and the successful checks on my data analysis procedures, the unusual atmosphere of WASP-18b is a compelling conclusion. Our observations also reveal the first instance where both absorption and emission features are seen in the spectrum of an exoplanet, both due to CO. The absorption at

$\sim$1.6 μm is caused by a weaker CO band compared to the emission in a stronger CO band in the 4.5 μm region. As shown by the contribution functions in Fig. 3, the 1.6 μm region in the spectrum probes the lower atmosphere due to the lower opacity compared to the 4.5 μm band which probes the upper atmosphere due to a higher opacity in that spectral region. Note that simultaneous absorption and emission in the same molecule is observed in the Earth's infrared spectrum, specifically in the 15 μm band of CO_2, due to the temperature structure at the tropopause and stratosphere (Hanel et al., 1972).

If confirmed, the atmospheric properties of WASP-18b open a new regime in the phase space of hot Jupiters. Classically, thermal inversions in hot Jupiters were suggested to be caused by TiO and VO in very high temperature atmospheres (Hubeny et al., 2003; Fortney et al., 2008). All studies so far have focused on the plausibility of TiO/VO as a function of various parameters and processes such as settling and cold traps (Spiegel et al., 2009), stellar chromospheric emission (Knutson et al., 2010), C/O ratio (Madhusudhan et al., 2011), dynamics (Parmentier et al., 2013; Menou, 2012), etc. For TiO/VO to be abundant enough to cause thermal inversions, the C/O balance must be approximately 0.5 or lower (Madhusudhan et al., 2011). Planets with high C/O ratios were not predicted to host thermal inversions since their TiO/VO abundances would be severely depleted (Madhusudhan, 2012); however, recent work suggests other processes, such as inefficient atmospheric cooling, could lead to an inverson (Mollière et al., 2015). Alternatively, oxygen-poor absorbers may play a similar role to TiO and VO (Zahnle et al., 2009b). The two hot Jupiters for which thermal inversions have been detected have both showed signatures of TiO/VO in their atmospheres: WASP-33b (Haynes et al., 2015) and WASP-121 (Evans et al., 2017). WASP-18b is the first system which shows a thermal inversion along with a high C/O ratio of ∼1 with no evidence for TiO/VO, and hence provides a new test case for theories of thermal inversions in hot Jupiters.

WASP-18b's unique atmospheric composition implies an interesting constraint for planetary formation theories. Its metal-enrichment is a factor of 1000 more than that predicted by the inverse mass-metallicity relationship for a $10M_J$ planet (Kreidberg et al., 2014). High metallicity and a C/O ratio of 1 are plausibly explained by formation from extremely CO-rich gas beyond the water condensation line (Madhusudhan et al., 2014a) or upper atmospheric enrichment in carbon and oxygen due to ablation of icy planetesimals during late-stage accretion (Pinhas et al., 2016). Future eclipse observations with the James Webb Space Telescope and improved modeling of giant planet accretion processes will help clarify the details of WASP-18b's formation history.

4.6.3 Addendum

I published the majority of the work on WASP-18b in a letter in 2017. Since then, previously unincorporated physics and opacities have been shown to be important to eclipse modeling of ultra-hot Jupiters like WASP-18b, and other groups have analyzed the planet through that lens. Here, I summarize their findings, what is contested, and what we agree on.

In 2018, Lothringer et al. (2018) and Parmentier et al. (2018) argued that molecular (especially water) dissociation and H- opacity are important physics when analyzing ultra-hot Jupiters. They demonstrate that at very high temperatures (typically above 3000K, which WASP-18b's upper dayside atmosphere is), molecular dissociation become important: not only do normally spectroscopically dominant molecules like water and TiO dissociate, but H_2 also dissociates and K and Na atoms are ionized. This allows, at certain temperatures and pressures, for H-atoms and electrons freed up from metal ionization to combine into H-, which adds continuum opacity (Parmentier et al., 2018). Further, the thermal structure of the dayside atmosphere is closely intertwined with its composition, since opacities impact the rate of heating and cooling at each pressure level. Consequently, the lack of an observed water feature is not necessarily due to a lack of elemental oxygen in the planet's atmosphere (i.e, a high C/O ratio). Instead, a variety of factors could mute the feature: water being dissociated and H- opacity dominating at 1.4μm, or water being dissocated high in the atmosphere and only being opaque in an isothermal section of the T-P profile. If water is present only in isothermal pressure layers, then both in the water band (higher altitudes) and outside the water band (deeper in the atmosphere) sample a Planck function of the same temperature, making the spectrum appear flat instead of as a bump.

Arcangeli et al. (2018) applied this analysis WASP-18b and disagreed with some conclusions from my paper (Sheppard et al., 2017). There was no issue conflict between data analyses — we derived consistent spectra. Arcangeli et al. (2018) incorporated this physics via a cloud-free radiative-convective-thermoequilibrium model called self-consistent CHIMERA (based on CHIMERA; Line et al., 2013).

This model is too computationally expensive to run a full retrieval on, so instead they generated grid points from the forward model with grid points in C/O, metallicity, and heat transport, and interpolated between points for fine enough sampling to perform a grid retrieval. They find that a self-consistent atmosphere can match the observed spectrum decently well if H- opacity and molecular dissociation are accounted for. This allows them to capture the flat WFC3 data and the bump at 4.5μm without necessitating a C/O=1 and a high metallicity. However, they do agree with both our thermal inversion detection, water non-detection in WFC3, and CO detection.

The three papers mentioned in the previous paragraph are excellent, and I agree chemically plausible models should be considered whenever possible. Free-chemistry retrievals, as used in Sheppard et al. (2017), offer flexibility to explore many potential atmospheric compositions and structures with no little prior constraint in a field with constant unexpected results. They have value, but they should be interpreted cautiously in tandem with chemically plausible models. I applied this lesson to the two other chapters in this dissertation, especially Chapter 3. I note that WASP-18b is too hot to retrieve with PLATON (T_{max}=3000 K).

WASP-18b is an important target for JWST, which should be able to easily differentiate between the atmosphere described in Sheppard et al. (2017) and that described in Arcangeli et al. (2018). To contextualize this, I used Pandexo (Batalha et al., 2017) to predict the spectrum assuming our derived properties are correct (e.g, C/O=1, high metallicity). This prediction is shown in Figure 4.8.

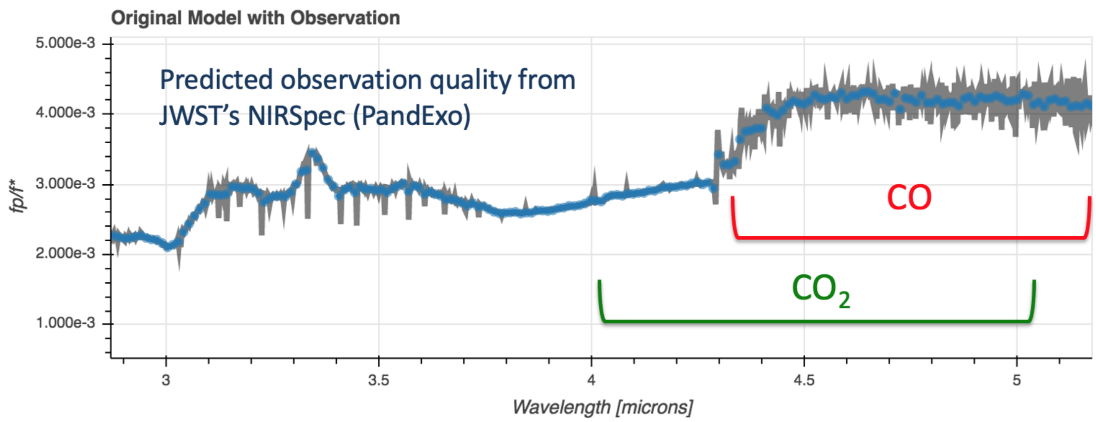


Figure 4.8: Pandexo generated spectrum for a JWST NIRSPEC observation. This emphasizes the utility of JWST, which will very easily be able to distinguish between the CO and CO_2, for example.

4.7 WASP-19b

4.7.1 Atmospheric Retrieval

I use PLATON v5 (Zhang et al., 2020b) to retrieve on the eclipse spectra of WASP-19b. PLATON assumes a chemical equilibrium framework, which ensures chemically plausible atmosphere. This naturally accounts for molecular dissociation and it allows for inclusion of H- opacity. The emission version of PLATON works very similarly to the transit version described in Section 3.6.1.

Due to the computational expense of self-consistent T-P profiles, PLATON only allows parametric T-P profiles. Consistent with WASP-18b, I use the empirical parameterization of Madhusudhan & Seager (2009), which treats the dayside atmosphere as a deep isothermal section, an intermediate section capable of capturing inversions, and a high-altitude, optically thin upper section. PLATON assumes a 1-D plane parallel geometry, with a log pressure grid extending from 10^{-6} to 10^{6} mbar.

The T-P parameters give a temperature at each pressure, which is combined with the elemental abundance in the atmosphere (from C/O and Z) and input into chemical equilibrium calculator ggchem (Woitke et al., 2018) to determine abundances at each pressure level. Temperatures are also used to generate cross-sections of each molecule from line-lists (mostly ExoMol (Tennyson & Yurchenko, 2018)), which are then combined with abundances to get opacities at each wavelength and pressure. Finally, hydrostatic equilibrium is invoked to convert pressure to height, and the emergent flux is calculated from those values. Stellar spectra for a given stellar radius and temperature are interpolated from BT-Settl (Allard et al., 2012). Finally, the eclipse depth at each wavelength is determined by $D_\lambda = (R_{p,\lambda}/R_s)^2 * F_{p,\lambda}/F_{s,\lambda}$.

The Bayesian sampler part of the retrieval is the nested sampling tool Dynesty (Speagle, 2020). For both planets, the 7-parameter prior volume (C/O, metallicity Z, R_p, 4 T-P parameters) is sampled to calculate the Bayesian evidence of the model and determine the posterior distributions of the parameters. I use uniform priors (log uniform when parameter covers many orders of magnitude) with limits set by computational limits of PLATON. Stellar radius, stellar temperature, and planet mass are fixed to TICv8 (Stassun et al., 2019) values (R_s=1.028$R_\odot$, M_p=1.114M_{Jup}, T_s=5503 K)

4.7.2 WASP-19b Results

The results for WASP-19b are shown in Figures 4.9 and 4.10. The model is an excellent fit to the data, with a $\chi^2_{\mathrm{red}} = 1.08$, consistent with expectations for the "correct" model. PLATON interprets the dip in the WFC3 data as water absorption, and predicts a similar dip around 3μm. It determines the fractional log water abundance to be $\log X_{H_2O}$=-3.64$^{+1.44}_{-0.72}$, consistent with both solar and stellar ([O/H]=0.18 (Brewer & Fischer, 2018a)) values.

I determine the detection significance of the water feature by re-running the retrieval with water opacity zeroed and comparing evidences. I find a weak (but existing!) 2.2σ water detection, favored by 4.5$\times$ over the no water model. This relatively low significance is driven by difficulty fitting the *Spitzer* data with water. To avoid possible biases due to offsets between instruments (e.g, the potentially biased absolute depth of WFC3), I re-run the retrievals on just the WFC3 data and find a stronger 2.8σ detection (odds ratio=14), implying water is 93% likely. Metallicity is mostly unconstrained, though C/O is forced to be less than 0.63, consistent with solar (0.53) and stellar ([C/H]=0.13, [O/H]=0.18, C/O=0.5 (Brewer & Fischer, 2018a)). This is likely driven by the presence of water and the lack of obvious CO or CO_2 features.

The presence of water and low C/O ratio agree well with transit observations, which consistently find strong water features and no CO/CO_2 absorption in the limb (e.g, Sing et al., 2016). The water abundance is consistent with the 0.2–30$\times$ solar value determined in transit analyses (Benneke, 2015). This weak-to-moderate detection places WASP-19b as the second hottest exoplanet with water absorption seen in eclipse spectrum, after Kepler-13Ab (Beatty et al., 2017). The spectroscopic WFC3 data significantly improve atmospheric constraints of WASP-19b: the water detection, water abundance, and low C/O ratio were not found by previous analyses

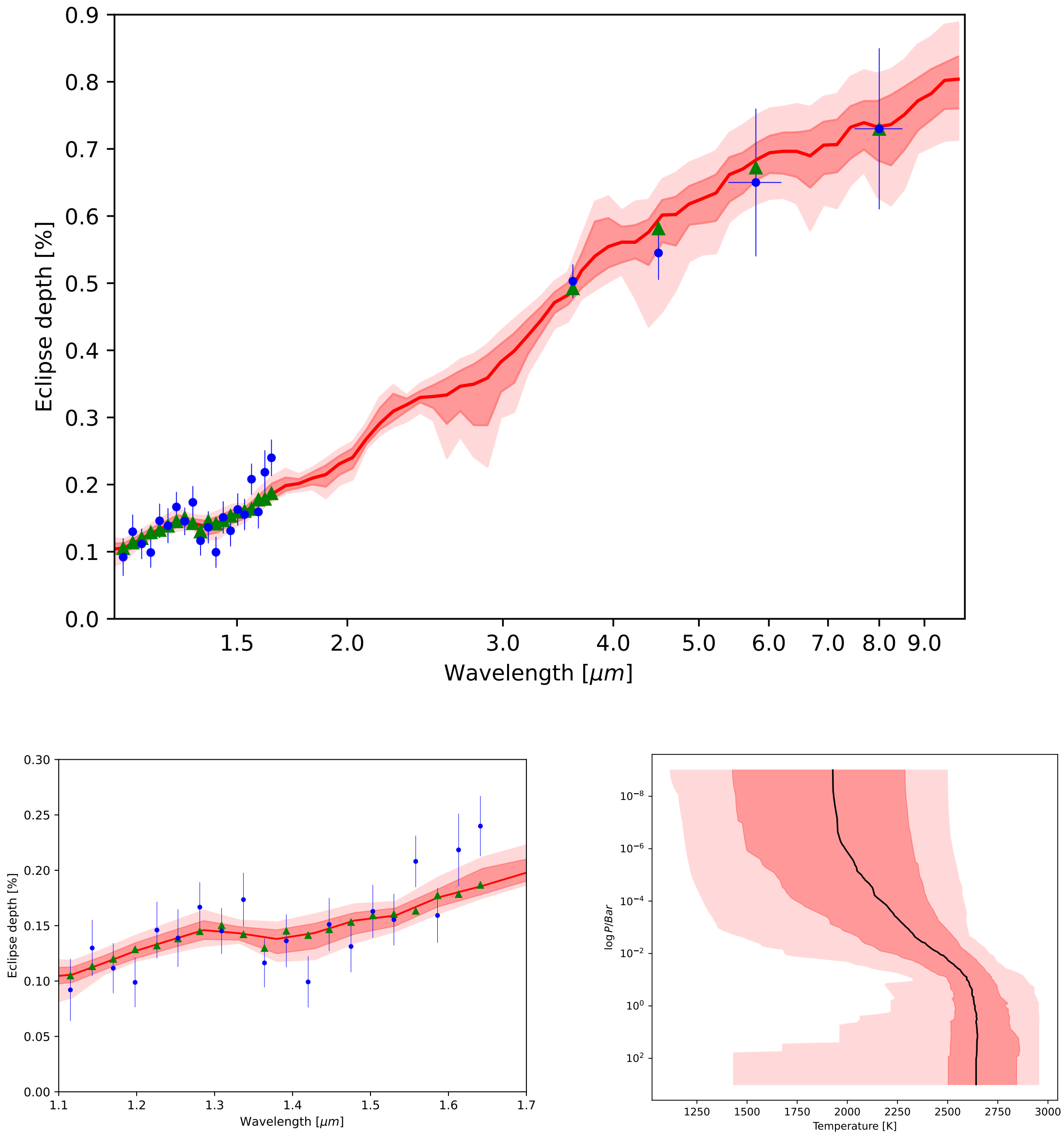


Figure 4.9: Median retrieved model (solid line) with 1 and 2-sigma contours (red; light red). The data is shown in blue, and the median-binned model is shown in green. **Top:** Full spectrum, the model is an excellent fit to the data. **Left:** Zoomed in WFC3 spectrum. The dip at 1.4μm is a water feature. **Right:** T-P profile, with lower pressures (higher altitudes) at the top. For WASP-19b, temperature decreases with height.

which depended only on *Spitzer* and two ground-based photometric points (Line et al., 2014; Madhusudhan, 2012). Those analyses found no evidence of water and determined that a high C/O ratio (≥ 0.99) best explained the data.

The thermal profile is well-constrained to be decreasing with height, with the deep-atmosphere temperature T_3 well constrained to be around 2600 K. The optically thin region temperature T_0 is less constrained, but definitively less than T_3. This decreasing profile is driven by water absorption, since absorption can only occur in atmospheres which decreasing T-P profiles (similar to cool gas in front of a hot star). This decreasing T-P profile is interesting in light of the debate on the presence of TiO in the limb of WASP-19b from transit observations (Sing et al., 2016; Sedaghati et al., 2017; Espinoza et al., 2019; Sedaghati et al., 2021). In the gaseous phase, TiO is an excellent high altitude optical absorber. Since stellar flux is greater in the optical than the IR, a surplus of flux is absorbed high in the atmosphere. In order to maintain thermal equilibrium, the temperature much increase such that flux can be re-radiated efficiently enough to match absorption. The lack of a high-altitude thermal inversion acts as evidence in favor of the TiO non-detections.

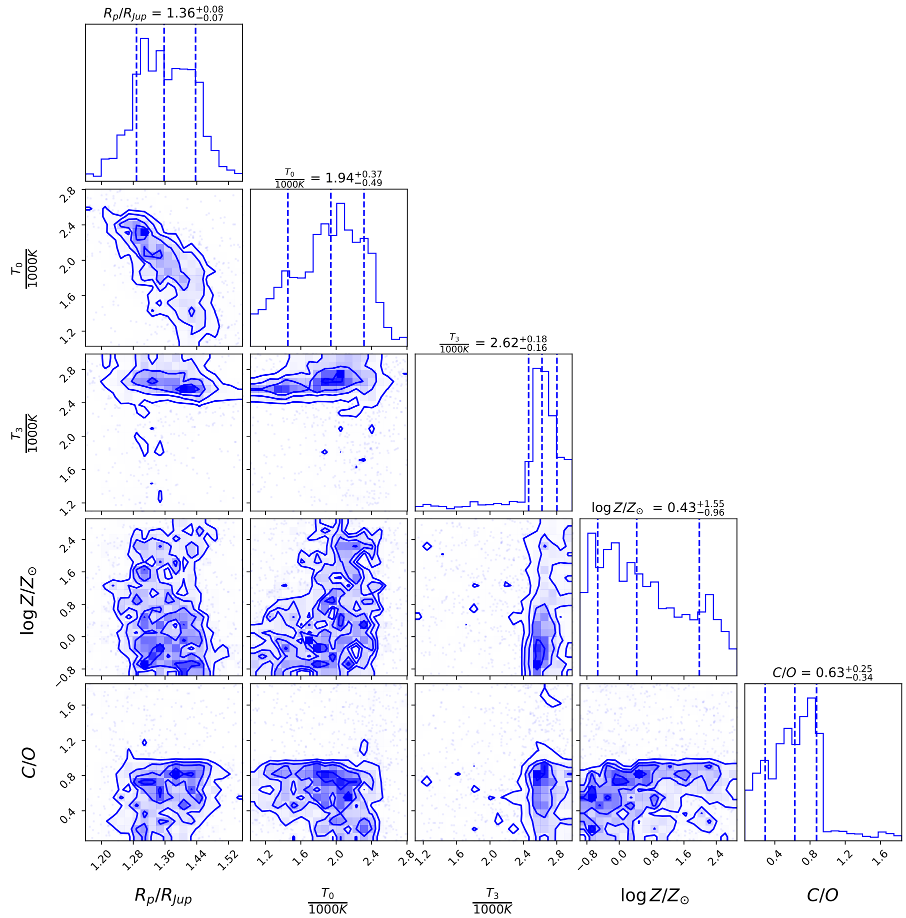


Figure 4.10: Corner Plot from WASP-19b Eclipse Retrieval

The equilibrium temperature and gravity of WASP-19b puts it in an interesting intersection of parameter space: it is on the borderline of where TiO is expected to be gaseous, and the borderline of significant water dissociation. This makes the water absorption detection especially interesting, since it is inconsistent with thermal inversions (and by extension gaseous TiO) and water dissociation. Parmentier et al. (2018) modeled expected water dissociation based on temperature and gravity, and determined that dissociation should be considered for planets on or above the 20% water dissociation contour. WASP-19b is almost directly along that contour (their Figure 13). It is possible that WASP-19b has efficient enough circulation where the dayside atmosphere is relatively cooler than the dayside-only distribution case, effectively pushing WASP-19b down to lower dayside temperatures and preventing significant water dissociation.

There are several reasons why TiO might not be present in a hot Jupiter's atmosphere. Spiegel et al. (2009) and Parmentier et al. (2016) suggest a "vertical" cold trap as a possible mechanism, where TiO randomly crosses a condensation boundary and becomes trapped in the interior, but WASP-19b does not have strong enough gravity to cause this (and is above the 1900 K threshold which is predicted to be necessary (Parmentier et al., 2016)). Another possibility is that the dayside atmosphere is so hot that TiO is thermally dissociated (Lothringer et al., 2018). WASP-19b's T-P profile does not reach temperatures hot enough to dissociate TiO. Further, if the temperature were high enough, then metal atoms and ion opacity in the optical would be sufficient for driving an inversion (Lothringer et al., 2018).

The more likely explanations are either a nightisde cold trap or photodis-association. Parmentier et al. (2013) hypothesized that as TiO circulates around the atmosphere it condenses on the cooler nightside, sinks into the interior, and is depleted from the spectroscopically active regions of the atmosphere. Beatty et al. (2017) applied this explanation to explain the water absorption in the ultra hot Jupiter Kepler-13Ab, and predicted that condensate particles with radii greater than 3μm are necessary to deplete WASP-19b of TiO an inhibit a thermal inversion. For comparison, Kepler-13Ab only needed particles of size 1μm, and they assumed planets which needed large particle sizes would likely exhibit thermal inversions. Still, this is a plausible solution. Alternatively, Knutson et al. (2010) presented a correlation between stellar activity and lack of inversions, arguing that the additional UV flux is photodissociating TiO. Given that WASP-19b is a very active star ($\log R_{HK}$=-4.5), this is a plausible explanation. WASP-19b is an interesting test case which demonstrates how secondary eclipse spectra and thermal structure can inform atmospheric physics.

Chapter 5: Summary and Future Work

This dissertation investigates the physics and composition of exoplanet atmospheres. Specifically, in it I characterize the atmosphere of five exoplanets, from Earth-sized planets in a multiplanet system to massive, ultrahot Jupiters. This analysis constrains the characteristics of those planets' atmospheres, an innately interesting problem which informs the diversity of exoplanet atmospheres and indicates which physical and chemical processes are important. Further, it contributes to the goals of the field at large by adding to the only ~ 30 planets with well-characterized (multi-space instrument) atmospheres (Figure 5.1). While we are a ways from statistically significant population studies (e.g, recent simulations indicate 500 planets may be necessary to confirm certain trends (Bean & FINESSE Science Team, 2017; Kempton et al., 2018)), adding to this sparsely-filled population with thoroughly investigated results is a major goal of the field. In enables better-refined models over a wider-range of parameter space. This applies to atmospheric physics models (temperature redistribution, T-P profiles), condensation models (rainout, cloud height), chemistry (chemical equilibrium, abundance profiles), and planet formation, among others. Further, current spectra derivations and retrieval analyses demonstrate the sensitivity of conclusions to model assumptions, which informs analyses of data from next-generation telescopes where systematic effects are likely to be more important. Finally, this analysis identifies interesting targets for JWST follow-up, helping to maximize observation efficiency.

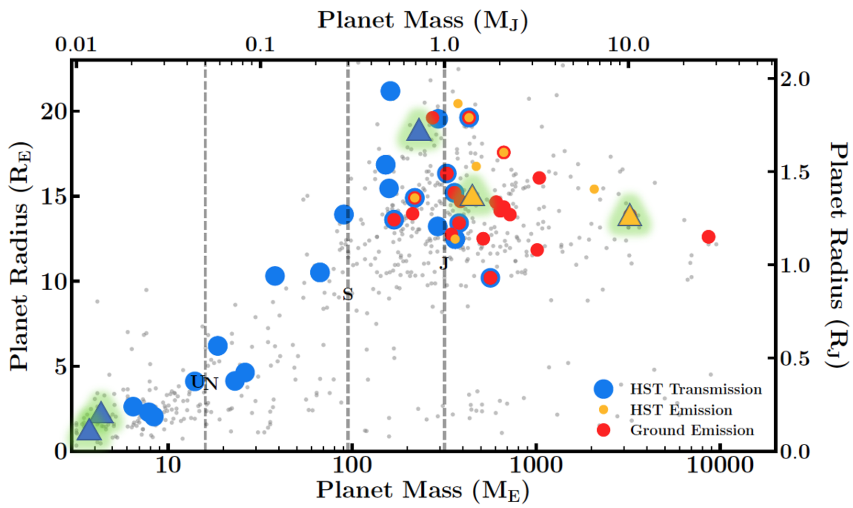


Figure 5.1: The sample of planets analyzed in this dissertation (glowing green triangles) compared to the population of planets with transit and eclipse spectroscopic observations, circa 2019. Original figure from Madhusudhan (2019b).

Here, I briefly summarize the process of characterizing an exoplanet atmosphere via transit or eclipse spectroscopy. Characterization utilized space-based telescope observations of planetary transits and eclipses. Leveraging the geometry of primary (or secondary) planetary transits to indirectly measure a planet's spectrum is known as transmission (emission) spectroscopy. The spaced-based instruments on HST and *Spitzer* are the golden standard for exoplanet light curves, as they are not subject to complications from Earth's atmosphere and can observe in the infrared with minimal thermal contamination. Images from the telescopes are processed and converted them into flux time series (light curves), which are split into wavelength bands and fit with a transit (eclipse) model to derive a transit (eclipse) depth at each wavelength. These flux time series (which are well-calibrated and frequently validated by instrument science reports[1]) thus provide an indirect planetary spectrum. This spectrum is a measure of stellar light blocked by the planet's atmosphere (or, in the case of eclipses, the planet flux emitted relative to stellar flux) as a function of wavelength.

This spectrum is compared to predicted spectra from forward models, which assume a 1D, plane-parallel atmosphere. These atmospheric models include opacities from roughly 30 abundant and spectroscopically active atomic and molecular species, including water, carbon dioxide, TiO, and CO. They also account for collision-induced absorption, Rayleigh scattering, cloud (condensate) opacity, and haze (photochemical byproduct) opacity. The atmospheric physics used for exoplanet atmosphere models are derived from stellar atmospheres, and have been validated on stars, brown dwarfs and Solar System planets (Hubbard et al., 2002; Tsuji et al., 1996; Allard et al., 1996; Burrows & Sharp, 1999). The 1D models have been cross-checked with full 3D models (GCMs) and exhibited agreement within observation uncertainties (Burrows et al., 2010; Fortney et al., 2010; Blecic et al., 2017). Finally, the temperature-pressure profile parameterizations have been validated against both Solar System planets and self-consistent radiative-convective models (Madhusudhan & Seager, 2009; Kempton et al., 2017).

By fitting bench-marked exoplanet atmosphere models to calibrated, high quality space-telescope data using a Bayesian sampler (e.g, `emcee` (Foreman-Mackey

[1]https://www.stsci.edu/hst/instrumentation/wfc3/documentation/instrument-science-reports-isrs

et al., 2013) or `dynesty` (Speagle, 2020)), I am able to constrain physical and chemical properties of several exoplanet atmospheres. This is done both via parameter estimation (e.g, temperature, metallicity, and C/O posterior distributions) and model comparison (e.g, determining detection significance by comparing the Bayesian evidence of a model with water opacity to the same model without water opacity), and simultaneously accounts for potential instrumental biases by marginalizing over those biases as nuisance parameters. The results for each of the five planets analyzed in this dissertation are summarized below and in Figure 5.2.

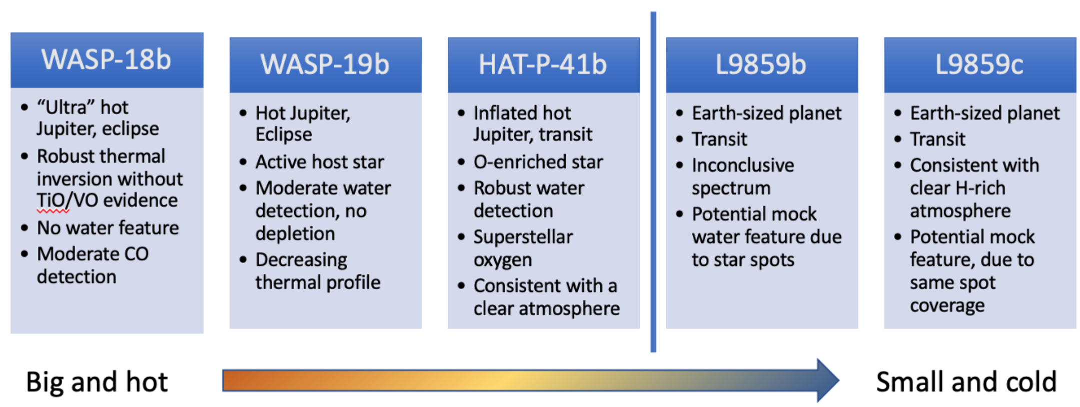


Figure 5.2: Summary of Results and characteristics of interest for each planet.

L9859 Planets: I derived transit spectra of two planets in the second-closest multi-planet system to Earth, using four HST WFC3 visits for L9859b and one of L9859c (Sec 2.4). I found evidence of structure in both spectra, which is unique for Earth-sized planets (ESPs) and potentially indicative of an atmosphere. Such an atmosphere would be the first observed around an ESP.

I used the atmospheric retrieval tool PLATON to perform Bayesian model comparison and determine the evidence in favor of an atmosphere on L9859 (L9859b is unable to retain the H_2-dominated atmosphere that PLATON assumes; Sec 2.5.1). I find an H_2-dominated atmosphere is weakly-preferred over a flat line (no atmosphere transiting a quiet star), and that L9859c is the first ESP with a water feature size that is consistent with a clear, H_2-dominated atmosphere (Sec 2.5).

However, dropping the assumption of a quiet host star impacted the results (Sec 2.5.2). I performed another L9859c retrieval assuming an atmosphere-free planet transiting an active host M-dwarf with variable star spot coverage fraction, finding this scenario to be weakly preferred over the H_2-dominated atmosphere. Further, I performed the same retrieval on L9859b and found that its structure is captured by an identical spot coverage fraction ($\sim 35\%$). This is evidence the structure in both spectra are "mock" features, meaning both planets are atmosphere-free bodies orbiting an active star. Though not definitive, this is the first observational evidence of "mock" spectral features on multiple planets in the same system.

Future Work: It is possible that the observed spectra are a combination of lower-magnitude "mock" features and planetary features in higher mean-molecular weight atmospheres. For example, instead of 35% spot coverage and no atmosphere, its possible that the coverage is 15% and L9859c hosts a steam atmosphere. Better understanding the host star spot coverage provides the most obvious motivation for follow-up work. There are currently two more HST WFC3 observations of the system scheduled to complete soon: another of L9859c, and one of the third planet L9859d, an expected mini-Neptune. The L9859c observation will clarify if the spectral feature is real. If real, the SNR will increase when combining observations, if not, then it will decrease. The L9859d observation will provide useful constraints on host star activity, since its bulk density indicates it hosts a significant H2-dominated atmosphere. An example of such a constraint is if a water feature is observed which is larger than possible due to the atmosphere alone — this would indicate a mock stellar feature is adding onto an existing planetary feature. Additionally, a more detailed modeling of the spot coverage on L9859 is warranted, similar to what Wakeford et al. (2019) did for the TRAPPIST system.

HAT-P-41b: I derived a wide-coverage (0.3–5μm) transit spectrum for the inflated hot Jupiter HAT-P-41b using data from HST STIS G430L (two visits), G750L (one), HST WFC3 G141 (one), and *Spitzer* IRAC1 and IRAC2 (one each) transit observations (Sec 3.5). I (and collaborators) took a multi-pronged approach and performed a suite of retrievals within different model paradigms (Sections 3.7 and 3.9) to allow for both inter- and intra-model comparisons (Sec 3.10).

While the underlying forward models (i.e, the physics of stellar light interacting with gaseous species in a exoplanet atmosphere) are well-validated, there are specific assumptions (e.g, chemical equilibrium, inclusion of certain opacity sources, or condensation scheme) which are debated. It is important to explore the sensitivity of conclusions to those assumptions, and so I argue the comprehensive approach we take is necessary to appropriately contextualize results. Of note, I consider the possibility of a uniform bias between different instruments and explore how such a bias could impact interpretation (Sections 3.6.1.2, 3.9.2, and 3.7.2.3). I argued that a bias is possible both from assuming point value estimates for orbital parameters and as potential relics from spectral derivation (since absolute depth is typically less well constrained than spectral shape). Accordingly, I recommend accounting for potential biases via physically-motivated Gaussian-prior offsets in spectral instruments (STIS/WFC3) in at least one retrieval in order to investigate the impact. This is especially important when there is only a single observation for a given instrument.

Both retrieval methods find a metal-rich atmosphere for almost all model assumptions (most likely O/H ratio of $\log_{10} Z/Z_{\odot} = 1.46^{+0.53}_{-0.68}$ and $\log_{10} Z/Z_{\odot} = 2.33^{+0.23}_{-0.25}$, for AURA and PLATON, respectively). This corresponds to a significantly super-stellar oxygen-enrichment, making it the hot Jupiter with the highest atmospheric metallicity to date (Secs 3.8 and 3.9). The metal-enrichment is driven by a 5σ detection of water as well as evidence of gas absorption in the optical (>2.7-σ detection) due to Na, AlO and/or VO/TiO, though no individual species is strongly detected and the two methods disagree on which species drives the feature. Both retrievals determine the transit spectrum to be consistent with a clear atmosphere, with no evidence of haze or high-altitude clouds. Interior modeling constraints on the maximum atmospheric metallicity ($\log_{10} Z/Z_{\odot} < 1.7$) favor the AURA results (Sec 3.10.2).

Future Work: A parallel study examined the HST UVIS spectrum of HAT-P-41b (Lewis et al., 2020), and a natural extension would be to combine the entire transit data set (UVIS, STIS, WFC3, and *Spitzer*) into a single analysis. Similarly, there are archival HST WFC3 and *Spitzer* eclipse data on HAT-P-41b, which can be used to further discern between the environments inferred by the different retrieval methods. A joint retrieval — where the transit and eclipse spectra are constrained simultaneously — would be especially informative. More generally, it would be interesting to re-analyze literature multi-instrument spectrum to determine if including an instrumental bias parameter significantly impacts conclusions.

WASP-18b and WASP-19b: I derived the HST WFC3 emission spectra for these two highly irradiated hot Jupiters and combined it with *Spitzer* eclipse data to retrieve their atmospheric properties (Sec 4.4). Most notably, I found robust evidence for a strong thermal inversion on the dayside atmosphere of the massive, ultra-hot WASP-18b (Sec 4.6). I found a non-detection of water, TiO, and VO, but a moderate detection of a CO via both an emission feature at 4.5 μm and a less convincing absorption feature at 1.6 μm. The derived composition and T/P profile suggest that WASP-18b is the first example of a planet with a non-oxide driven thermal inversion.

In WASP-19b, I found moderate evidence (2.8σ) of water absorption and a decreasing T-P profile (Sec 4.7). Additionally, I derived a water abundance ($\log X_{H_2O}$=-3.64$^{+1.44}_{-0.72}$) and C/O ratio (C/O<0.63) which are consistent with both host star abundances and those found in the transit analysis (Benneke, 2015). My constraints based on the spectroscopic HST eclipse data overturn the results of previous eclipse analyses, which relied only on four *Spitzer* and two ground-based photometric points (Madhusudhan, 2012; Line et al., 2014). The presence of water and decreasing thermal profile have interesting implications on TiO-driven inversions (Fortney et al., 2008) and water dissociation (Parmentier et al., 2018), since WASP-19b is hot enough (2100 K) that both are expected to play a role.

Future work: Like the other planets in this dissertation, analysis of additional data is a natural extension. For WASP-19b, that involves a joint transit/eclipse spectra retrieval, which can leverage the extra information to achieve additional constraints (e.g, Kreidberg et al., 2014b). For both, JWST observations will provide vital insight into their atmosphere's natures, and Figure 4.8 gives an example of how JWST's high resolution will easily break current degeneracies. In the big picture, a uniform analysis of the roughly 30 eclipse observations on the MAST archive can add to these case studies and best constrain the atmospheric questions most relevant to eclipses. An example would be to update the Knutson et al. (2010) study, which correlated stellar activity to thermal inversions based on *Spitzer* photometric data alone. The additional resolving power of spectral HST data allows for more strict constraints on population-wide trends.

Analysis Pipeline: Finally, I developed the codes I used to analyze data in this dissertation into a Python 3 pipeline, nicknamed `DEFLATE`, which is downloadable on Github[2]. I described the pipeline in detail in Sections 2.3 and 2.4. Within the dissertation, I validated my pipeline against literature spectra and verified my derived depths with a suite of quality-of-it diagnostics.

The pipeline is highly customizable, allowing for exploration of the impact of both data processing and light curve fitting assumptions on the derived transit spectrum. It converts telescope fits image files to light curves (non-trivial for spatial scan mode), and uses marginalization to fit those light curves to both determine orbital properties (namely radius, transit/eclipse time, and optionally linear limb-darkening coefficient, inclination, or a/R_s) and to derive a transit spectrum. Is also provides a suite of diagnostics to verify light curve fits.

Future work: Given the flexibility of the systematic model grid, this pipeline

[2]https://github.com/AstroSheppard/WFC3-analysis

can be readily expanded to work with JWST data once available, and will be especially useful in determining which auxiliary parameters (such as orbital phase or wavelength shift) are relevant to JWST light curve analyses. More immediately, the flexible and customizable nature of the code make it ideal for creating a WFC3 exoplanet spectral library. While spectra are published in the literature, a uniform analysis approach is necessary to ensure minimal bias due to analysis method in population studies. A uniform spectral library will give modelers access to an unprecedented amount of spectroscopic exoplanet data, enabling more frequent comparative exoplanetology.

Appendix A: Chapter 3 Supplementary Material

A.1 L9859b White Light Curves

These figures show the band-integrated data for each of the three L9859b observations not shown in the chapter (Section 2.4). The band-integrated light curve is calculated by summing the integrated flux of every pixel in the *final.fits* exposures, which gives the total photons observed over the course of the exposure. The de-trended light curve is the data divided by the highest-weight systematic model. The bottom panel shows the residuals between the data and the highest-weight light curve model. For each L9859b observation, the first orbit and first few exposures of each orbit are removed.

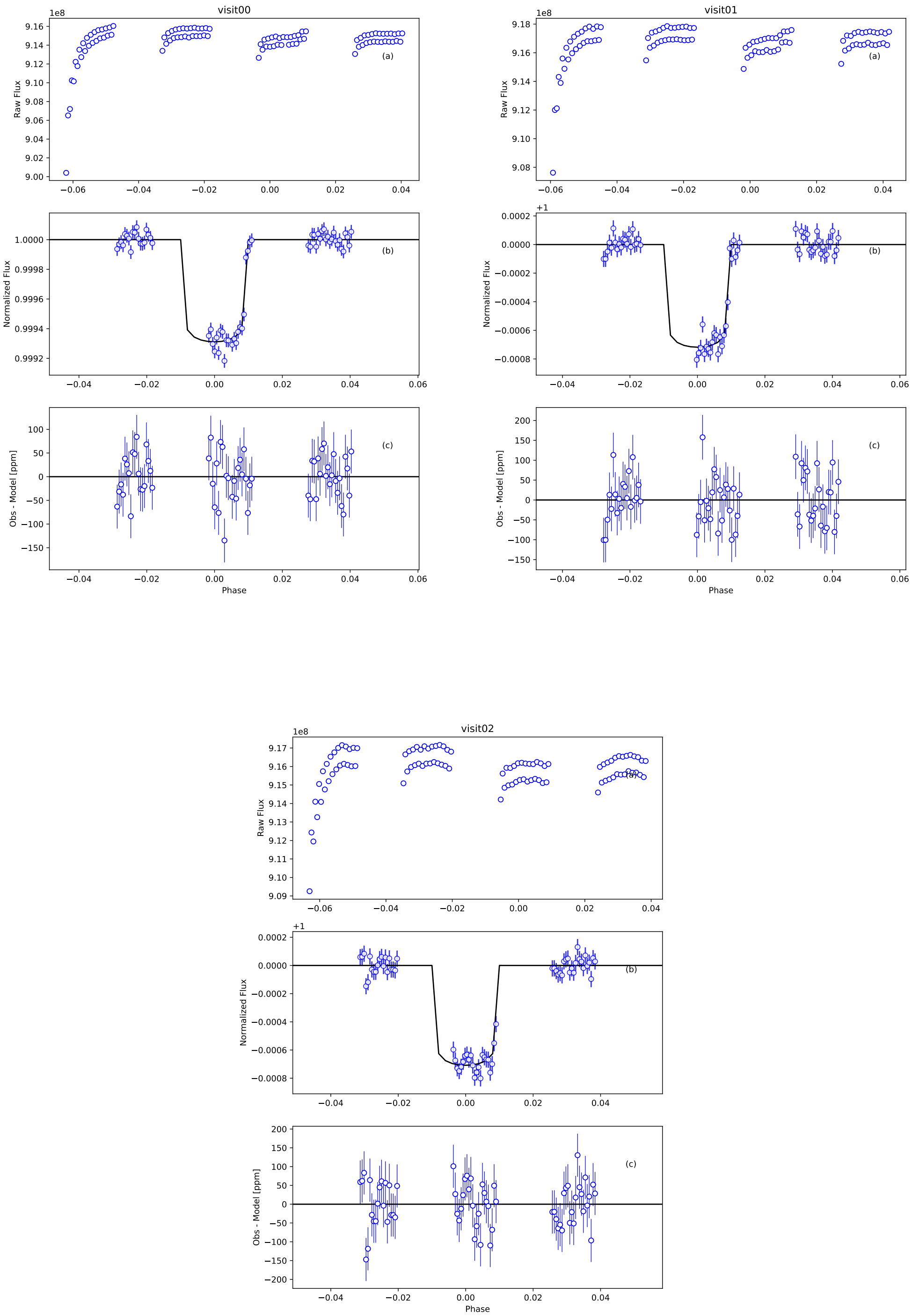


Figure A.1: Visualization of white light curve fit for the highest weighted systematic model for L9859b visits 00, 01, and 02. Panel (a) shows the band-integrated light curve. Panel (b) shows the de-trended light curve and the best fitting transit model. Note that this is illustrative — the instrumental effect and transit model parameters are fit for simultaneously. Panel (c) shows the residuals between the data and the best-fitting model.

A.1.1 MCMC Validation Figures

These are the additional figures from the L9859c whitelight fit. First, I visualize the result in Figure A.2 by projecting random samples from the posterior onto the space of the data.

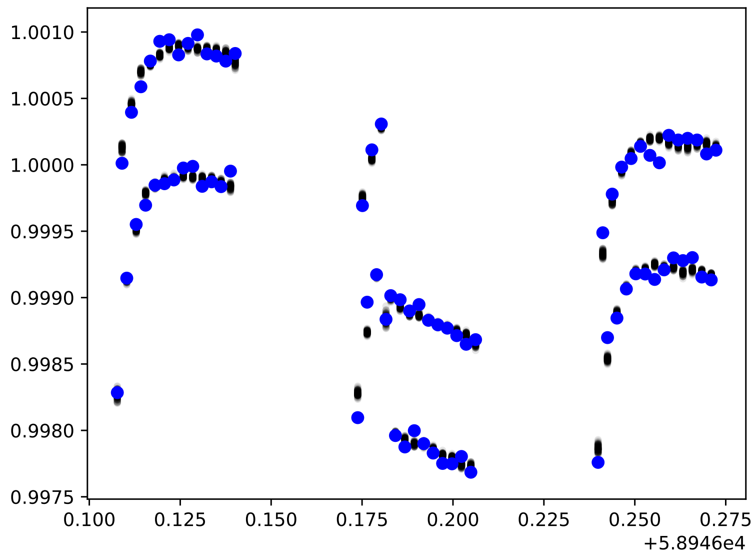


Figure A.2: Distribution of model (black points) based on parameters derived in MCMC fit compared to data (blue). Y-axis is normalized flux and x-axis is MJD time.

Next, I provide the full corner plot. Transit depth has no strong dependence or correlation with any systematic parameter.

Finally, I prove covergence via autocorrelation. In ensemble MCMC samplers like `emcee`, the Gelman-Rubin statistic is not valid since the chains are not independent. Instead, for adequate sampling to achieve a small enough computational error, it is a good rule of thumb to run chains for at least $50\times\tau$, the autocorrelation time. This is not known *a priori* and must be estimated, as in Figure A.4. This process is

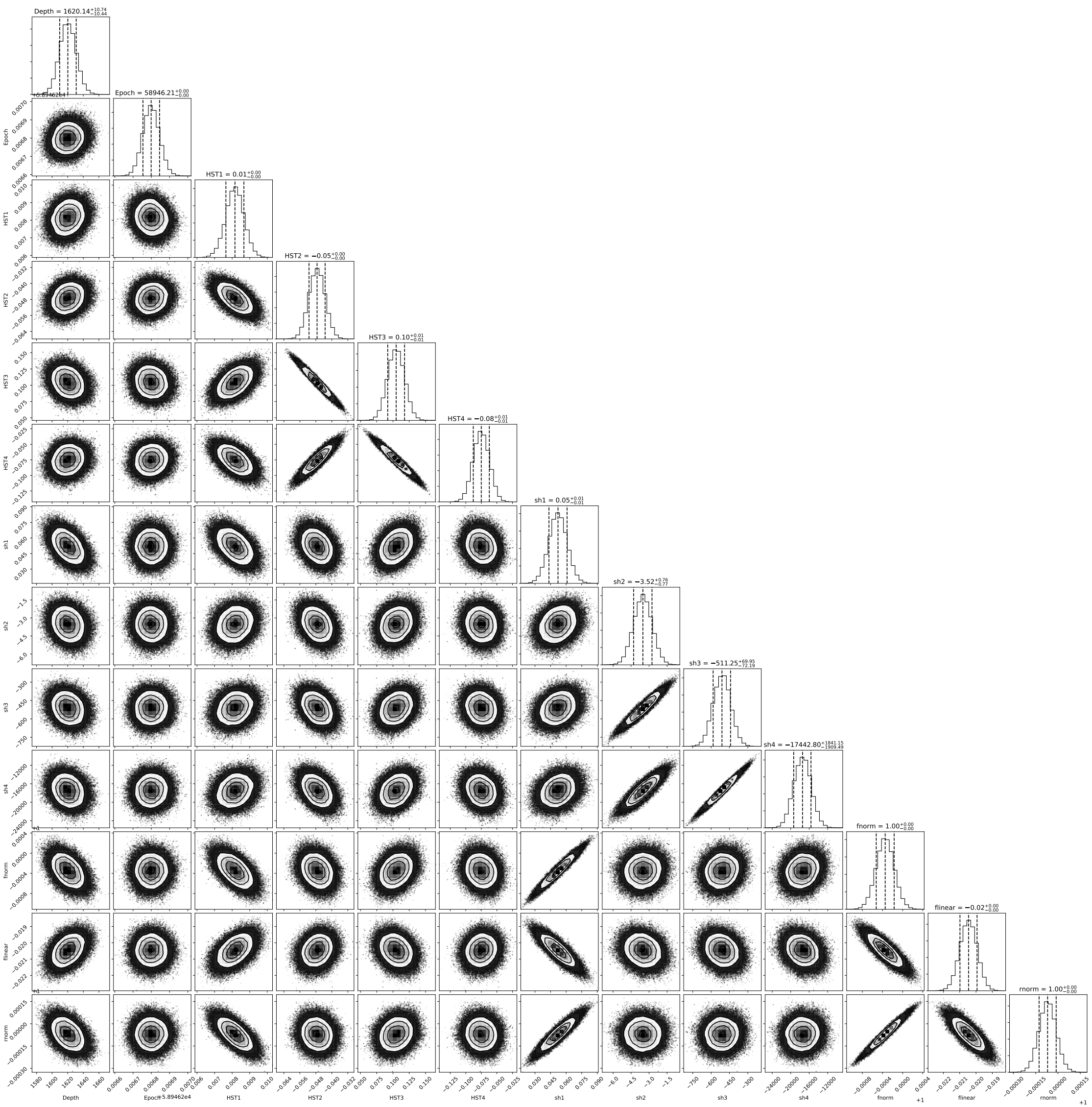


Figure A.3: L9859c Transit MCMC Full Corner Plot explained in this tutorial https://emcee.readthedocs.io/en/stable/tutorials/autocorr/. At each chain length, the autocorrelation time is estimated for each parameter. When these estimates flatten out and cross the N/50 line, then the autocorrelation estimate becomes reliable. It is most important that depth converges, but ideally all parameters will. In this case, the autocorrelation time is about 100, meaning 5000 samples is the minimum for convergence. I run the chain for 20000 steps for 2.5×nDimensions walkers, and conservatively remove the first 2000 as burn-in.

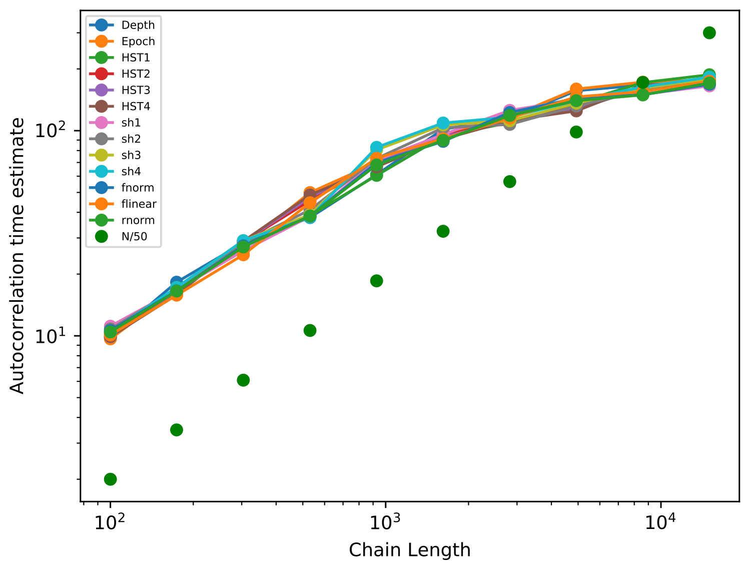


Figure A.4: L9859c Transit MCMC Proof of Convergence

A.2 L9859 Transit HST Spectrophotometric Light Curve Fits

The figures in this section show the de-trended light curve data and best-fit transit model for every spectral bin for each observation. These are analogous to panel (b) of the figures in Section A.1. They illustrate the fit used to derive the transit spectrum (Section 2.4).

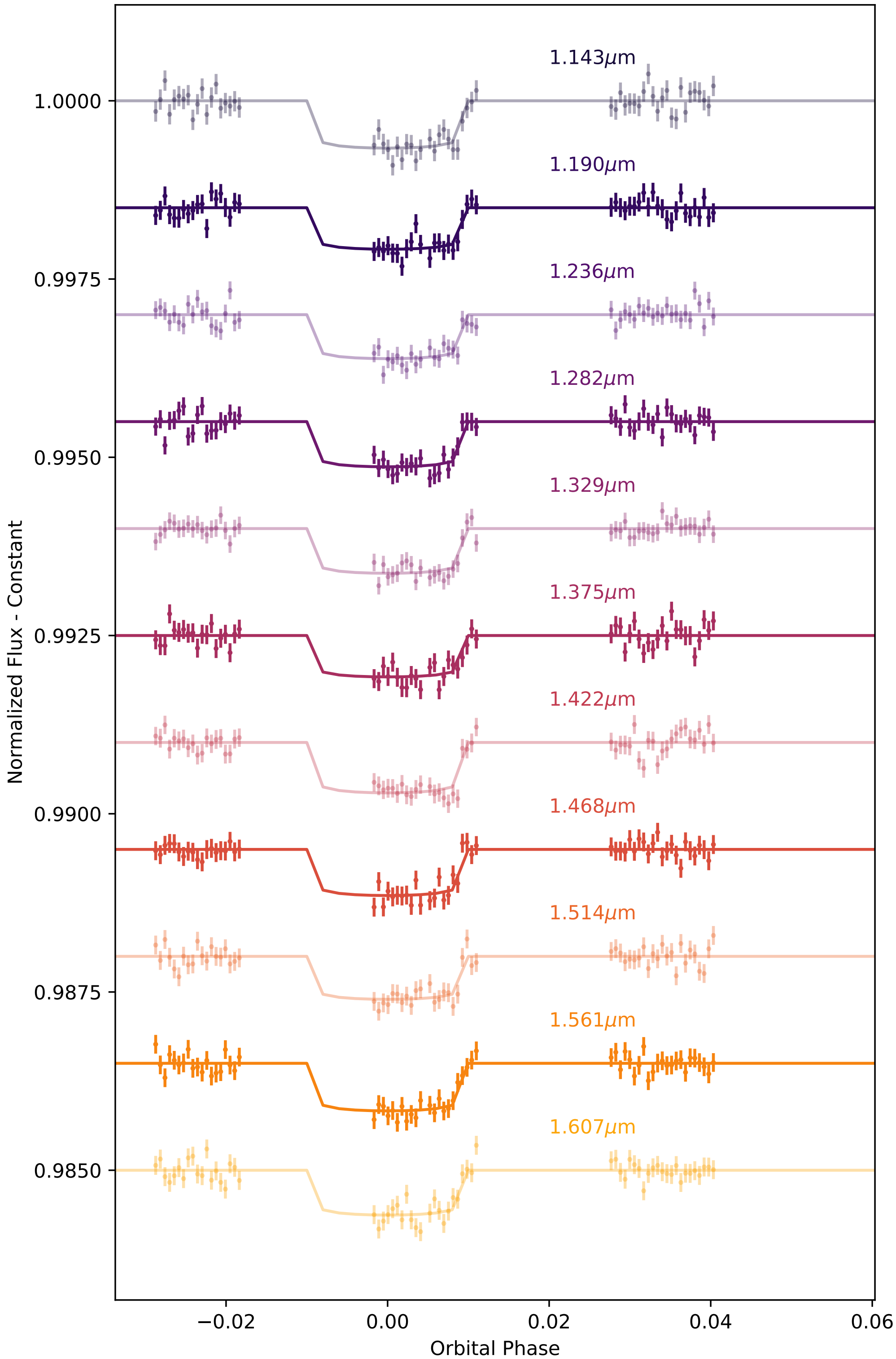


Figure A.5: Spectral light curves for L9859b, visit 00.

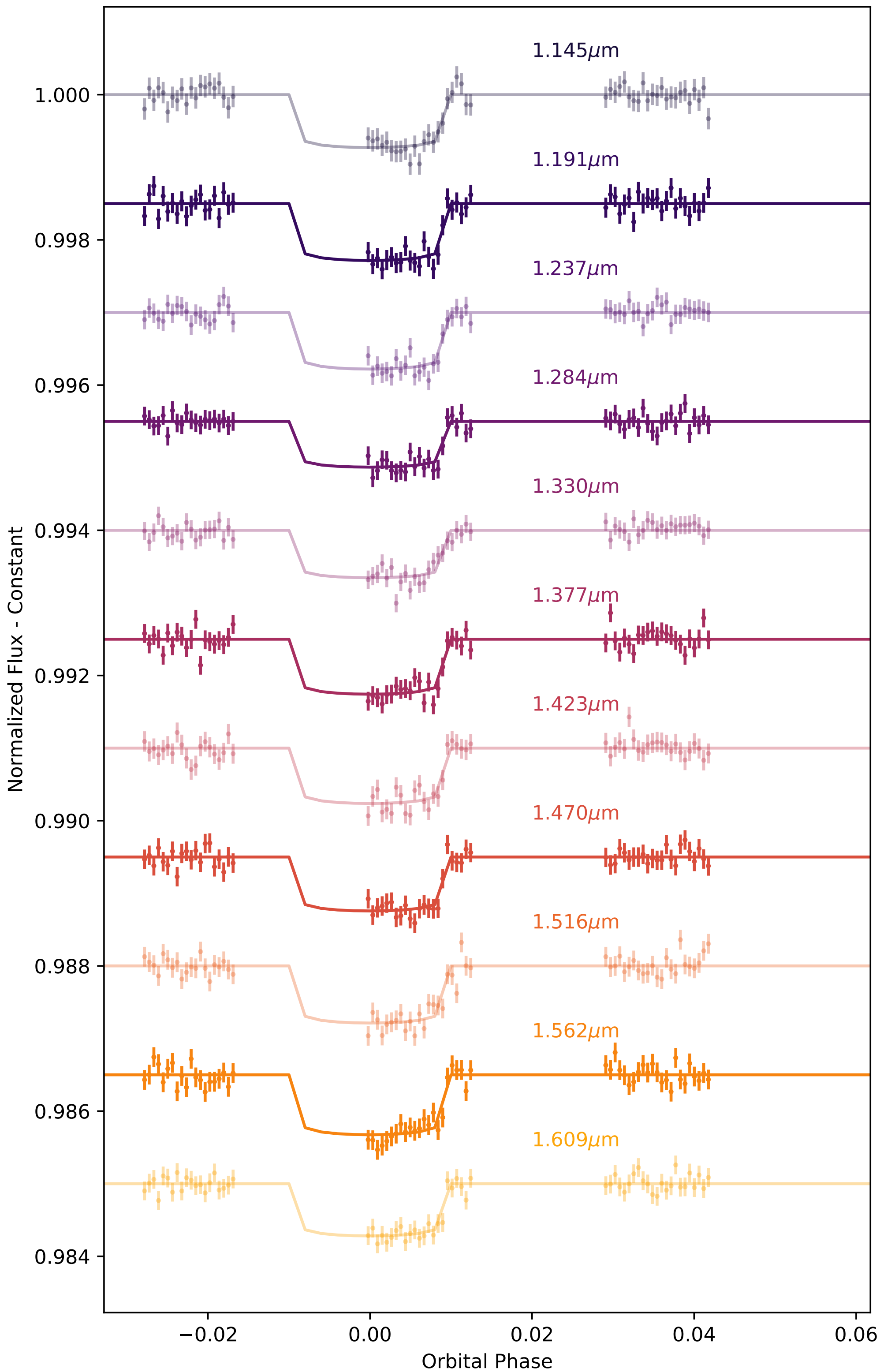


Figure A.6: Spectral light curves for L9859b, visit 01.

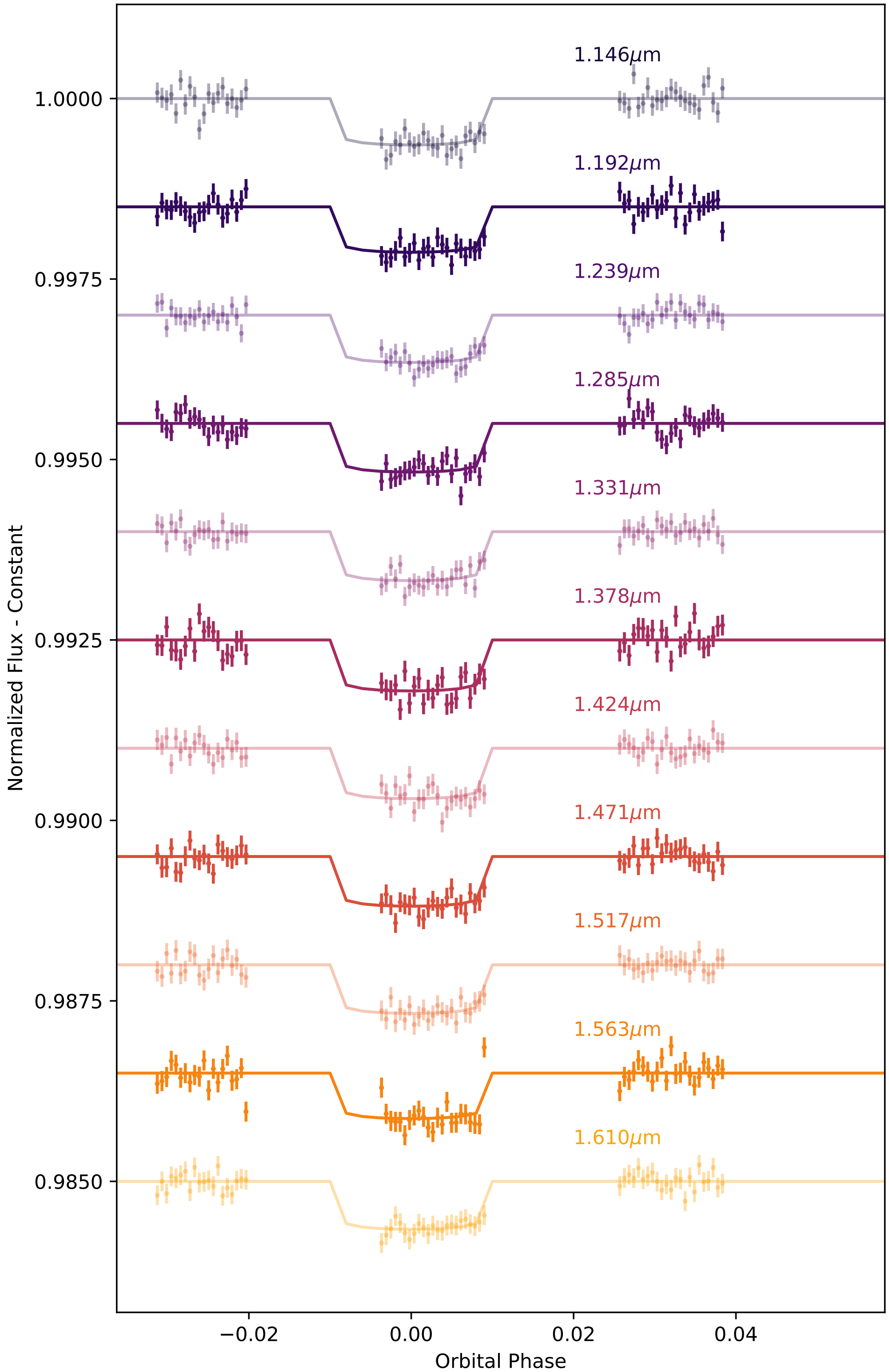


Figure A.7: Spectral light curves for L9859b, visit 02.

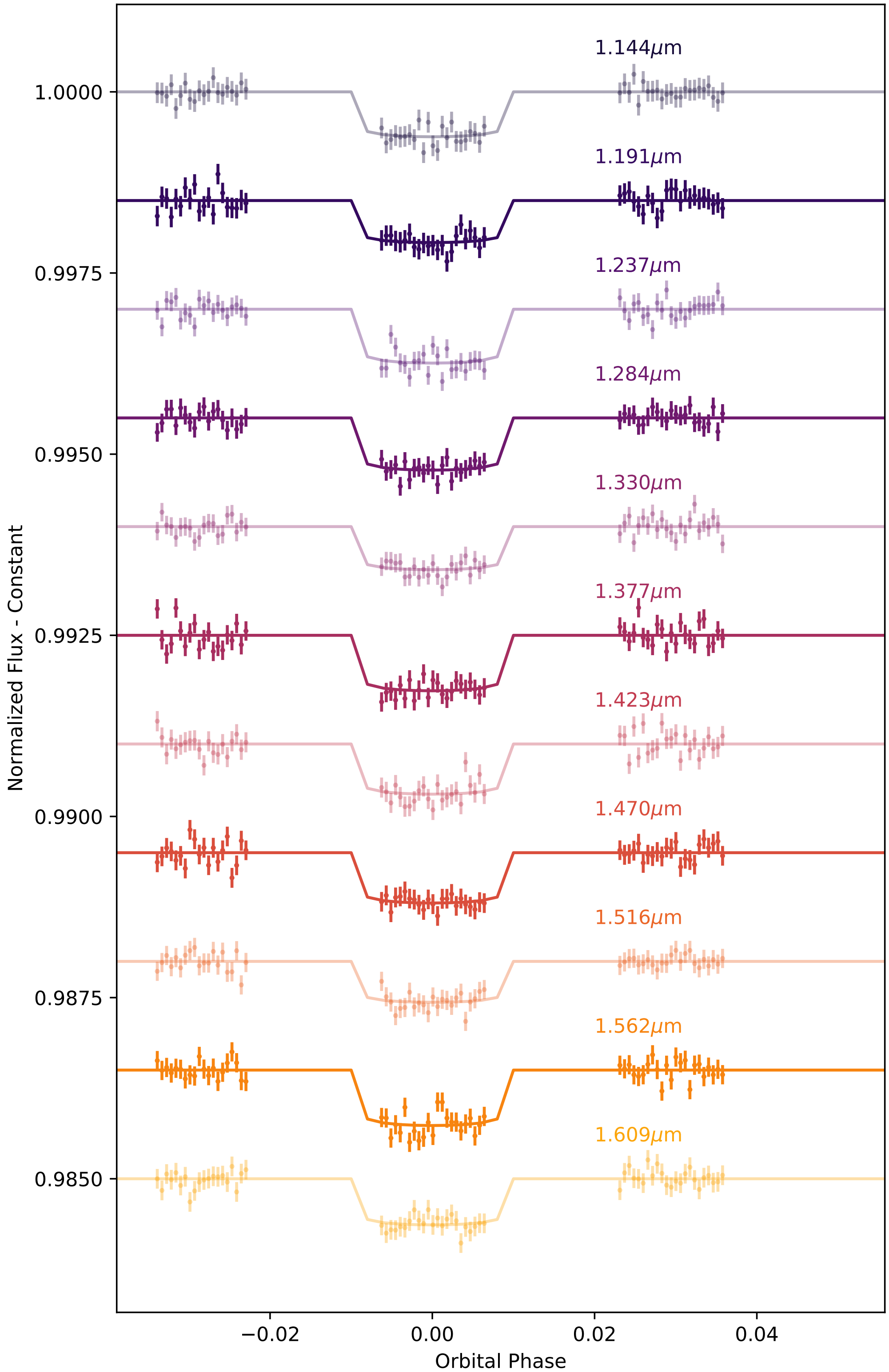
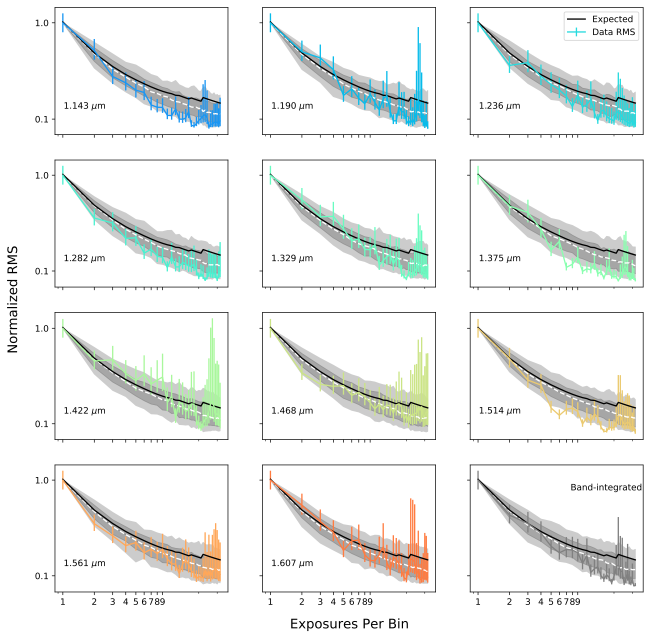


Figure A.8: Spectral light curves for L9859b, visit 03.

A.3 Red Noise Diagnostic Figures

The figures in this section visualize correlated noise analysis. See Section 2.4.3.1 for detailed discussion on how to interpret these figures.



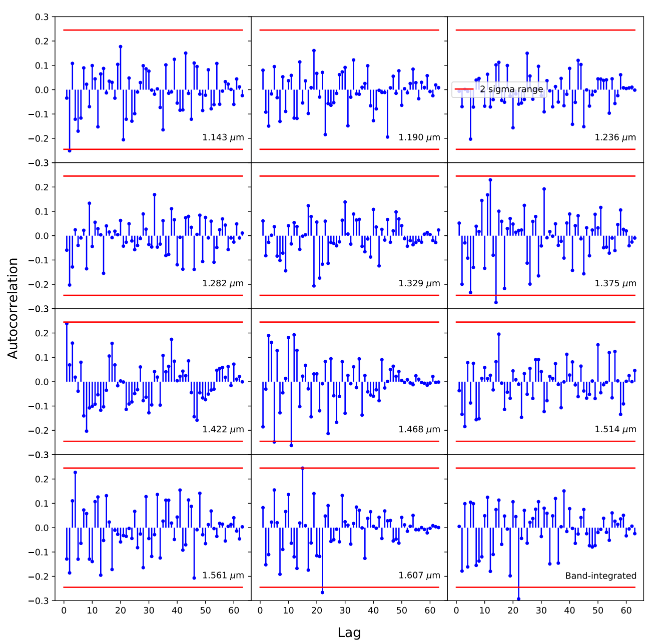
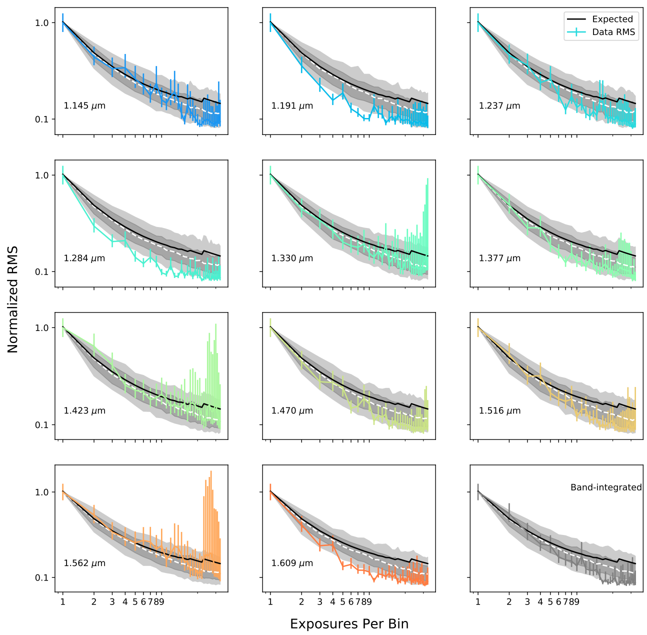


Figure A.9: Correlated Noise Diagnostic Figures for L9859b Visit 00. **Top:** Binning analysis for each spectral bin (see Section 2.4.3.1. **Bottom:** Autocorrelation function for each spectral bin.



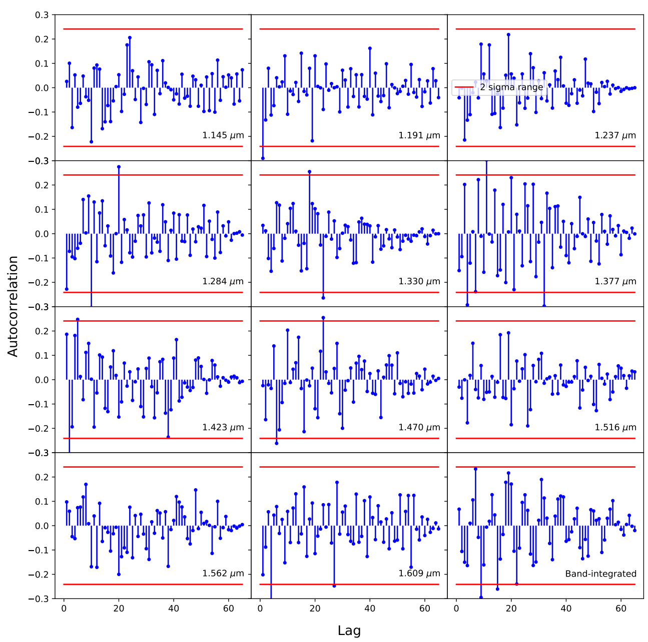


Figure A.10: Correlated Noise Diagnostic Figures for L9859b Visit 01. **Top:** Binning analysis for each spectral bin (see Section 2.4.3.1. **Bottom:** Autocorrelation function for each spectral bin.

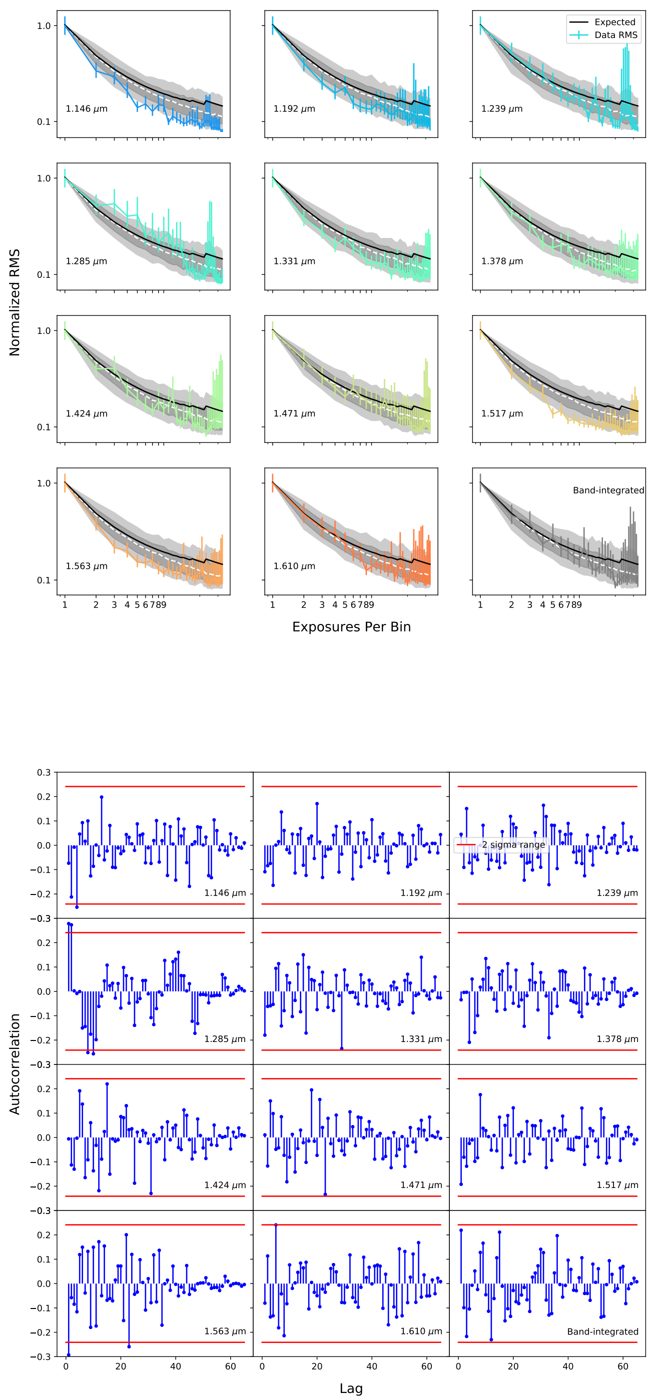


Figure A.11: Correlated Noise Diagnostic Figures for L9859b Visit 02. **Top:** Binning analysis for each spectral bin (see Section 2.4.3.1. **Bottom:** Autocorrelation function for each spectral bin.

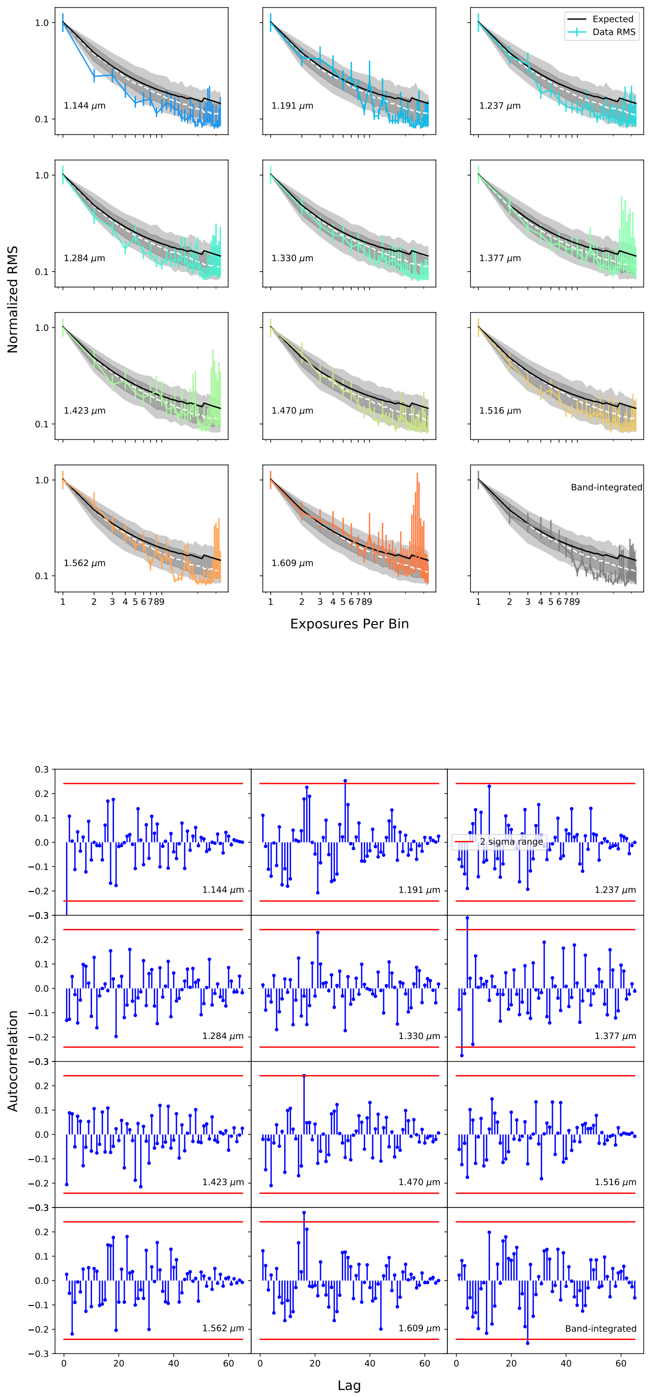


Figure A.12: Correlated Noise Diagnostic Figures for L9859b Visit 03. **Top:** Binning analysis for each spectral bin (see Section 2.4.3.1. **Bottom:** Autocorrelation function for each spectral bin.

Appendix B: Chapter 4 Supplementary Material

B.1 HAT-P-41b Transit HST Spectrophotometric Light Curve Fits

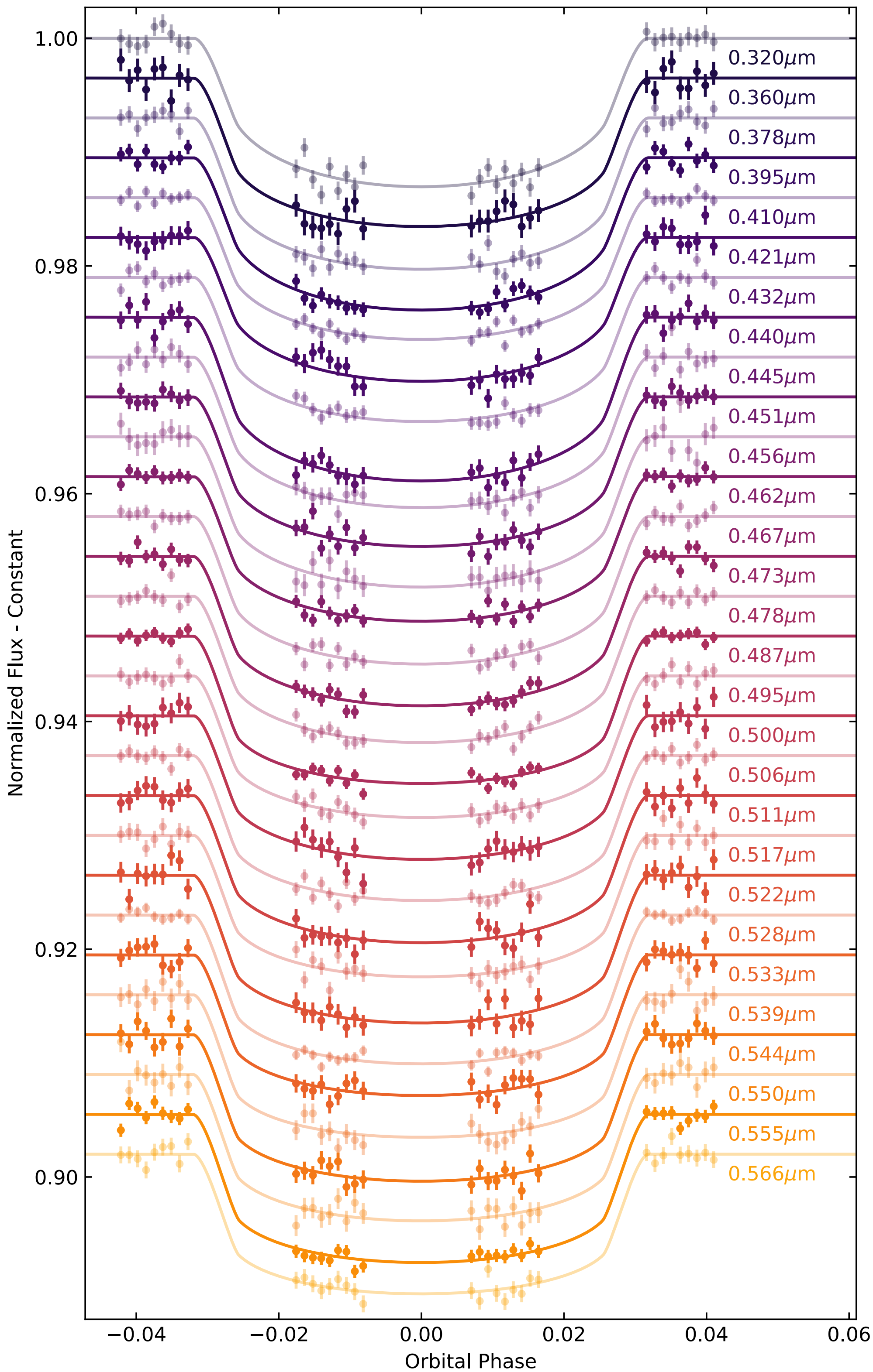


Figure B.1: Spectral light curves for STIS G430L, visit 83.

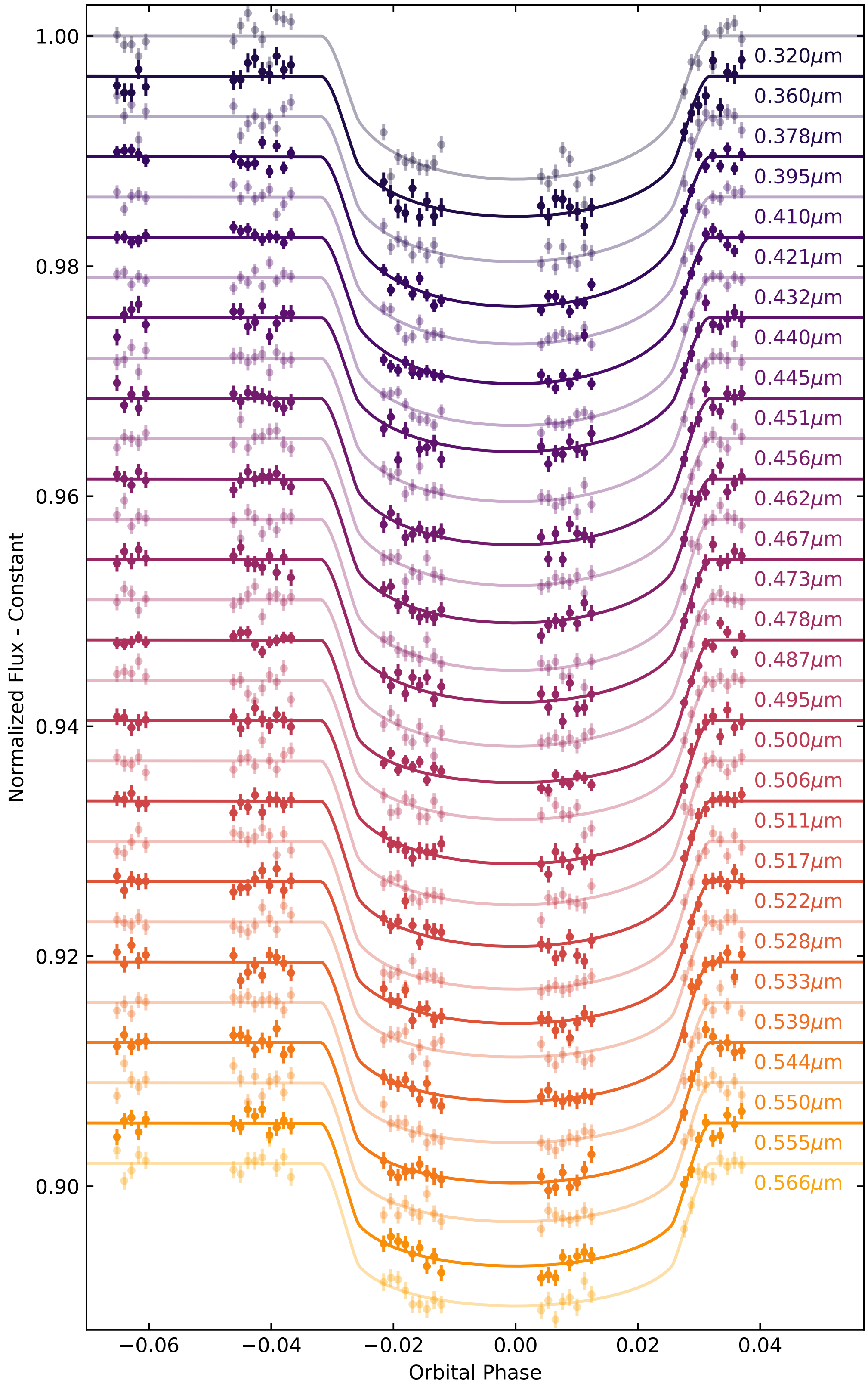


Figure B.2: Spectral light curves for STIS G430L, visit 84.

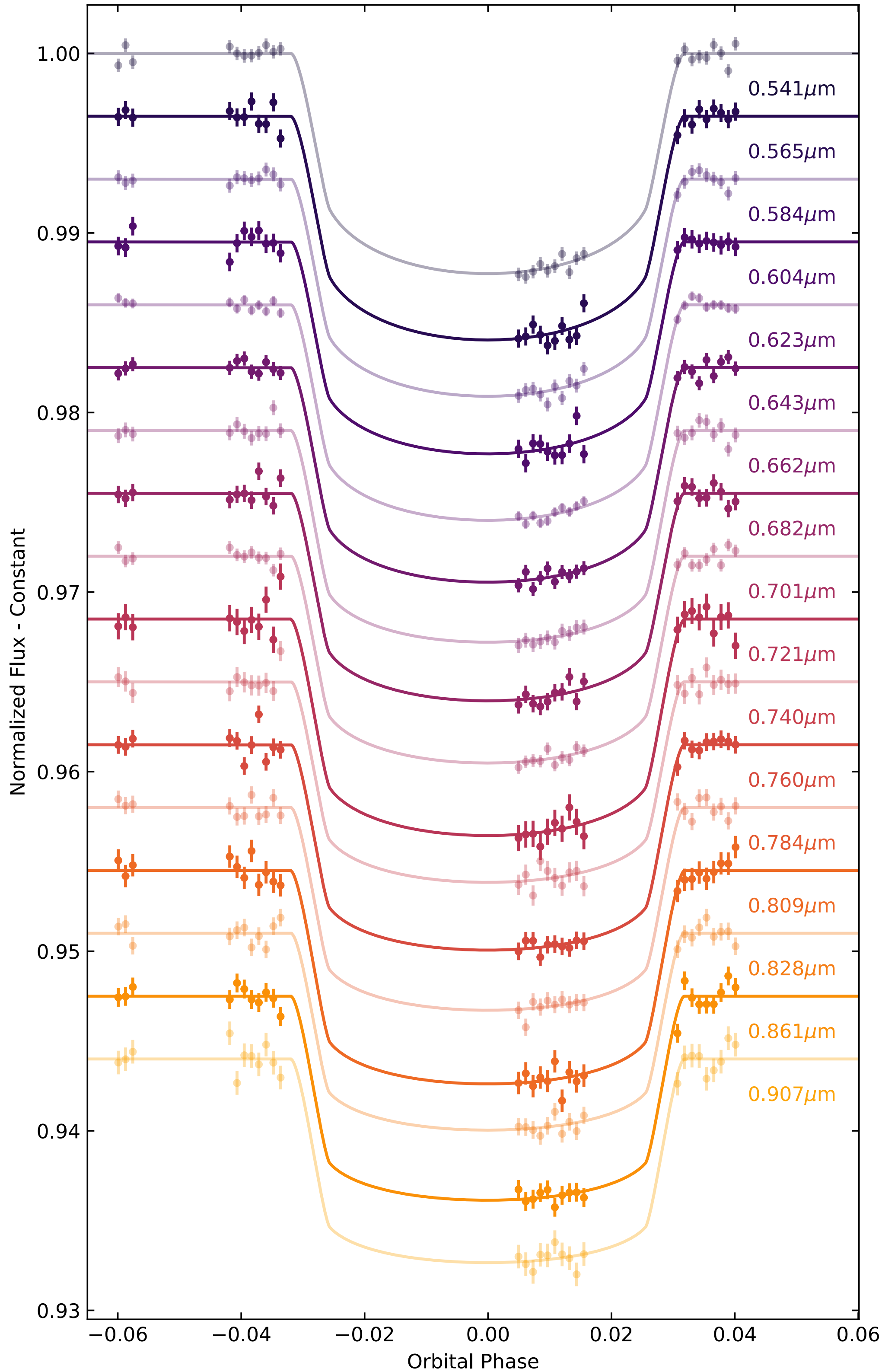


Figure B.3: Spectral light curves for STIS G750L, visit 85.

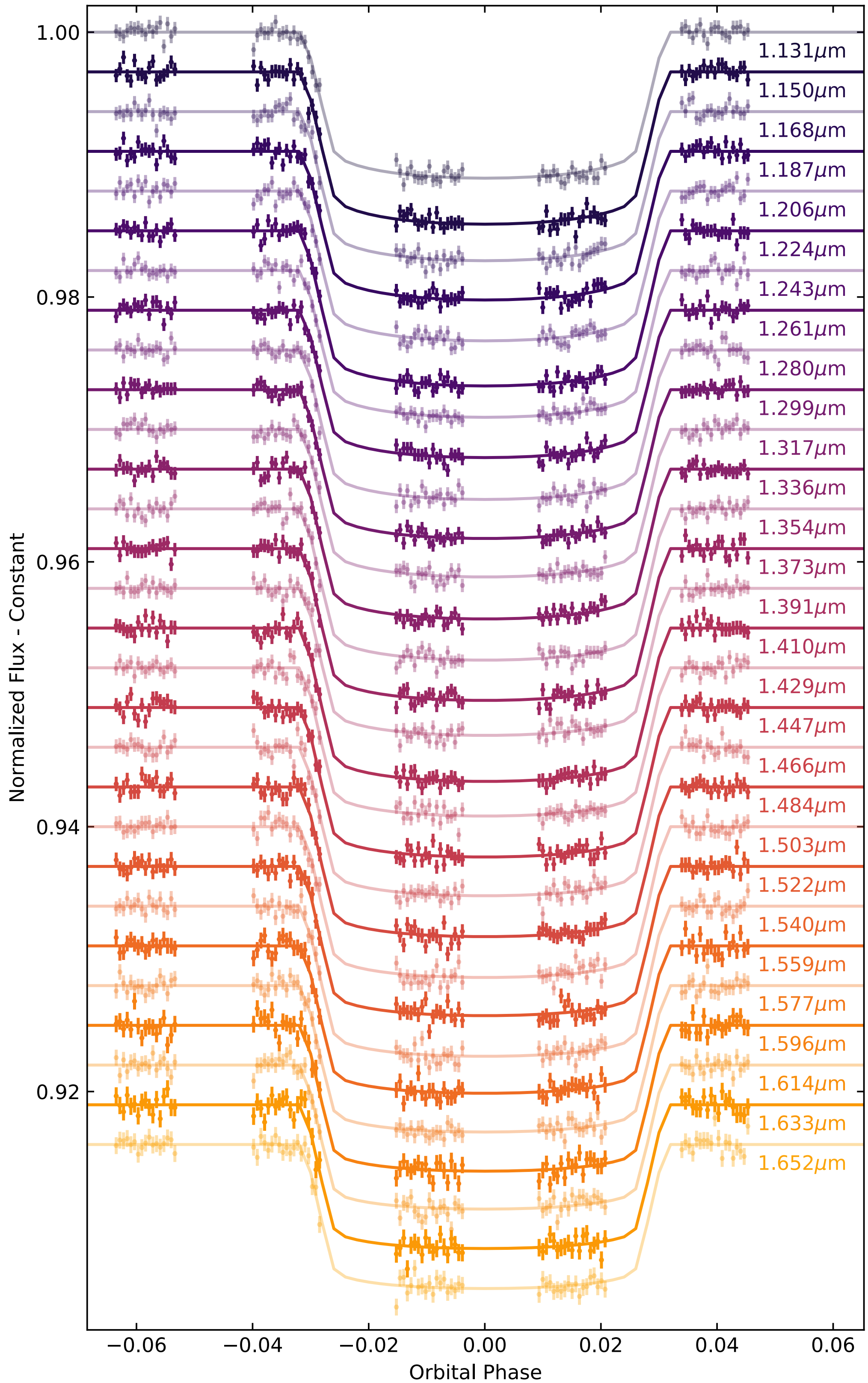


Figure B.4: Spectral light curves for single WFC3 visit.

Appendix C: Facilities and Software

1. Archival Data Used in Thesis

   All data are from the MAST Archive[1], and are provided with an associated proposal ID (GO), Principal Investigator (PI), and Digitial Object Identifier (DOI).

   - HAT-P-41b: HST STIS and WFC3 (GO 14767, PI Sing) and *Spitzer* (GO 13044, PI Deming) transit data: DOI 10.17909/t9-fg9z-er59
   - L9859 System: HST WFC3 (GO 15856, PI Barclay) transit data: DOI 10.17909/t9-xf60-w063
   - WASP-18b: HST WFC3 (GO 13467, PI Bean) eclipse data: DOI 10.17909/t9-4pmh-fd65
   - WASP-19b: HST WFC3 (GO 13431, PI Huitson) eclipse data: DOI 10.17909/t9-xnwm-hd84
   - WASP-18b and WASP-19b *Spitzer* (GO 60185, PI Maxted) eclipse data: DOI 10.17909/t9-kavg-yj38

2. Open source software used in thesis

[1] `https://archive.stsci.edu/`

- IRAF (Tody, 1986, 1993)
- SciPy (Jones et al., 2001–)
- Matplotlib (Hunter, 2007)
- `nestle` (`https://github.com/kbarbary/nestle`)
- `dynesty` (Higson et al., 2019)
- `BATMAN` (Kreidberg, 2015)
- Kapetyn (Terlouw & Vogelaar, 2015)
- `Corner.py` (Foreman-Mackey, 2016)
- `PLATON` (Zhang et al., 2019, 2020b)
- NumPy (Harris et al., 2020)
- Pandas (The Pandas Development Team, 2020)
- `mc3` (`https://github.com/pcubillos/mc3`)

3. Software I developed used in thesis

    - `DEFLATE` (`https://github.com/AstroSheppard/WFC3-analysis`)

Bibliography

Akaike, H. 1974, IEEE Transactions on Automatic Control, 19, 716

Alexoudi, X., Mallonn, M., von Essen, C., Turner, J. D., Keles, E., Southworth, J., Mancini, L., Ciceri, S., Granzer, T., Denker, C., Dineva, E., & Strassmeier, K. G. 2018, A&A, 620, A142, arXiv:1810.02172

Allan, D. W. 1966, IEEE Proceedings, 54, 221

Allard, F., Hauschildt, P. H., Baraffe, I., & Chabrier, G. 1996, ApJL, 465, L123

Allard, F., Homeier, D., & Freytag, B. 2012, Philosophical Transactions of the Royal Society of London Series A, 370, 2765, arXiv:1112.3591

Allard, N. F., Kielkopf, J. F., & Allard, F. 2007, European Physical Journal D, 44, 507

Allard, N. F., Spiegelman, F., & Kielkopf, J. F. 2016, A&A, 589, A21

Allard, N. F., Spiegelman, F., Leininger, T., & Molliere, P. 2019, A&A, 628, A120, arXiv:1908.01989

Amundsen, D. S., Baraffe, I., Tremblin, P., Manners, J., Hayek, W., Mayne, N. J., & Acreman, D. M. 2014, A&A, 564, A59, arXiv:1402.0814

Anderson, D. R., Smith, A. M. S., Madhusudhan, N., Wheatley, P. J., Collier Cameron, A., Hellier, C., Campo, C., Gillon, M., Harrington, J., Maxted, P. F. L., Pollacco, D., Queloz, D., Smalley, B., Triaud, A. H. M. J., & West, R. G. 2013, MNRAS, 430, 3422, arXiv:1112.5145

Andrae, R., Fouesneau, M., Creevey, O., Ordenovic, C., Mary, N., Burlacu, A., Chaoul, L., Jean-Antoine-Piccolo, A., Kordopatis, G., Korn, A., Lebreton, Y., Panem, C., Pichon, B., Thévenin, F., Walmsley, G., & Bailer-Jones, C. A. L. 2018, A&A, 616, A8, arXiv:1804.09374

Andrae, R., Schulze-Hartung, T., & Melchior, P. 2010, arXiv e-prints, arXiv:1012.3754, arXiv:1012.3754

Arcangeli, J., Désert, J.-M., Line, M. R., Bean, J. L., Parmentier, V., Stevenson, K. B., Kreidberg, L., Fortney, J. J., Mansfield, M., & Showman, A. P. 2018, ApJL, 855, L30, arXiv:1801.02489

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481, arXiv:0909.0948

Atreya, S. K., Crida, A., Guillot, T., Lunine, J. I., Madhusudhan, N., & Mousis, O. 2016, arXiv e-prints, arXiv:1606.04510, arXiv:1606.04510

Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58, arXiv:1804.10121

Barber, R. J., Strange, J. K., Hill, C., Polyansky, O. L., Mellau, G. C., Yurchenko, S. N., & Tennyson, J. 2014, MNRAS, 437, 1828, arXiv:1311.1328

Barber, R. J., Tennyson, J., Harris, G. J., & Tolchenov, R. N. 2006, MNRAS, 368, 1087, arXiv:astro-ph/0601236

Baron, E., Chen, B., & Hauschildt, P. H. 2010, PHOENIX: A General-purpose State-of-the-art Stellar and Planetary Atmosphere Code

Barstow, J. K., Aigrain, S., Irwin, P. G. J., & Sing, D. K. 2017, ApJ, 834, 50, arXiv:1610.01841

Barstow, J. K., Changeat, Q., Garland, R., Line, M. R., Rocchetto, M., & Waldmann, I. P. 2020, MNRAS, 493, 4884, arXiv:2002.01063

Batalha, N. E., Mandell, A., Pontoppidan, K., Stevenson, K. B., Lewis, N. K., Kalirai, J., Earl, N., Greene, T., Albert, L., & Nielsen, L. D. 2017, PASP, 129, 064501, arXiv:1702.01820

Batygin, K., Bodenheimer, P. H., & Laughlin, G. P. 2016, ApJ, 829, 114, arXiv:1511.09157

Bean, J. & FINESSE Science Team. 2017, in American Astronomical Society Meeting Abstracts, Vol. 229, American Astronomical Society Meeting Abstracts #229, 301.08

Bean, J. L., Raymond, S. N., & Owen, J. E. 2021, Journal of Geophysical Research (Planets), 126, e06639, arXiv:2010.11867

Beatty, T. G., Madhusudhan, N., Tsiaras, A., Zhao, M., Gilliland, R. L., Knutson, H. A., Shporer, A., & Wright, J. T. 2017, AJ, 154, 158, arXiv:1612.06409

Benneke, B. 2015, arXiv e-prints, arXiv:1504.07655, arXiv:1504.07655

Benneke, B., Knutson, H. A., Lothringer, J., Crossfield, I. J. M., Moses, J. I., Morley, C., Kreidberg, L., Fulton, B. J., Dragomir, D., Howard, A. W., Wong, I., Désert, J.-M., McCullough, P. R., Kempton, E. M. R., Fortney, J., Gilliland , R., Deming, D., & Kammer, J. 2019a, Nature Astronomy, 3, 813, arXiv:1907.00449

Benneke, B. & Seager, S. 2013, ApJ, 778, 153, arXiv:1306.6325

Benneke, B., Wong, I., Piaulet, C., Knutson, H. A., Lothringer, J., Morley, C. V., Crossfield, I. J. M., Gao, P., Greene, T. P., Dressing, C., Dragomir, D., Howard, A. W., McCullough, P. R., Kempton, E. M. R., Fortney, J. J., & Fraine, J. 2019b, ApJL, 887, L14, arXiv:1909.04642

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

Birkby, J. L. 2018, arXiv e-prints, arXiv:1806.04617, arXiv:1806.04617

Blecic, J., Dobbs-Dixon, I., & Greene, T. 2017, ApJ, 848, 127, arXiv:1803.06678

Bolmont, E., Selsis, F., Owen, J. E., Ribas, I., Raymond, S. N., Leconte, J., & Gillon, M. 2017, MNRAS, 464, 3728, arXiv:1605.00616

Booth, R. A., Clarke, C. J., Madhusudhan, N., & Ilee, J. D. 2017, MNRAS, 469, 3994, arXiv:1705.03305

Brewer, J. M. & Fischer, D. A. 2018a, ApJS, 237, 38, arXiv:1804.00673

—. 2018b, VizieR Online Data Catalog, J/ApJS/237/38

Brown, T. M. 2001, ApJ, 553, 1006, arXiv:astro-ph/0101307

Buchner, J., Georgakakis, A., Nandra, K., Hsu, L., Rangel, C., Brightman, M., Merloni, A., Salvato, M., Donley, J., & Kocevski, D. 2014, A&A, 564, A125, arXiv:1402.0004

Burrows, A., Hubeny, I., Budaj, J., Knutson, H. A., & Charbonneau, D. 2007, The Astrophysical Journal, 668, L171

Burrows, A., Rauscher, E., Spiegel, D. S., & Menou, K. 2010, ApJ, 719, 341, arXiv:1005.0346

Burrows, A. & Sharp, C. M. 1999, ApJ, 512, 843, arXiv:astro-ph/9807055

Carone, L., Baeyens, R., Mollière, P., Barth, P., Vazan, A., Decin, L., Sarkis, P., Venot, O., & Henning, T. 2020, MNRAS, 496, 3582, arXiv:1904.13334

Cartier, K. M. S., Beatty, T. G., Zhao, M., Line, M., Ngo, H., Mawet, D., Stassun, K. G., Wright, J. T., Kreidberg, L., Fortney, J., & Knutson, H. 2017, The Astronomical Journal, 153, 34

Castelli, F. & Kurucz, R. L. 2004, arXiv.org, 5087, arXiv:astro-ph/0405087

Chachan, Y., Knutson, H. A., Gao, P., Kataria, T., Wong, I., Henry, G. W., Benneke, B., Zhang, M., Barstow, J., Bean, J. L., Mikal-Evans, T., Lewis, N. K., Mansfield, M., López-Morales, M., Nikolov, N., Sing, D. K., & Wakeford, H. 2019a, AJ, 158, 244, arXiv:1910.07523

—. 2019b, AJ, 158, 244, arXiv:1910.07523

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, The Astrophysical Journal, 626, 523

Charbonneau, D., Berta, Z. K., Irwin, J., Burke, C. J., Nutzman, P., Buchhave, L. A., Lovis, C., Bonfils, X., Latham, D. W., Udry, S., Murray-Clay, R. A., Holman, M. J., Falco, E. E., Winn, J. N., Queloz, D., Pepe, F., Mayor, M., Delfosse, X., & Forveille, T. 2009, Nature, 462, 891, arXiv:0912.3229

Charbonneau, D., Brown, T. M., Noyes, R. W., & Gilliland, R. L. 2002, ApJ, 568, 377, arXiv:astro-ph/0111544

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

Claret, A., Hauschildt, P. H., & Witte, S. 2012, A&A, 546, A14

—. 2013, A&A, 552, A16

Cloutier, R., Astudillo-Defru, N., Bonfils, X., Jenkins, J. S., Berdiñas, Z., Ricker, G., Vanderspek, R., Latham, D. W., Seager, S., Winn, J., Jenkins, J. M., Almenara, J. M., Bouchy, F., Delfosse, X., Díaz, M. R., Díaz, R. F., Doyon, R., Figueira, P., Forveille, T., Kurtovic, N. T., Lovis, C., Mayor, M., Menou, K., Morgan, E., Morris, R., Muirhead, P., Murgas, F., Pepe, F., Santos, N. C., Ségransan, D., Smith, J. C., Tenenbaum, P., Torres, G., Udry, S., Vezie, M., & Villasenor, J. 2019, A&A, 629, A111, arXiv:1905.10669

Cowan, N. B. & Agol, E. 2011, ApJ, 729, 54, arXiv:1001.0012

Crossfield, I. J. M. 2015, PASP, 127, 941, arXiv:1507.03966

Crouzet, N., McCullough, P. R., Deming, D., & Madhusudhan, N. 2014, ApJ, 795, 166, arXiv:1409.4000

Cubillos, P., Harrington, J., Loredo, T. J., Lust, N. B., Blecic, J., & Stemm, M. 2017, AJ, 153, 3, arXiv:1610.01336

de Wit, J., Wakeford, H. R., Lewis, N. K., Delrez, L., Gillon, M., Selsis, F., Leconte, J., Demory, B.-O., Bolmont, E., Bourrier, V., Burgasser, A. J., Grimm, S., Jehin, E., Lederer, S. M., Owen, J. E., Stamenković, V., & Triaud, A. H. M. J. 2018, Nature Astronomy, 2, 214, arXiv:1802.02250

Deming, D., Knutson, H., Kammer, J., Fulton, B. J., Ingalls, J., Carey, S., Burrows, A., Fortney, J. J., Todorov, K., Agol, E., Cowan, N., Desert, J.-M., Fraine, J., Langton, J., Morley, C., & Showman, A. P. 2015, ApJ, 805, 132, arXiv:1411.7404

Deming, D., Knutson, H., Kammer, J., Fulton, B. J., Ingalls, J., Carey, S., Burrows, A., Fortney, J. J., Todorov, K., Agol, E., Cowan, N., Desert, J.-M., Fraine, J., Langton, J., Morley, C., & Showman, A. P. 2015, The Astrophysical Journal, 805, 132

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

Deming, D. & Sheppard, K. 2017, ApJL, 841, L3, arXiv:1705.00625

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., Mand ell, A. M., Ranjan, S., Charbonneau, D., Clampin, M., Seager, S., & Showman, A. P. 2013, ApJ, 774, 95, arXiv:1302.1141

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, The Astrophysical Journal, 774, 95

Deming, L. D. & Seager, S. 2017, Journal of Geophysical Research (Planets), 122, 53

Diamond-Lowe, H., Berta-Thompson, Z., Charbonneau, D., & Kempton, E. M. R. 2018, AJ, 156, 42, arXiv:1805.07328

Duncan, D. K., Vaughan, A. H., Wilson, O. C., Preston, G. W., Frazer, J., Lanning, H., Misch, A., Mueller, J., Soyumer, D., Woodard, L., Baliunas, S. L., Noyes, R. W., Hartmann, L. W., Porter, A., Zwaan, C., Middelkoop, F., Rutten, R. G. M., & Mihalas, D. 1991, ApJS, 76, 383

Eaton, J. A., Henry, G. W., & Fekel, F. C. 2003, Astrophysics and Space Science Library, Vol. 288, Advantages of Automated Observing with Small Telescopes (Kluwer Academic Publishers), 189

Espinoza, N. & Jones, K. 2021, arXiv e-prints, arXiv:2106.15687, arXiv:2106.15687

Espinoza, N., Rackham, B. V., Jordán, A., Apai, D., López-Morales, M., Osip, D. J., Grimm, S. L., Hoeijmakers, J., Wilson, P. A., Bixel, A., McGruder, C., Rodler, F., Weaver, I., Lewis, N. K., Fortney, J. J., & Fraine, J. 2019, MNRAS, 482, 2065, arXiv:1807.10652

Evans, D. F., Southworth, J., Maxted, P. F. L., Skottfelt, J., Hundertmark, M., Jørgensen, U. G., Dominik, M., Alsubai, K. A., Andersen, M. I., Bozza, V., Bramich, D. M., Burgdorf, M. J., Ciceri, S., D'Ago, G., Figuera Jaimes, R., Gu, S. H., Haugbølle, T., Hinse, T. C., Juncher, D., Kains, N., Kerins, E., Korhonen, H., Kuffmeier, M., Mancini, L., Peixinho, N., Popovas, A., Rabus, M., Rahvar, S., Schmidt, R. W., Snodgrass, C., Starkey, D., Surdej, J., Tronsgaard, R., von Essen, C., Wang, Y.-B., & Wertz, O. 2016, A&A, 589, A58, arXiv:1603.03274

Evans, D. F., Southworth, J., Smalley, B., Jørgensen, U. G., Dominik, M., Andersen, M. I., Bozza, V., Bramich, D. M., Burgdorf, M. J., Ciceri, S., D'Ago, G., Figuera Jaimes, R., Gu, S. H., Hinse, T. C., Henning, T., Hundertmark, M., Kains, N., Kerins, E., Korhonen, H., Kokotanekova, R., Kuffmeier, M., Longa-Peña, P., Mancini, L., MacKenzie, J., Popovas, A., Rabus, M., Rahvar, S., Sajadian, S., Snodgrass, C., Skottfelt, J., Surdej, J., Tronsgaard, R., Unda-Sanzana, E., von Essen, C., Wang, Y.-B., & Wertz, O. 2018, A&A, 610, A20, arXiv:1709.07476

Evans, T. M., Sing, D. K., Kataria, T., Goyal, J., Nikolov, N., Wakeford, H. R., Deming, D., Marley, M. S., Amundsen, D. S., Ballester, G. E., Barstow, J. K., Ben-Jaffel, L., Bourrier, V., Buchhave, L. A., Cohen, O., Ehrenreich, D., García Muñoz, A., Henry, G. W., Knutson, H., Lavvas, P., Lecavelier Des Etangs, A., Lewis, N. K., López-Morales, M., Mandell, A. M., Sanz-Forcada, J., Tremblin, P., & Lupu, R. 2017, Nature, 548, 58, arXiv:1708.01076

Evans, T. M., Sing, D. K., Kataria, T., Goyal, J., Nikolov, N., Wakeford, H. R., Deming, D., Marley, M. S., Amundsen, D. S., Ballester, G. E., Barstow, J. K., Ben-Jaffel, L., Bourrier, V., Buchhave, L. A., Cohen, O., Ehrenreich, D., Garcia Munoz, A., Henry, G. W., Knutson, H., Lavvas, P., Etangs, A. L. d., Lewis, N. K., López-Morales, M., Mandell, A. M., Sanz-Forcada, J., Tremblin, P., & Lupu, R. 2017, Nature, 548, 58

Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2013, ArXiv e-prints, arXiv:1306.2144

Fisher, C. & Heng, K. 2018, MNRAS, 481, 4698, arXiv:1809.06894

Ford, E. B. 2005, The Astronomical Journal, 129, 1706

Foreman-Mackey, D. 2016, The Journal of Open Source Software, 1, 24

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, arXiv:1202.3665

Fortney, J. J. 2005, MNRAS, 364, 649, arXiv:astro-ph/0509292

Fortney, J. J., Dawson, R. I., & Komacek, T. D. 2021, Journal of Geophysical Research (Planets), 126, e06629, arXiv:2102.05064

Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, The Astrophysical Journal, 678, 1419

Fortney, J. J., Mordasini, C., Nettelmann, N., Kempton, E. M. R., Greene, T. P., & Zahnle, K. 2013, ApJ, 775, 80, arXiv:1306.4329

Fortney, J. J., Shabram, M., Showman, A. P., Lian, Y., Freedman, R. S., Marley, M. S., & Lewis, N. K. 2010, ApJ, 709, 1396, arXiv:0912.2350

Fraine, J., Deming, D., Benneke, B., Knutson, H., Jordán, A., Espinoza, N., Madhusudhan, N., Wilkins, A., & Todorov, K. 2014, Nature, 513, 526, arXiv:1409.8349

Fu, G., Deming, D., Knutson, H., Madhusudhan, N., Mandell, A., & Fraine, J. 2017, ApJL, 847, L22, arXiv:1709.07385

Fulton, B. J., Petigura, E. A., Howard, A. W., Isaacson, H., Marcy, G. W., Cargile, P. A., Hebb, L., Weiss, L. M., Johnson, J. A., Morton, T. D., Sinukoff, E., Crossfield, I. J. M., & Hirsch, L. A. 2017, AJ, 154, 109, arXiv:1703.10375

Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., Babusiaux, C., Bailer-Jones, C. A. L., Biermann, M., Evans, D. W., Eyer, L., Jansen, F., Jordi, C., Klioner, S. A., Lammers, U., Lindegren, L., Luri, X., Mignard, F., Panem, C., Pourbaix, D., Randich, S., Sartoretti, P., Siddiqui, H. I., Soubiran, C., van Leeuwen, F., Walton, N. A., Arenou, F., Bastian, U., Cropper, M., Drimmel, R., Katz, D., Lattanzi, M. G., Bakker, J., Cacciari, C., Castañeda, J., Chaoul, L., Cheek, N., De Angeli, F., Fabricius, C., Guerra, R., Holl, B., Masana, E., Messineo, R., Mowlavi, N., Nienartowicz, K., Panuzzo, P., Portell, J., Riello, M., Seabroke, G. M., Tanga, P., Thévenin, F., Gracia-Abril, G., Comoretto, G., Garcia-Reinaldos, M., Teyssier, D., Altmann, M., Andrae, R., Audard, M., Bellas-Velidis, I., Benson, K., Berthier, J., Blomme, R., Burgess, P., Busso, G., Carry, B., Cellino, A., Clementini, G., Clotet, M., Creevey, O., Davidson, M., De Ridder, J., Delchambre, L., Dell'Oro, A., Ducourant, C., Fernández-Hernández, J., Fouesneau, M., Frémat, Y., Galluccio, L., García-Torres, M., González-Núñez, J., González-Vidal, J. J., Gosset, E., Guy, L. P., Halbwachs, J. L., Hambly, N. C., Harrison, D. L., Hernández, J., Hestroffer, D., Hodgkin, S. T., Hutton, A., Jasniewicz, G., Jean-Antoine-Piccolo, A., Jordan, S., Korn, A. J., Krone-Martins, A., Lanzafame, A. C., Lebzelter, T., Löffler, W., Manteiga, M., Marrese, P. M., Martín-Fleitas, J. M., Moitinho, A., Mora, A., Muinonen, K., Osinde, J., Pancino, E., Pauwels, T., Petit, J. M., Recio-Blanco, A., Richards, P. J., Rimoldini, L., Robin, A. C., Sarro, L. M., Siopis, C., Smith, M., Sozzetti, A., Süveges, M., Torra, J., van Reeven, W., Abbas, U., Abreu Aramburu, A., Accart, S., Aerts, C., Altavilla, G., Álvarez, M. A., Alvarez, R., Alves, J., Anderson, R. I., Andrei, A. H., Anglada Varela, E., Antiche, E., Antoja, T., Arcay, B., Astraatmadja, T. L., Bach, N., Baker, S. G., Balaguer-Núñez, L., Balm, P., Barache, C., Barata, C., Barbato, D., Barblan, F., Barklem, P. S., Barrado, D., Barros, M., Barstow, M. A., Bartholomé Muñoz, S., Bassilana, J. L., Becciani, U., Bellazzini, M., Berihuete, A., Bertone, S., Bianchi, L., Bienaymé, O., Blanco-Cuaresma, S., Boch, T., Boeche, C., Bombrun, A., Borrachero, R., Bossini, D., Bouquillon, S., Bourda, G., Bragaglia, A., Bramante, L., Breddels, M. A., Bressan, A., Brouillet, N., Brüsemeister, T., Brugaletta, E., Bucciarelli, B., Burlacu, A., Busonero, D., Butkevich, A. G., Buzzi, R., Caffau, E., Cancelliere, R., Cannizzaro, G., Cantat-Gaudin, T., Carballo, R., Carlucci, T., Carrasco, J. M., Casamiquela, L., Castellani, M., Castro-Ginard, A., Charlot, P., Chemin, L., Chiavassa, A., Cocozza, G., Costigan, G., Cowell, S., Crifo, F., Crosta, M., Crowley, C., Cuypers, J., Dafonte, C., Damerdji, Y., Dapergolas, A., David, P., David, M., de Laverny, P., De Luise, F., De March, R., de Martino, D., de Souza, R., de Torres, A., Debosscher, J., del Pozo, E., Delbo, M., Delgado, A., Delgado, H. E., Di Matteo, P., Diakite, S., Diener, C., Distefano, E., Dolding, C., Drazinos, P., Durán, J., Edvardsson, B., Enke, H., Eriksson, K., Esquej, P., Eynard Bontemps, G., Fabre, C., Fabrizio, M., Faigler, S., Falcão, A. J., Farràs Casas, M., Federici, L., Fedorets, G., Fernique, P., Figueras, F., Filippi, F., Findeisen, K., Fonti, A., Fraile, E., Fraser, M., Frézouls, B., Gai, M., Galleti, S., Garabato, D., García-Sedano, F., Garofalo, A., Garralda, N., Gavel, A., Gavras, P., Gerssen, J., Geyer, R., Giacobbe, P., Gilmore, G., Girona, S., Giuffrida, G., Glass, F., Gomes, M., Granvik, M., Gueguen, A., Guerrier, A., Guiraud, J., Gutiérrez-Sánchez, R., Haigron, R., Hatzidimitriou, D., Hauser, M., Haywood, M., Heiter, U., Helmi, A., Heu, J., Hilger, T., Hobbs, D., Hofmann, W., Holland, G., Huckle, H. E., Hypki, A., Icardi, V., Janßen, K., Jevardat de Fombelle, G., Jonker, P. G., Juhász, Á. L., Julbe, F., Karampelas, A., Kewley, A., Klar, J., Kochoska, A., Kohley, R., Kolenberg, K., Kontizas, M., Kontizas, E., Koposov, S. E., Kordopatis, G., Kostrzewa-Rutkowska, Z., Koubsky, P., Lambert, S., Lanza, A. F., Lasne, Y., Lavigne, J. B., Le Fustec, Y., Le Poncin-Lafitte, C., Lebreton, Y., Leccia, S., Leclerc, N., Lecoeur-Taibi, I., Lenhardt, H., Leroux, F., Liao, S., Licata, E., Lindstrøm, H. E. P., Lister, T. A., Livanou, E., Lobel, A., López, M., Managau, S., Mann, R. G., Mantelet, G., Marchal, O., Marchant, J. M., Marconi, M., Marinoni, S., Marschalkó, G., Marshall, D. J., Martino, M., Marton, G., Mary, N., Massari, D., Matijevič, G., Mazeh, T., McMillan, P. J., Messina, S., Michalik, D., Millar, N. R., Molina, D., Molinaro, R., Molnár, L., Montegriffo, P., Mor, R., Morbidelli, R., Morel, T., Morris, D., Mulone, A. F., Muraveva, T., Musella, I., Nelemans, G., Nicastro, L., Noval, L., O'Mullane, W., Ordénovic, C., Ordóñez-Blanco, D., Osborne, P., Pagani, C., Pagano, I., Pailler, F., Palacin, H., Palaversa, L., Panahi, A., Pawlak, M., Piersimoni, A. M., Pineau, F. X., Plachy, E., Plum, G., Poggio, E., Poujoulet, E., Prša, A., Pulone, L., Racero, E., Ragaini, S., Rambaux, N., Ramos-Lerate, M., Regibo, S., Reylé, C., Riclet, F., Ripepi, V., Riva, A., Rivard, A., Rixon, G., Roegiers, T., Roelens, M., Romero-Gómez, M., Rowell, N., Royer, F., Ruiz-Dern, L., Sadowski, G., Sagristà Sellés, T., Sahlmann, J., Salgado, J., Salguero, E., Sanna, N., Santana-Ros, T., Sarasso, M., Savietto, H., Schultheis, M., Sciacca, E., Segol, M., Segovia, J. C., Ségransan, D., Shih, I. C., Siltala, L., Silva, A. F., Smart, R. L., Smith, K. W., Solano, E., Solitro, F., Sordo, R., Soria Nieto, S., Souchay, J., Spagna, A., Spoto, F., Stampa, U., Steele, I. A., Steidelmüller, H., Stephenson, C. A., Stoev, H., Suess, F. F., Surdej, J., Szabados, L., Szegedi-Elek, E., Tapiador, D., Taris, F., Tauran, G., Taylor, M. B., Teixeira, R., Terrett, D., Teyssand ier, P., Thuillot, W., Titarenko, A., Torra Clotet, F., Turon, C., Ulla, A., Utrilla, E., Uzzi, S., Vaillant, M., Valentini, G., Valette, V., van Elteren, A., Van Hemelryck, E., van Leeuwen, M., Vaschetto, M., Vecchiato, A., Veljanoski, J., Viala, Y., Vicente, D., Vogt, S., von Essen, C., Voss, H., Votruba, V., Voutsinas, S., Walmsley, G., Weiler, M., Wertz, O., Wevers, T., Wyrzykowski, Ł., Yoldas, A., Žerjal, M., Ziaeepour, H., Zorec, J., Zschocke, S., Zucker, S., Zurbach, C., & Zwitter, T. 2018, A&A, 616, A1, arXiv:1804.09365

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., Brown, A. G. A., Vallenari, A., Babusiaux, C., Bailer-Jones, C. A. L., Bastian, U., Biermann, M., Evans, D. W., Eyer, L., Jansen, F., Jordi, C., Klioner, S. A., Lammers, U., Lindegren, L., Luri, X., Mignard, F., Milligan, D. J., Panem, C., Poinsignon, V., Pourbaix, D., Randich, S., Sarri, G., Sartoretti, P., Siddiqui, H. I., Soubiran, C., Valette, V., van Leeuwen, F., Walton, N. A., Aerts, C., Arenou, F., Cropper, M., Drimmel, R., Høg, E., Katz, D., Lattanzi, M. G., O'Mullane, W., Grebel, E. K., Holland, A. D., Huc, C., Passot, X., Bramante, L., Cacciari, C., Castañeda, J., Chaoul, L., Cheek, N., De Angeli, F., Fabricius, C., Guerra, R., Hernández, J., Jean-Antoine-Piccolo, A., Masana, E., Messineo, R., Mowlavi, N., Nienartowicz, K., Ordóñez-Blanco, D., Panuzzo, P., Portell, J., Richards, P. J., Riello, M., Seabroke, G. M., Tanga, P., Thévenin, F., Torra, J., Els, S. G., Gracia-Abril, G., Comoretto, G., Garcia-Reinaldos, M., Lock, T., Mercier, E., Altmann, M., Andrae, R., Astraatmadja, T. L., Bellas-Velidis, I., Benson, K., Berthier, J., Blomme, R., Busso, G., Carry, B., Cellino, A., Clementini, G., Cowell, S., Creevey, O., Cuypers, J., Davidson, M., De Ridder, J., de Torres, A., Delchambre, L., Dell'Oro, A., Ducourant, C., Frémat, Y., García-Torres, M., Gosset, E., Halbwachs, J. L., Hambly, N. C., Harrison, D. L., Hauser, M., Hestroffer, D., Hodgkin, S. T., Huckle, H. E., Hutton, A., Jasniewicz, G., Jordan, S., Kontizas, M., Korn, A. J., Lanzafame, A. C., Manteiga, M., Moitinho, A., Muinonen, K., Osinde, J., Pancino, E., Pauwels, T., Petit, J. M., Recio-Blanco, A., Robin, A. C., Sarro, L. M., Siopis, C., Smith, M., Smith, K. W., Sozzetti, A., Thuillot, W., van Reeven, W., Viala, Y., Abbas, U., Abreu Aramburu, A., Accart, S., Aguado, J. J., Allan, P. M., Allasia, W., Altavilla, G., Álvarez, M. A., Alves, J., Anderson, R. I., Andrei, A. H., Anglada Varela, E., Antiche, E., Antoja, T., Antón, S., Arcay, B., Atzei, A., Ayache, L., Bach, N., Baker, S. G., Balaguer-Núñez, L., Barache, C., Barata, C., Barbier, A., Barblan, F., Baroni, M., Barrado y Navascués, D., Barros, M., Barstow, M. A., Becciani, U., Bellazzini, M., Bellei, G., Bello García, A., Belokurov, V., Bendjoya, P., Berihuete, A., Bianchi, L., Bienaymé, O., Billebaud, F., Blagorodnova, N., Blanco-Cuaresma, S., Boch, T., Bombrun, A., Borrachero, R., Bouquillon, S., Bourda, G., Bouy, H., Bragaglia, A., Breddels, M. A., Brouillet, N., Brüsemeister, T., Bucciarelli, B., Budnik, F., Burgess, P., Burgon, R., Burlacu, A., Busonero, D., Buzzi, R., Caffau, E., Cambras, J., Campbell, H., Cancelliere, R., Cantat-Gaudin, T., Carlucci, T., Carrasco, J. M., Castellani, M., Charlot, P., Charnas, J., Charvet, P., Chassat, F., Chiavassa, A., Clotet, M., Cocozza, G., Collins, R. S., Collins, P., Costigan, G., Crifo, F., Cross, N. J. G., Crosta, M., Crowley, C., Dafonte, C., Damerdji, Y., Dapergolas, A., David, P., David, M., De Cat, P., de Felice, F., de Laverny, P., De Luise, F., De March, R., de Martino, D., de

Souza, R., Debosscher, J., del Pozo, E., Delbo, M., Delgado, A., Delgado, H. E., di Marco, F., Di Matteo, P., Diakite, S., Distefano, E., Dolding, C., Dos Anjos, S., Drazinos, P., Durán, J., Dzigan, Y., Ecale, E., Edvardsson, B., Enke, H., Erdmann, M., Escolar, D., Espina, M., Evans, N. W., Eynard Bontemps, G., Fabre, C., Fabrizio, M., Faigler, S., Falcão, A. J., Farràs Casas, M., Faye, F., Federici, L., Fedorets, G., Fernández-Hernández, J., Fernique, P., Fienga, A., Figueras, F., Filippi, F., Findeisen, K., Fonti, A., Fouesneau, M., Fraile, E., Fraser, M., Fuchs, J., Furnell, R., Gai, M., Galleti, S., Galluccio, L., Garabato, D., García-Sedano, F., Garé, P., Garofalo, A., Garralda, N., Gavras, P., Gerssen, J., Geyer, R., Gilmore, G., Girona, S., Giuffrida, G., Gomes, M., González-Marcos, A., González-Núñez, J., González-Vidal, J. J., Granvik, M., Guerrier, A., Guillout, P., Guiraud, J., Gúrpide, A., Gutiérrez-Sánchez, R., Guy, L. P., Haigron, R., Hatzidimitriou, D., Haywood, M., Heiter, U., Helmi, A., Hobbs, D., Hofmann, W., Holl, B., Holland , G., Hunt, J. A. S., Hypki, A., Icardi, V., Irwin, M., Jevardat de Fombelle, G., Jofré, P., Jonker, P. G., Jorissen, A., Julbe, F., Karampelas, A., Kochoska, A., Kohley, R., Kolenberg, K., Kontizas, E., Koposov, S. E., Kordopatis, G., Koubsky, P., Kowalczyk, A., Krone-Martins, A., Kudryashova, M., Kull, I., Bachchan, R. K., Lacoste-Seris, F., Lanza, A. F., Lavigne, J. B., Le Poncin-Lafitte, C., Lebreton, Y., Lebzelter, T., Leccia, S., Leclerc, N., Lecoeur-Taibi, I., Lemaitre, V., Lenhardt, H., Leroux, F., Liao, S., Licata, E., Lindstrøm, H. E. P., Lister, T. A., Livanou, E., Lobel, A., Löffler, W., López, M., Lopez-Lozano, A., Lorenz, D., Loureiro, T., MacDonald, I., Magalhães Fernandes, T., Managau, S., Mann, R. G., Mantelet, G., Marchal, O., Marchant, J. M., Marconi, M., Marie, J., Marinoni, S., Marrese, P. M., Marschalkó, G., Marshall, D. J., Martín-Fleitas, J. M., Martino, M., Mary, N., Matijevič, G., Mazeh, T., McMillan, P. J., Messina, S., Mestre, A., Michalik, D., Millar, N. R., Miranda, B. M. H., Molina, D., Molinaro, R., Molinaro, M., Molnár, L., Moniez, M., Montegriffo, P., Monteiro, D., Mor, R., Mora, A., Morbidelli, R., Morel, T., Morgenthaler, S., Morley, T., Morris, D., Mulone, A. F., Muraveva, T., Musella, I., Narbonne, J., Nelemans, G., Nicastro, L., Noval, L., Ordénovic, C., Ordieres-Meré, J., Osborne, P., Pagani, C., Pagano, I., Pailler, F., Palacin, H., Palaversa, L., Parsons, P., Paulsen, T., Pecoraro, M., Pedrosa, R., Pentikäinen, H., Pereira, J., Pichon, B., Piersimoni, A. M., Pineau, F. X., Plachy, E., Plum, G., Poujoulet, E., Prša, A., Pulone, L., Ragaini, S., Rago, S., Rambaux, N., Ramos-Lerate, M., Ranalli, P., Rauw, G., Read, A., Regibo, S., Renk, F., Reylé, C., Ribeiro, R. A., Rimoldini, L., Ripepi, V., Riva, A., Rixon, G., Roelens, M., Romero-Gómez, M., Rowell, N., Royer, F., Rudolph, A., Ruiz-Dern, L., Sadowski, G., Sagristà Sellés, T., Sahlmann, J., Salgado, J., Salguero, E., Sarasso, M., Savietto, H., Schnorhk, A., Schultheis, M., Sciacca, E., Segol, M., Segovia, J. C., Segransan, D., Serpell, E., Shih, I. C., Smareglia, R., Smart, R. L., Smith, C., Solano, E., Solitro, F., Sordo, R., Soria Nieto, S., Souchay, J., Spagna, A., Spoto, F., Stampa, U., Steele, I. A., Steidelmüller, H., Stephenson, C. A., Stoev, H., Suess, F. F., Süveges, M., Surdej, J., Szabados, L., Szegedi-Elek, E., Tapiador, D., Taris, F., Tauran, G., Taylor, M. B., Teixeira, R., Terrett, D., Tingley, B., Trager, S. C., Turon, C., Ulla, A., Utrilla, E., Valentini, G., van Elteren, A., Van Hemelryck, E., van Leeuwen, M., Varadi, M., Vecchiato, A., Veljanoski, J., Via, T., Vicente, D., Vogt, S., Voss, H., Votruba, V., Voutsinas, S., Walmsley, G., Weiler, M., Weingrill, K., Werner, D., Wevers, T., Whitehead, G., Wyrzykowski, Ł., Yoldas, A., Žerjal, M., Zucker, S., Zurbach, C., Zwitter, T., Alecu, A., Allen, M., Allende Prieto, C., Amorim, A., Anglada-Escudé, G., Arsenijevic, V., Azaz, S., Balm, P., Beck, M., Bernstein, H. H., Bigot, L., Bijaoui, A., Blasco, C., Bonfigli, M., Bono, G., Boudreault, S., Bressan, A., Brown, S., Brunet, P. M., Bunclark, P., Buonanno, R., Butkevich, A. G., Carret, C., Carrion, C., Chemin, L., Chéreau, F., Corcione, L., Darmigny, E., de Boer, K. S., de Teodoro, P., de Zeeuw, P. T., Delle Luche, C., Domingues, C. D., Dubath, P., Fodor, F., Frézouls, B., Fries, A., Fustes, D., Fyfe, D., Gallardo, E., Gallegos, J., Gardiol, D., Gebran, M., Gomboc, A., Gómez, A., Grux, E., Gueguen, A., Heyrovsky, A., Hoar, J., Iannicola, G., Isasi Parache, Y., Janotto, A. M., Joliet, E., Jonckheere, A., Keil, R., Kim, D. W., Klagyivik, P., Klar, J., Knude, J., Kochukhov, O., Kolka, I., Kos, J., Kutka, A., Lainey, V., LeBouquin, D., Liu, C., Loreggia, D., Makarov, V. V., Marseille, M. G., Martayan, C., Martinez-Rubi, O., Massart, B., Meynadier, F., Mignot, S., Munari, U., Nguyen, A. T., Nordlander, T., Ocvirk, P., O'Flaherty, K. S., Olias Sanz, A., Ortiz, P., Osorio, J., Oszkiewicz, D., Ouzounis, A., Palmer, M., Park, P., Pasquato, E., Peltzer, C., Peralta, J., Péturaud, F., Pieniluoma, T., Pigozzi, E., Poels, J., Prat, G., Prod'homme, T., Raison, F., Rebordao, J. M., Risquez, D., Rocca-Volmerange, B., Rosen, S., Ruiz-Fuertes, M. I., Russo, F., Sembay, S., Serraller Vizcaino, I., Short, A., Siebert, A., Silva, H., Sinachopoulos, D., Slezak, E., Soffel, M., Sosnowska, D., Straižys, V., ter Linden, M., Terrell, D., Theil, S., Tiede, C., Troisi, L., Tsalmantza, P., Tur, D., Vaccari, M., Vachier, F., Valles, P., Van Hamme, W., Veltz, L., Virtanen, J., Wallut, J. M., Wichmann, R., Wilkinson, M. I., Ziaeepour, H., & Zschocke, S. 2016, A&A, 595, A1, arXiv:1609.04153

Gandhi, S. & Madhusudhan, N. 2017, MNRAS, 472, 2334, arXiv:1706.02302

—. 2018, MNRAS, 474, 271, arXiv:1710.06433

Garhart, E., Deming, D., Mandell, A., Knutson, H. A., Wallack, N., Burrows, A., Fortney, J. J., Hood, C., Seay, C., Sing, D. K., Benneke, B., Fraine, J. D., Kataria, T., Lewis, N., Madhusudhan, N., McCullough, P., Stevenson, K. B., & Wakeford, H. 2020, AJ, 159, 137, arXiv:1901.07040

Gelman, A. & Rubin, D. B. 1992, Statistical Science, 7, 457

Gelman, A. & Rubin, D. B. 1992, Statistical Science, 7, 457

Gennaro, M. & et al. 2018, WFC3 Data Handbook v. 4.0

Gibson, N. P. 2014a, MNRAS, 445, 3401, arXiv:1409.5668

—. 2014b, MNRAS, 445, 3401, arXiv:1409.5668

Gillon, M., Triaud, A. H. M. J., Demory, B.-O., Jehin, E., Agol, E., Deck, K. M., Lederer, S. M., de Wit, J., Burdanov, A., Ingalls, J. G., Bolmont, E., Leconte, J., Raymond, S. N., Selsis, F., Turbet, M., Barkaoui, K., Burgasser, A., Burleigh, M. R., Carey, S. J., Chaushev, A., Copperwheat, C. M., Delrez, L., Fernandes, C. S., Holdsworth, D. L., Kotze, E. J., Van Grootel, V., Almleaky, Y., Benkhaldoun, Z., Magain, P., & Queloz, D. 2017, Nature, 542, 456, arXiv:1703.01424

Goodman, J. & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Goulding, N. T., Barnes, J. R., Pinfield, D. J., Kovács, G., Birkby, J., Hodgkin, S., Catalán, S., Sipőcz, B., Jones, H. R. A., Del Burgo, C., Jeffers, S. V., Nefs, S., Gálvez-Ortiz, M. C., & Martin, E. L. 2012, MNRAS, 427, 3358, arXiv:1211.5288

Goyal, J. M., Mayne, N., Sing, D. K., Drummond, B., Tremblin, P., Amundsen, D. S., Evans, T., Carter, A. L., Spake, J., Baraffe, I., Nikolov, N., Manners, J., Chabrier, G., & Hebrard, E. 2019, MNRAS, 486, 783

Griffith, C. A. 2014, Philosophical Transactions of the Royal Society of London Series A, 372, 20130086, arXiv:1312.3988

Guillot, T. 2010, A&A, 520, A27, arXiv:1006.4702

Hanel, R. A., Conrath, B. J., Kunde, V. G., Prabhakara, C., Revah, I., Salomonson, V. V., & Wolford, G. 1972, JGR, 77, 2629

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del R'ıo, J. F., Wiebe, M., Peterson, P., G'erard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., & Oliphant, T. E. 2020, Nature, 585, 357

Hartman, J. D., Bakos, G. Á., Béky, B., Torres, G., Latham, D. W., Csubry, Z., Penev, K., Shporer, A., Fulton, B. J., Buchhave, L. A., Johnson, J. A., Howard, A. W., Marcy, G. W., Fischer, D. A., Kovács, G., Noyes, R. W., Esquerdo, G. A., Everett, M., Szklenár, T., Quinn, S. N., Bieryla, A., Knox, R. P., Hinz, P., Sasselov, D. D., Fűrész, G., Stefanik, R. P., Lázár, J., Papp, I., & Sári, P. 2012, AJ, 144, 139, arXiv:1207.3344

Haynes, K., Mandell, A. M., Madhusudhan, N., Deming, D., & Knutson, H. 2015, The Astrophysical Journal, 806, 146

Haynes, K., Mandell, A. M., Madhusudhan, N., Deming, D., & Knutson, H. 2015, ApJ, 806, 146, arXiv:1505.01490

Hebb, L., Collier-Cameron, A., Triaud, A. H. M. J., Lister, T. A., Smalley, B., Maxted, P. F. L., Hellier, C., Anderson, D. R., Pollacco, D., Gillon, M., Queloz, D., West, R. G., Bentley, S., Enoch, B., Haswell, C. A., Horne, K., Mayor, M., Pepe, F., Segransan, D., Skillen, I., Udry, S., & Wheatley, P. J. 2010, ApJ, 708, 224, arXiv:1001.0403

Hellier, C., Anderson, D. R., Collier Cameron, A., Gillon, M., Hebb, L., Maxted, P. F. L., Queloz, D., Smalley, B., Triaud, A. H. M. J., West, R. G., Wilson, D. M., Bentley, S. J., Enoch, B., Horne, K., Irwin, J., Lister, T. A., Mayor, M., Parley, N., Pepe, F., Pollacco, D. L., Segransan, D., Udry, S., & Wheatley, P. J. 2009, Nature, 460, 1098

Heng, K. 2018, Research Notes of the American Astronomical Society, 2, 128, arXiv:1807.06102

Heng, K. & Lyons, J. R. 2016, ApJ, 817, 149, arXiv:1507.01944

Henry, G. W. 1999, PASP, 111, 845

Higson, E., Handley, W., Hobson, M., & Lasenby, A. 2019, Statistics and Computing, 29, 891, arXiv:1704.03459

Howe, A. R., Adams, F. C., & Meyer, M. R. 2020, ApJ, 894, 130

Hubbard, W. B., Burrows, A., & Lunine, J. I. 2002, ARA&A, 40, 103

Hubeny, I., Burrows, A., & Sudarsky, D. 2003, The Astrophysical Journal, 594, 1011

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, arXiv:1307.2083

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90

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, JQSRT, 109, 1136

Isaacson, H. & Fischer, D. 2010, ApJ, 725, 875, arXiv:1009.2301

Johnson, M. C., Cochran, W. D., Addison, B. C., Tinney, C. G., & Wright, D. J. 2017, AJ, 154, 137, arXiv:1708.01291

Jones, E., Oliphant, T., Peterson, P., et al. 2001–, SciPy: Open source scientific tools for Python

Kammer, J. A., Knutson, H. A., Line, M. R., Fortney, J. J., Deming, D., Burrows, A., Cowan, N. B., Triaud, A. H. M. J., Agol, E., Desert, J.-M., Fulton, B. J., Howard, A. W., Laughlin, G. P., Lewis, N. K., Morley, C. V., Moses, J. I., Showman, A. P., & Todorov, K. O. 2015, The Astrophysical Journal, 810, 118

Kasting, J. F. 1988, Icarus, 74, 472

Kataria, T., Showman, A. P., Fortney, J. J., Stevenson, K. B., Line, M. R., Kreidberg, L., Bean, J. L., & Désert, J.-M. 2015, ApJ, 801, 86, arXiv:1410.2382

Kempton, E. M. R., Bean, J. L., Louie, D. R., Deming, D., Koll, D. D. B., Mansfield, M., Christiansen, J. L., López-Morales, M., Swain, M. R., Zellem, R. T., Ballard, S., Barclay, T., Barstow, J. K., Batalha, N. E., Beatty, T. G., Berta-Thompson, Z., Birkby, J., Buchhave, L. A., Charbonneau, D., Cowan, N. B., Crossfield, I., de Val-Borro, M., Doyon, R., Dragomir, D., Gaidos, E., Heng, K., Hu, R., Kane, S. R., Kreidberg, L., Mallonn, M., Morley, C. V., Narita, N., Nascimbeni, V., Pallé, E., Quintana, E. V., Rauscher, E., Seager, S., Shkolnik, E. L., Sing, D. K., Sozzetti, A., Stassun, K. G., Valenti, J. A., & von Essen, C. 2018, PASP, 130, 114401, arXiv:1805.03671

Kempton, E. M. R., Lupu, R., Owusu-Asare, A., Slough, P., & Cale, B. 2017, PASP, 129, 044402, arXiv:1611.03871

Kilpatrick, B. M., Cubillos, P. E., Stevenson, K. B., Lewis, N. K., Wakeford, H. R., MacDonald, R. J., Madhusudhan, N., Blecic, J., Bruno, G., Burrows, A., Deming, D., Heng, K., Line, M. R., Morley, C. V., Parmentier, V., Tucker, G. S., Valenti, J. A., Waldmann, I. P., Bean, J. L., Beichman, C., Fraine, J., Krick, J. E., Lothringer, J. D., & Mandell, A. M. 2018, AJ, 156, 103, arXiv:1704.07421

Kitzmann, D. & Heng, K. 2018, MNRAS, 475, 94, arXiv:1710.04946

Knutson, H. A., Howard, A. W., & Isaacson, H. 2010, The Astrophysical Journal, 720, 1569

Kopp, G., Lawrence, G., & Rottman, G. 2005, , 230, 129

Kopparapu, R. K., Ramirez, R., Kasting, J. F., Eymet, V., Robinson, T. D., Mahadevan, S., Terrien, R. C., Domagal-Goldman, S., Meadows, V., & Deshpande, R. 2013, ApJ, 765, 131, arXiv:1301.6674

Kostov, V. B., Schlieder, J. E., Barclay, T., Quintana, E. V., Colón, K. D., Brande, J., Collins, K. A., Feinstein, A. D., Hadden, S., Kane, S. R., Kreidberg, L., Kruse, E., Lam, C., Matthews, E., Montet, B. T., Pozuelos, F. J., Stassun, K. G., Winters, J. G., Ricker, G., Vanderspek, R., Latham, D., Seager, S., Winn, J., Jenkins, J. M., Afanasev, D., Armstrong, J. J. D., Arney, G., Boyd, P., Barentsen, G., Barkaoui, K., Batalha, N. E., Beichman, C., Bayliss, D., Burke, C., Burdanov, A., Cacciapuoti, L., Carson, A., Charbonneau, D., Christiansen, J., Ciardi, D., Clampin, M., Collins, K. I., Conti, D. M., Coughlin, J., Covone, G., Crossfield, I., Delrez, L., Domagal-Goldman, S., Dressing, C., Ducrot, E., Essack, Z., Everett, M. E., Fauchez, T., Foreman-Mackey, D., Gan, T., Gilbert, E., Gillon, M., Gonzales, E., Hamann, A., Hedges, C., Hocutt, H., Hoffman, K., Horch, E. P., Horne, K., Howell, S., Hynes, S., Ireland, M., Irwin, J. M., Isopi, G., Jensen, E. L. N., Jehin, E., Kaltenegger, L., Kielkopf, J. F., Kopparapu, R., Lewis, N., Lopez, E., Lissauer, J. J., Mann, A. W., Mallia, F., Mandell, A., Matson, R. A., Mazeh, T., Monsue, T., Moran, S. E., Moran, V., Morley, C. V., Morris, B., Muirhead, P., Mukai, K., Mullally, S., Mullally, F., Murray, C., Narita, N., Palle, E., Pidhorodetska, D., Quinn, D., Relles, H., Rinehart, S., Ritsko, M., Rodriguez, J. E., Rowden, P., Rowe, J. F., Sebastian, D., Sefako, R., Shahaf, S., Shporer, A., Tañón Reyes, N., Tenenbaum, P., Ting, E. B., Twicken, J. D., van Belle, G. T., Vega, L., Volosin, J., Walkowicz, L. M., & Youngblood, A. 2019, AJ, 158, 32, arXiv:1903.08017

Kreidberg, L. 2015, PASP, 127, 1161, arXiv:1507.08285

—. 2018, Exoplanet Atmosphere Measurements from Transmission Spectroscopy and Other Planet Star Combined Light Observations, 100

Kreidberg, L., Bean, J. L., Désert, J.-M., Benneke, B., Deming, D., Stevenson, K. B., Seager, S., Berta-Thompson, Z., Seifahrt, A., & Homeier, D. 2014a, Nature, 505, 69, arXiv:1401.0022

Kreidberg, L., Bean, J. L., Désert, J.-M., Line, M. R., Fortney, J. J., Madhusudhan, N., Stevenson, K. B., Showman, A. P., Charbonneau, D., McCullough, P. R., Seager, S., Burrows, A., Henry, G. W., Williamson, M., Kataria, T., & Homeier, D. 2014b, ApJL, 793, L27, arXiv:1410.2255

Kreidberg, L., Bean, J. L., Desert, J.-M., Line, M. R., Fortney, J. J., Madhusudhan, N., Stevenson, K. B., Showman, A. P., Charbonneau, D., McCullough, P. R., Seager, S., Burrows, A., Henry, G. W., Williamson, M., Kataria, T., & Homeier, D. 2014, The Astrophysical Journal, 793, L27

Kreidberg, L., Line, M. R., Bean, J. L., Stevenson, K. B., Désert, J.-M., Madhusudhan, N., Fortney, J. J., Barstow, J. K., Henry, G. W., Williamson, M. H., & Showman, A. P. 2015, ApJ, 814, 66, arXiv:1504.05586

Kuntschner, H., Bushouse, H., Kümmel, M., & Walsh, J. R. 2009, WFC3 SMOV proposal 11552: Calibration of the G141 grism, Space Telescope WFC Instrument Science Report

Lavvas, P., Koskinen, T., & Yelle, R. V. 2014, ApJ, 796, 15

Lee, J. M., Fletcher, L. N., & Irwin, P. G. J. 2012, MNRAS, 420, 170, arXiv:1110.2934

Lewis, N. K., Showman, A. P., Fortney, J. J., Knutson, H. A., & Marley, M. S. 2014, ApJ, 795, 150, arXiv:1409.5108

Lewis, N. K., Wakeford, H. R., MacDonald, R. J., Goyal, J. M., Sing, D. K., Barstow, J., Powell, D., Kataria, T., Mishra, I., Marley, M. S., Batalha, N. E., Moses, J. I., Gao, P., Wilson, T. J., Chubb, K. L., Mikal-Evans, T., Nikolov, N., Pirzkal, N., Spake, J. J., Stevenson, K. B., Valenti, J., & Zhang, X. 2020, ApJL, 902, L19

Li, C., Ingersoll, A., Bolton, S., Levin, S., Janssen, M., Atreya, S., Lunine, J., Steffes, P., Brown, S., Guillot, T., Allison, M., Arballo, J., Bellotti, A., Adumitroaie, V., Gulkis, S., Hodges, A., Li, L., Misra, S., Orton, G., Oyafuso, F., Santos-Costa, D., Waite, H., & Zhang, Z. 2020, Nature Astronomy, 4, 609, arXiv:2012.10305

Libby-Roberts, J. E., Berta-Thompson, Z. K., Diamond-Lowe, H., Gully-Santiago, M. A., Irwin, J. M., Kempton, E. M. R., Rackham, B. V., Charbonneau, D., Desert, J.-M., Dittmann, J. A., Hofmann, R., Morley, C. V., & Newton, E. R. 2021, arXiv e-prints, arXiv:2105.10487, arXiv:2105.10487

Line, M. R., Knutson, H., Wolf, A. S., & Yung, Y. L. 2014, ApJ, 783, 70, arXiv:1309.6663

Line, M. R. & Parmentier, V. 2016, ApJ, 820, 78, arXiv:1511.09443

Line, M. R., Wolf, A. S., Zhang, X., Knutson, H., Kammer, J. A., Ellison, E., Deroo, P., Crisp, D., & Yung, Y. L. 2013, ApJ, 775, 137, arXiv:1304.5561

Line, M. R., Zhang, X., Vasisht, G., Natraj, V., Chen, P., & Yung, Y. L. 2012, ApJ, 749, 93, arXiv:1111.2612

Lissauer, J. J. & Stevenson, D. J. Protostars and Planets V, ed. , B. ReipurthD. Jewitt & K. Keil, 591

Lothringer, J. D., Barman, T., & Koskinen, T. 2018, ArXiv e-prints, arXiv:1805.00038

MacDonald, R. J. & Madhusudhan, N. 2017, MNRAS, 469, 1979, arXiv:1701.01113

Madhusudhan, N. 2012, The Astrophysical Journal, 758, 36

Madhusudhan, N. 2012, ApJ, 758, 36, arXiv:1209.2412

—. 2018, Atmospheric Retrieval of Exoplanets, 104

—. 2019a, ARA&A, 57, 617, arXiv:1904.03190

—. 2019b, ARA&A, 57, 617, arXiv:1904.03190

Madhusudhan, N., Agúndez, M., Moses, J. I., & Hu, Y. 2016, Space Science Reviews, 205, 285, arXiv:1604.06092

Madhusudhan, N., Amin, M. A., & Kennedy, G. M. 2014a, ApJL, 794, L12, arXiv:1408.3668

—. 2014b, ApJL, 794, L12, arXiv:1408.3668

Madhusudhan, N., Bitsch, B., Johansen, A., & Eriksson, L. 2017, MNRAS, 469, 4102, arXiv:1611.03083

Madhusudhan, N., Crouzet, N., McCullough, P. R., Deming, D., & Hedges, C. 2014c, ApJL, 791, L9, arXiv:1407.6054

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, arXiv:1012.1603

Madhusudhan, N. & Seager, S. 2009, ApJ, 707, 24, arXiv:0910.1347

Madhusudhan, N. & Seager, S. 2009, The Astrophysical Journal, 707, 24

Madhusudhan, N. & Seager, S. 2010a, ApJ, 725, 261, arXiv:1010.4585

—. 2010b, ApJ, 725, 261, arXiv:1010.4585

Magic, Z., Chiavassa, A., Collet, R., & Asplund, M. 2015, A&A, 573, A90, arXiv:1403.3487

Mandal, S., Hegde, M., Samanta, T., Hazra, G., Banerjee, D., & Ravindra, B. 2017, A&A, 601, A106, arXiv:1608.04665

Mandel, K. & Agol, E. 2002, ApJL, 580, L171, arXiv:astro-ph/0210099

Mandel, K. & Agol, E. 2002, The Astrophysical Journal, 580, L171

Mandell, A. M., Haynes, K., Sinukoff, E., Madhusudhan, N., Burrows, A., & Deming, D. 2013, ApJ, 779, 128, arXiv:1310.2949

Mandell, A. M., Haynes, K., Sinukoff, E., Madhusudhan, N., Burrows, A., & Deming, D. 2013, The Astrophysical Journal, 779, 128

Mansfield, M., Bean, J. L., Line, M. R., Parmentier, V., Kreidberg, L., Désert, J.-M., Fortney, J. J., Stevenson, K. B., Arcangeli, J., & Dragomir, D. 2018, AJ, 156, 10, arXiv:1805.00424

Markwardt, C. B. Astronomical Society of the Pacific Conference Series, Vol. 411, , Astronomical Data Analysis Software and Systems XVIII, ed. D. A. BohlenderD. Durand & P. Dowler, 251

Maxted, P. F. L., Anderson, D. R., Doyle, A. P., Gillon, M., Harrington, J., Iro, N., Jehin, E., Lafreniere, D., Smalley, B., & Southworth, J. 2013, Monthly Notices of the Royal Astronomical Society, 428, 2645

McComas, D. J., Alexander, N., Allegrini, F., Bagenal, F., Beebe, C., Clark, G., Crary, F., Desai, M. I., De Los Santos, A., Demkee, D., Dickinson, J., Everett, D., Finley, T., Gribanova, A., Hill, R., Johnson, J., Kofoed, C., Loeffler, C., Louarn, P., Maple, M., Mills, W., Pollock, C., Reno, M., Rodriguez, B., Rouzaud, J., Santos-Costa, D., Valek, P., Weidner, S., Wilson, P., Wilson, R. J., & White, D. 2017, Space Science Reviews, 213, 547

McCullough, P. & MacKenty, J. 2012, Considerations for using Spatial Scans with WFC3, Space Telescope WFC Instrument Science Report

McKemmish, L. K., Yurchenko, S. N., & Tennyson, J. 2016, MNRAS, 463, 771, arXiv:1609.06120

Menou, K. 2012, ApJL, 754, L9, arXiv:1206.3363

Misener, W. & Schlichting, H. E. 2021, MNRAS, 503, 5658, arXiv:2103.09212

Mollière, P., van Boekel, R., Dullemond, C., Henning, T., & Mordasini, C. 2015, ApJ, 813, 47, arXiv:1509.07523

Moran, S. E., Hörst, S. M., Batalha, N. E., Lewis, N. K., & Wakeford, H. R. 2018, AJ, 156, 252, arXiv:1810.05210

Mordasini, C. 2014, A&A, 572, A118, arXiv:1406.4127

Mordasini, C., van Boekel, R., Mollière, P., Henning, T., & Benneke, B. 2016, ApJ, 832, 41, arXiv:1609.03019

Morley, C. V., Knutson, H., Line, M., Fortney, J. J., Thorngren, D., Marley, M. S., Teal, D., & Lupu, R. 2017, AJ, 153, 86, arXiv:1610.07632

Moses, J. I., Madhusudhan, N., Visscher, C., & Freedman, R. S. 2013, The Astrophysical Journal, 763, 25

Mugnai, L. V., Modirrousta-Galian, D., Edwards, B., Changeat, Q., Bouwman, J., Morello, G., Al-Refaie, A., Baeyens, R., Bieger, M. F., Blain, D., Gressier, A., Guilluy, G., Jaziri, Y., Kiefer, F., Morvan, M., Pluriel, W., Poveda, M., Skaf, N., Whiteford, N., Wright, S., Yip, K. H., Zingales, T., Charnay, B., Drossart, P., Leconte, J., Venot, O., Waldmann, I., & Beaulieu, J.-P. 2021, AJ, 161, 284, arXiv:2104.01873

National Academies of Sciences, E. & Medicine. 2018, Exoplanet Science Strategy (Washington, DC: The National Academies Press)

Newton, E. R., Irwin, J., Charbonneau, D., Berta-Thompson, Z. K., Dittmann, J. A., & West, A. A. 2016, ApJ, 821, 93, arXiv:1511.00957

Nikolov, N., Sing, D. K., Burrows, A. S., Fortney, J. J., Henry, G. W., Pont, F., Ballester, G. E., Aigrain, S., Wilson, P. A., Huitson, C. M., Gibson, N. P., Désert, J. M., Lecavelier Des Etangs, A., Showman, A. P., Vidal-Madjar, A., Wakeford, H. R., & Zahnle, K. 2015, MNRAS, 447, 463, arXiv:1411.4567

Nikolov, N., Sing, D. K., Goyal, J., Henry, G. W., Wakeford, H. R., Evans, T. M., López-Morales, M., García Muñoz, A., Ben-Jaffel, L., Sanz-Forcada, J., Ballester, G. E., Kataria, T., Barstow, J. K., Bourrier, V., Buchhave, L. A., Cohen, O., Deming, D., Ehrenreich, D., Knutson, H., Lavvas, P., Lecavelier des Etangs, A., Lewis, N. K., Mandell, A. M., & Williamson, M. H. 2018, MNRAS, 474, 1705, arXiv:1711.00859

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, arXiv:1310.0083

Noyes, R. W., Hartmann, L. W., Baliunas, S. L., Duncan, D. K., & Vaughan, A. H. 1984, ApJ, 279, 763

Nymeyer, S., Harrington, J., Hardy, R. A., Stevenson, K. B., Campo, C. J., Madhusudhan, N., Collier-Cameron, A., Loredo, T. J., Blecic, J., Bowman, W. C., Britt, C. B. T., Cubillos, P., Hellier, C., Gillon, M., Maxted, P. F. L., Hebb, L., Wheatley, P. J., Pollacco, D., & Anderson, D. R. 2011, ApJ, 742, 35, arXiv:1005.1017

Öberg, K. I. & Bergin, E. A. 2016, ApJL, 831, L19, arXiv:1610.07859

Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, ApJL, 743, L16, arXiv:1110.5567

Owen, T., Mahaffy, P., Niemann, H. B., Atreya, S., Donahue, T., Bar-Nun, A., & de Pater, I. 1999, Nature, 402, 269

Parmentier, V., Fortney, J. J., Showman, A. P., Morley, C., & Marley, M. S. 2016, ApJ, 828, 22, arXiv:1602.03088

Parmentier, V., Line, M. R., Bean, J. L., Mansfield, M., Kreidberg, L., Lupu, R., Visscher, C., Desert, J.-M., Fortney, J. J., Deleuil, M., Arcangeli, J., Showman, A. P., & Marley, M. S. 2018, ArXiv e-prints, arXiv:1805.00096

Parmentier, V., Showman, A. P., & Lian, Y. 2013, Astronomy and Astrophysics, 558, A91

Parviainen, H. Bayesian Methods for Exoplanet Science, ed. , H. J. DeegJ. A. Belmonte, 149

Patrascu, A. T., Yurchenko, S. N., & Tennyson, J. 2015, MNRAS, 449, 3613, arXiv:1504.02938

Pidhorodetska, D., Moran, S. E., Schwieterman, E. W., Barclay, T., Fauchez, T. J., Lewis, N. K., Quintana, E. V., Villanueva, G. L., Domagal-Goldman, S. D., Schlieder, J. E., Gilbert, E. A., Kane, S. R., & Kostov, V. B. 2021, arXiv e-prints, arXiv:2106.00685, arXiv:2106.00685

Pinhas, A., Madhusudhan, N., & Clarke, C. 2016, MNRAS, 463, 4516, arXiv:1609.02143

Pinhas, A., Madhusudhan, N., Gandhi, S., & MacDonald, R. 2019, MNRAS, 482, 1485, arXiv:1811.00011

Pinhas, A., Rackham, B. V., Madhusudhan, N., & Apai, D. 2018, MNRAS, 480, 5314, arXiv:1808.10017

Pollack, J. B., Hubickyj, O., Bodenheimer, P., Lissauer, J. J., Podolak, M., & Greenzweig, Y. 1996, Icarus, 124, 62

Pont, F., Sing, D. K., Gibson, N. P., Aigrain, S., Henry, G., & Husnoo, N. 2013, MNRAS, 432, 2917, arXiv:1210.4163

Pont, F., Zucker, S., & Queloz, D. 2006, MNRAS, 373, 231, arXiv:astro-ph/0608597

Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, ApJ, 853, 122, arXiv:1711.05691

—. 2019, AJ, 157, 96, arXiv:1812.06184

Richard, C., Gordon, I. E., Rothman, L. S., Abel, M., Frommhold, L., Gustafsson, M., Hartmann, J.-M., Hermans, C., Lafferty, W. J., Orton, G. S., Smith, K. M., & Tran, H. 2012, JQSRT, 113, 1276

Robinson, T. D. & Catling, D. C. 2014, Nature Geoscience, 7, 12, arXiv:1312.6859

Rodgers, C. D. 2000, Inverse Methods for Atmospheric Sounding: Theory and Practice

Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Boudon, V., Brown, L. R., Campargue, A., Champion, J. P., Chance, K., Coudert, L. H., Dana, V., Devi, V. M., Fally, S., Flaud, J. M., Gamache, R. R., Goldman, A., Jacquemart, D., Kleiner, I., Lacome, N., Lafferty, W. J., Mandin, J. Y., Massie, S. T., Mikhailenko, S. N., Miller, C. E., Moazzen-Ahmadi, N., Naumenko, O. V., Nikitin, A. V., Orphal, J., Perevalov, V. I., Perrin, A., Predoi-Cross, A., Rinsland, C. P., Rotger, M., Šimečková, M., Smith, M. A. H., Sung, K., Tashkun, S. A., Tennyson, J., Toth, R. A., Vandaele, A. C., & Vander Auwera, J. 2009, JQSRT, 110, 533

Rothman, L. S., Gordon, I. E., Barber, R. J., Dothe, H., Gamache, R. R., Goldman, A., Perevalov, V. I., Tashkun, S. A., & Tennyson, J. 2010, JQSRT, 111, 2139

Rothman, L. S., Gordon, I. E., Barber, R. J., Dothe, H., Gamache, R. R., Goldman, A., Perevalov, V. I., Tashkun, S. A., & Tennyson, J. 2010, Journal of Quantitative Spectroscopy and Radiative Transfer, 111, 2139

Safronov, V. S. & Zvjagina, E. V. 1969, Icarus, 10, 109

Schwarz, G. 1978, Annals of Statistics, 6, 461

Schwenke, D. W. 1998, Faraday Discussions, 109, 321

Seager, S. & Sasselov, D. D. 2000, ApJ, 537, 916, arXiv:astro-ph/9912241

Sedaghati, E., Boffin, H. M. J., MacDonald, R. J., Gandhi, S., Madhusudhan, N., Gibson, N. P., Oshagh, M., Claret, A., & Rauer, H. 2017, Nature, 549, 238, arXiv:1709.04118

Sedaghati, E., MacDonald, R. J., Casasayas-Barris, N., Hoeijmakers, H. J., Boffin, H. M. J., Rodler, F., Brahm, R., Jones, M., Sánchez-López, A., Carleo, I., Figueira, P., Mehner, A., & López-Puertas, M. 2021, MNRAS, 505, 435, arXiv:2103.12858

Shapiro, S. S. & Wilk, M. B. 1965, Biometrika, 52, 591

Sheppard, K. B., Mandell, A. M., Tamburo, P., Gand hi, S., Pinhas, A., Madhusudhan, N., & Deming, D. 2017, ApJL, 850, L32, arXiv:1711.10491

Sheppard, K. B., Welbanks, L., Mandell, A. M., Madhusudhan, N., Nikolov, N., Deming, D., Henry, G. W., Williamson, M. H., Sing, D. K., López-Morales, M., Ih, J., Sanz-Forcada, J., Lavvas, P., Ballester, G. E., Evans, T. M., García Muñoz, A., & dos Santos, L. A. 2021, AJ, 161, 51, arXiv:2010.09659

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. 2011, A&A, 527, A73, arXiv:1008.4795

Sing, D. K., Fortney, J. J., Nikolov, N., Wakeford, H. R., Kataria, T., Evans, T. M., Aigrain, S., Ballester, G. E., Burrows, A. S., Deming, D., Désert, J.-M., Gibson, N. P., Henry, G. W., Huitson, C. M., Knutson, H. A., Lecavelier Des Etangs, A., Pont, F., Showman, A. P., Vidal-Madjar, A., Williamson, M. H., & Wilson, P. A. 2016, Nature, 529, 59, arXiv:1512.04341

Sing, D. K., Wakeford, H. R., Showman, A. P., Nikolov, N., Fortney, J. J., Burrows, A. S., Ballester, G. E., Deming, D., Aigrain, S., Désert, J. M., Gibson, N. P., Henry, G. W., Knutson, H., Lecavelier des Etangs, A., Pont, F., Vidal-Madjar, A., Williamson, M. W., & Wilson, P. A. 2015, MNRAS, 446, 2428, arXiv:1410.7611

Skilling, J. American Institute of Physics Conference Series, Vol. 735, , American Institute of Physics Conference Series, ed. R. FischerR. Preuss & U. V. Toussaint, 395–405

Soubiran, C., Le Campion, J.-F., Brouillet, N., & Chemin, L. 2016, A&A, 591, A118, arXiv:1605.07384

Speagle, J. S. 2020, MNRAS, 493, 3132, arXiv:1904.02180

Spiegel, D. S., Silverio, K., & Burrows, A. 2009, The Astrophysical Journal, 699, 1487

Stassun, K. G. 2019, VizieR Online Data Catalog, IV/38

Stassun, K. G., Collins, K. A., & Gaudi, B. S. 2017, AJ, 153, 136, arXiv:1609.04389

Stassun, K. G., Oelkers, R. J., Paegert, M., Torres, G., Pepper, J., De Lee, N., Collins, K., Latham, D. W., Muirhead, P. S., Chittidi, J., Rojas-Ayala, B., Fleming, S. W., Rose, M. E., Tenenbaum, P., Ting, E. B., Kane, S. R., Barclay, T., Bean, J. L., Brassuer, C. E., Charbonneau, D., Ge, J., Lissauer, J. J., Mann, A. W., McLean, B., Mullally, S., Narita, N., Plavchan, P., Ricker, G. R., Sasselov, D., Seager, S., Sharma, S., Shiao, B., Sozzetti, A., Stello, D., Vanderspek, R., Wallace, G., & Winn, J. N. 2019, AJ, 158, 138, arXiv:1905.10694

Stevenson, K. B. 2016, ApJL, 817, L16, arXiv:1601.03492

Stevenson, K. B., Désert, J.-M., Line, M. R., Bean, J. L., Fortney, J. J., Showman, A. P., Kataria, T., Kreidberg, L., McCullough, P. R., Henry, G. W., Charbonneau, D., Burrows, A., Seager, S., Madhusudhan, N., Williamson, M. H., & Homeier, D. 2014, Science, 346, 838, arXiv:1410.2241

Swain, M. R., Estrela, R., Roudier, G. M., Sotin, C., Rimmer, P. B., Valio, A., West, R., Pearson, K., Huber-Feely, N., & Zellem, R. T. 2021, AJ, 161, 213, arXiv:2103.05657

Tamburo, P., Mandell, A. M., Deming, D., & Garhart, E. 2017, The Astronomical Journal, submitted

Tennyson, J. & Sutcliffe, B. T. 1982, , 77, 4061

Tennyson, J. & Yurchenko, S. 2018, Atoms, 6, 26, arXiv:1805.03711

Tennyson, J. & Yurchenko, S. N. 2012, MNRAS, 425, 21, arXiv:1204.0124

Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., Barton, E. J., Chubb, K. L., Coles, P. A., Diamantopoulou, S., Gorman, M. N., Hill, C., Lam, A. Z., Lodi, L., McKemmish, L. K., Na, Y., Owens, A., Polyansky, O. L., Rivlin, T., Sousa-Silva, C., Underwood, D. S., Yachmenev, A., & Zak, E. 2016, Journal of Molecular Spectroscopy, 327, 73, arXiv:1603.05890

Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., Barton, E. J., Chubb, K. L., Coles, P. A., Diamantopoulou, S., Gorman, M. N., Hill, C., Lam, A. Z., Lodi, L., McKemmish, L. K., Na, Y., Owens, A., Polyansky, O. L., Rivlin, T., Sousa-Silva, C., Underwood, D. S., Yachmenev, A., & Zak, E. 2016, Journal of Molecular Spectroscopy, 327, 73

Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., Clark, V. H. J., Chubb, K. L., Conway, E. K., Dewan, A., Gorman, M. N., Hill, C., Lynas-Gray, A. E., Mellor, T., McKemmish, L. K., Owens, A., Polyansky, O. L., Semenov, M., Somogyi, W., Tinetti, G., Upadhyay, A., Waldmann, I., Wang, Y., Wright, S., & Yurchenko, O. P. 2020, JQSRT, 255, 107228, arXiv:2007.13022

Terlouw, J. P. & Vogelaar, M. G. R. 2015, Kapteyn Package, version 2.3b6, Kapteyn Astronomical Institute, Groningen, available from `http://www.astro.rug.nl/software/kapteyn/`

The Pandas Development Team. 2020, pandas-dev/pandas: Pandas

Thorngren, D. & Fortney, J. J. 2019, ApJL, 874, L31, arXiv:1811.11859

Thorngren, D. P., Fortney, J. J., Murray-Clay, R. A., & Lopez, E. D. 2016, ApJ, 831, 64, arXiv:1511.07854

Tody, D. 1986, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 627, Instrumentation in astronomy VI, ed. D. L. Crawford, 733

Tody, D. Astronomical Society of the Pacific Conference Series, Vol. 52, , Astronomical Data Analysis Software and Systems II, ed. R. J. HanischR. J. V. Brissenden & J. Barnes, 173

Tregloan-Reed, J., Southworth, J., & Tappert, C. 2013, MNRAS, 428, 3671, arXiv:1211.0864

Tremblin, P., Amundsen, D. S., Chabrier, G., Baraffe, I., Drummond, B., Hinkley, S., Mourier, P., & Venot, O. 2016, ApJL, 817, L19, arXiv:1601.03652

Tremblin, P., Amundsen, D. S., Mourier, P., Baraffe, I., Chabrier, G., Drummond, B., Homeier, D., & Venot, O. 2015, ApJL, 804, L17, arXiv:1504.03334

Tremblin, P., Chabrier, G., Mayne, N. J., Amundsen, D. S., Baraffe, I., Debras, F., Drummond, B., Manners, J., & Fromang, S. 2017, ApJ, 841, 30, arXiv:1704.05440

Trotta, R. 2008, Contemporary Physics, 49, 71, arXiv:0803.4089

Tsiaras, A., Waldmann, I. P., Rocchetto, M., Varley, R., Morello, G., Damiano, M., & Tinetti, G. 2016, ApJ, 832, 202, arXiv:1511.07796

Tsiaras, A., Waldmann, I. P., Zingales, T., Rocchetto, M., Morello, G., Damiano, M., Karpouzas, K., Tinetti, G., McKemmish, L. K., Tennyson, J., & Yurchenko, S. N. 2018, AJ, 155, 156, arXiv:1704.05413

Tsuji, T., Ohnaka, K., Aoki, W., & Nakajima, T. 1996, A&A, 308, L29

Venot, O. & Agúndez, M. 2015, Experimental Astronomy, 40, 469, arXiv:1406.6566

von Zahn, U., Hunten, D. M., & Lehmacher, G. 1998, JGR, 103, 22815

Wakeford, H. R., Lewis, N. K., Fowler, J., Bruno, G., Wilson, T. J., Moran, S. E., Valenti, J., Batalha, N. E., Filippazzo, J., Bourrier, V., Hörst, S. M., Lederer, S. M., & de Wit, J. 2019, AJ, 157, 11, arXiv:1811.04877

Wakeford, H. R. & Sing, D. K. 2015, A&A, 573, A122, arXiv:1409.7594

Wakeford, H. R., Sing, D. K., Deming, D., Gibson, N. P., Fortney, J. J., Burrows, A. S., Ballester, G., Nikolov, N., Aigrain, S., Henry, G., Knutson, H., Lecavelier des Etangs, A., Pont, F., Showman, A. P., Vidal-Madjar, A., & Zahnle, K. 2013, MNRAS, 435, 3481, arXiv:1308.2106

Wakeford, H. R., Sing, D. K., Deming, D., Lewis, N. K., Goyal, J., Wilson, T. J., Barstow, J., Kataria, T., Drummond, B., Evans, T. M., Carter, A. L., Nikolov, N., Knutson, H. A., Ballester, G. E., & Mand ell, A. M. 2018, AJ, 155, 29, arXiv:1711.10529

Wakeford, H. R., Sing, D. K., Evans, T., Deming, D., & Mandell, A. 2016a, ApJ, 819, 10, arXiv:1601.02587

—. 2016b, ApJ, 819, 10, arXiv:1601.02587

Wakeford, H. R., Sing, D. K., Kataria, T., Deming, D., Nikolov, N., Lopez, E. D., Tremblin, P., Amundsen, D. S., Lewis, N. K., Mandell, A. M., Fortney, J. J., Knutson, H., Benneke, B., & Evans, T. M. 2017, Science, 356, 628, arXiv:1705.04354

Wakeford, H. R., Sing, D. K., Stevenson, K. B., Lewis, N. K., Pirzkal, N., Wilson, T. J., Goyal, J., Kataria, T., Mikal-Evans, T., Nikolov, N., & Spake, J. 2020, AJ, 159, 204, arXiv:2003.00536

Waldmann, I. P., Rocchetto, M., Tinetti, G., Barton, E. J., Yurchenko, S. N., & Tennyson, J. 2015a, ApJ, 813, 13, arXiv:1508.07591

Waldmann, I. P., Tinetti, G., Rocchetto, M., Barton, E. J., Yurchenko, S. N., & Tennyson, J. 2015b, ApJ, 802, 107, arXiv:1409.2312

Watson, A. J., Donahue, T. M., & Walker, J. C. G. 1981, Icarus, 48, 150

Welbanks, L. & Madhusudhan, N. 2019, AJ, 157, 206, arXiv:1904.05356

Welbanks, L., Madhusudhan, N., Allard, N. F., Hubeny, I., Spiegelman, F., & Leininger, T. 2019, ApJL, 887, L20, arXiv:1912.04904

Wetherill, G. W. & Stewart, G. R. 1993, Icarus, 106, 190

Wilkins, A. N., Deming, D., Madhusudhan, N., Burrows, A., Knutson, H., McCullough, P., & Ranjan, S. 2014, ApJ, 783, 113, arXiv:1401.4464

Winn, J. N., Fabrycky, D., Albrecht, S., & Johnson, J. A. 2010, ApJL, 718, L145, arXiv:1006.4161

Woitke, P., Helling, C., Hunter, G. H., Millard, J. D., Turner, G. E., Worters, M., Blecic, J., & Stock, J. W. 2018, A&A, 614, A1, arXiv:1712.01010

Wöllert, M. & Brandner, W. 2015, A&A, 579, A129, arXiv:1506.05456

Wong, I., Knutson, H. A., Kataria, T., Lewis, N. K., Burrows, A., Fortney, J. J., Schwartz, J., Shporer, A., Agol, E., Cowan, N. B., Deming, D., Désert, J.-M., Fulton, B. J., Howard, A. W., Langton, J., Laughlin, G., Showman, A. P., & Todorov, K. 2016, ApJ, 823, 122

Wright, N. J., Drake, J. J., Mamajek, E. E., & Henry, G. W. 2011, ApJ, 743, 48, arXiv:1109.4634

Yurchenko, S. N., Barber, R. J., & Tennyson, J. 2011, MNRAS, 413, 1828, arXiv:1011.1569

Yurchenko, S. N. & Tennyson, J. 2014, MNRAS, 440, 1649, arXiv:1401.4852

Zahnle, K., Marley, M. S., Freedman, R. S., Lodders, K., & Fortney, J. J. 2009a, ApJL, 701, L20, arXiv:0903.1663

—. 2009b, ApJL, 701, L20, arXiv:0903.1663

Zhang, M., Chachan, Y., Kempton, E. M. R., & Knutson, H. A. 2019, PASP, 131, 034501, arXiv:1811.11761

Zhang, M., Chachan, Y., Kempton, E. M. R., Knutson, H. A., & Chang, W. H. 2020a, ApJ, 899, 27, arXiv:2004.09513

—. 2020b, ApJ, 899, 27, arXiv:2004.09513

Zhou, Y., Apai, D., Lew, B. W. P., & Schneider, G. 2017, AJ, 153, 243, arXiv:1703.01301