

ACCESS: Ground-based Optical Transmission Spectroscopy of the Hot Jupiter WASP-4b

ALEX BIXEL,^{1,2} BENJAMIN V. RACKHAM,^{1,2} DÁNIEL APAI,^{1,2,3} NÉSTOR ESPINOZA,^{4,5,6} MERCEDES LÓPEZ-MORALES,^{7,2}
DAVID OSIP,⁸ ANDRÉS JORDÁN,^{5,6,4,2} CHIMA MCGRUDER,⁷ AND IAN WEAVER⁷

¹*Department of Astronomy/Steward Observatory, The University of Arizona, 933 N. Cherry Avenue, Tucson, AZ 85721, USA*

²*Earths in Other Solar Systems Team, NASA Nexus for Exoplanet System Science*

³*Lunar and Planetary Laboratory, The University of Arizona, 1640 E. University Blvd, AZ 85721, USA*

⁴*Max-Planck-Institut für Astronomie, Königstuhl 17, 69117, Heidelberg, Germany*

⁵*Instituto de Astrofísica, Facultad de Física, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 782-0436 Macul, Santiago, Chile*

⁶*Millennium Institute of Astrophysics, Av. Vicuña Mackenna 4860, 782-0436 Macul, Santiago, Chile*

⁷*Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 01238, USA*

⁸*Las Campanas Observatory, Carnegie Institution of Washington, Colina el Pino, Casilla 601, La Serena, Chile*

ABSTRACT

We present an optical transmission spectrum of the atmosphere of WASP-4b obtained through observations of four transits with *Magellan*/IMACS. Using a Bayesian approach to atmospheric retrieval, we find that the data are best described by a uniform opacity atmosphere, but the presence of potassium cannot be ruled out. Atmospheres with scattering hazes are weakly disfavored by the data. Our models include a component to model the transit light source effect (spectral contamination from unocculted spots on the stellar photosphere), which we show can have a marked impact on the observed transmission spectrum for reasonable spot covering fractions ($< 10\%$); this is the first such analysis for WASP-4b. We are also able to fit for the size and temperature contrast of spots observed during the second and third transits, finding evidence for both small, cool and large, warm spot-like features on the photosphere. Finally, we compare our results to those published by [Huitson et al. \(2017\)](#) using *Gemini*/GMOS, and find that our data are in agreement. Our combined results suggest that the atmosphere is largely opaque, but that there may exist a narrow region of atomic absorption above the optically thick layer.

Keywords: planets and satellites: atmospheres — planets and satellites: individual (WASP-4b) — stars: activity — stars: starspots — techniques: spectroscopic

1. INTRODUCTION

WASP-4b is a $1.4R_J$ hot Jupiter orbiting a G7V star with a period of 1.34 d, equilibrium temperature of ~ 1750 K, and transit depth $(R_p/R_s)^2 = 2.4\%$ ([Wilson et al. 2008](#)). It has been observed in transit over three dozen times, offering strong constraints on its orbit ([Hoyer et al. 2013](#)), the spot activity and relative rotation of its host star ([Sanchis-Ojeda et al. 2011](#)), and also placing upper limits on its transit timing variation amplitude ([Nikolov et al. 2012](#)). Secondary eclipse photometry with *Spitzer* reveals little evidence for a temperature inversion in the planet’s atmosphere ([Beer et al. 2011](#)).

Given the tight constraints on its transit depth and orbital properties, WASP-4b is a natural target for spectroscopic transit observations. Transit spectroscopy can be used to constrain the composition and structure of the planet’s up-

per atmosphere and test for the presence of clouds, scattering hazes, and atomic or molecular absorbers (e.g., [Seager & Sasselov 2000](#); [Brown 2001](#)). These observations also offer insight into the stellar photosphere, i.e., through the measurement of star spot temperatures within the transit chord (e.g., [Pont et al. 2008](#); [Sing et al. 2011](#); [Béky et al. 2014](#)). However, signals from the photosphere can be degenerate with those from the planet’s atmosphere, leading to contrasting interpretations of planetary and stellar origins for features in optical transmission spectra (e.g., [Pont et al. 2013](#); [McCullough et al. 2014](#)). It is therefore critical to demonstrate methods for accounting for this degeneracy as the field moves toward smaller targets and increasingly precise observations (e.g., [Rackham et al. 2018](#)).

Transit spectroscopy of WASP-4b has been attempted with *HST*/WFC3 ([Ranjan et al. 2014](#)), but was unsuccessful due to detector saturation. Recently, a *Gemini*/GMOS optical transmission spectrum of WASP-4b has been published by [Huit-](#)

son et al. (2017), who measure a nearly uniform opacity from 440–940 nm, suggesting the presence of high-altitude clouds and a possible sodium absorption feature.

As part of the Arizona-CfA-Católica-Carnegie Exoplanet Spectroscopy Survey (ACCESS, Rackham et al. 2017), we have observed four transits of WASP-4b with *Magellan*/IMACS. We have previously demonstrated the use of this instrument for transit spectroscopy using our custom data reduction pipeline (Jordán et al. 2013; Rackham et al. 2017; Espinoza et al. 2018). In this paper, we present an optical transmission spectrum from 450–900 nm, interpret our results using a Bayesian retrieval framework introduced in Espinoza et al. (2018) (submitted to MNRAS), and compare and contrast our findings with those of Huitson et al. (2017). We find that the data slightly favor a uniform opacity atmosphere, but the presence of potassium cannot be ruled out.

2. OBSERVATIONS

We conducted spectroscopic observations of WASP-4 on the nights of 24 September 2013, 17 October 2013, 14 August 2015, and 26 September 2015 (hereafter Transits 1–4) using the Inamori-Magellan Areal Camera & Spectrograph (IMACS, Dressler et al. 2011) on the 6.5m Magellan-Baade telescope at Las Campanas Observatory in Chile. We observed using multi-slit masks in the $f/2$ mode, and observations of HeNeAr and quartz calibration lamps before and after the observations allowed for wavelength calibration and flat-field correction. The key parameters of our observations are listed in Table 1, and we assess the data quality for each night in Section 3.7.

2.1. 24 September 2013, 17 October 2013, and 14 August 2015

On the first three nights we used a setup consisting of a 400–1000 nm spectroscopic filter, a 300 lines/mm grism with a blaze angle of 17.5 degrees, and 20" wide spectral slits for the target and reference stars. Twelve reference stars were observed, although only one was used in the final data reduction for reasons discussed in Section 3.2. Most of the stellar spectra were dispersed across two chips. While the lunar sky background was minimal on the first and third nights, it was substantial during the second, and we take this into account in our data reduction pipeline. Finally, the observations on 25 September 2013 did not commence until shortly after ingress, so we are only able to retrieve a partial transit for the night.

2.2. 26 September 2015

Our instrument setup for the fourth observation was modified from the previous three. We used a 570–980 nm order-blocking filter for the purpose of eliminating higher-order interference towards red wavelengths, 10" wide slits, and a

150 lines/mm grism with a blaze angle of 18.8 degrees. This setup allowed for more tightly dispersed spectra which fall on a single chip, thereby reducing detector-to-detector variations in the spectra and avoiding chip gaps. The moon was in full phase and separated from the target by 40 degrees, contributing significantly to the ambient sky background, which we again account for in our data reduction pipeline.

3. DATA REDUCTION & TRANSIT FITTING

3.1. Data reduction

We reduce the raw data using our custom Python-based pipeline which has been used for similar observations of WASP-6b (Jordán et al. 2013), GJ 1214b (Rackham et al. 2017), and WASP-19b (Espinoza et al. 2018); the detailed functions of the pipeline, including standard bias and flat calibration, bad pixel and cosmic ray correction, sky subtraction, spectrum extraction, and wavelength calibration are reviewed by Jordán et al. (2013) and Rackham et al. (2017). In our reduction, we choose not to apply a flat field correction; when testing reductions both with and without the correction, we find that it does not noticeably improve our results for most transits and introduces systematic errors into the light curve of the second night.

We have updated our sky-subtraction algorithm to better account for asymmetric background profiles due to scattered light within the instrument. We use quartz lamp flat-field exposures to characterize the shape of the scattered light profile in each slit using a 6th-order polynomial, then fit an amplitude for the scattered light profile alongside a uniform sky level at each position along the spectrum for every exposure.

The output of the pipeline is a set of reduced wavelength-binned and integrated (“white”) light curves for the target and reference stars.

3.2. Reference star selection

The shape of the target’s light curve is complicated by instrumental and atmospheric effects, such as changes in air-mass or transparency. To calibrate out these effects, we simultaneously observed 12 reference stars of comparable optical apparent magnitudes and color ratios using multi-slit masks. Of these, two reached the saturation limit of the detector and were not usable.

Working with the remaining 10 reference stars, we correct the target light curve by dividing by a set of reference star light curves, and our choice of which reference stars to use is primarily constrained by the shape of the out-of-transit baseline, which we expect to be flat or consistent with a polynomial of no more than second order in time. We consider also the level of noise which a star introduces into the light curve and the similarity of its spectrum to that of the target. By these measures, we find that a single reference star (2MASS J23341836-4204509) which occupies the same pair of detec-

Table 1. Instrument setup and transit model characteristics for each of our four observations. All observations made use of the IMACS f/2 camera with multi-slit masks.

Transit	Date (start of night)	Filter	Grism (l/mm)	R_p/R_s (white light)	Scatter / Photon noise
1	24 September 2013	Spectroscopic f/2	300	$0.1533^{+0.0013}_{-0.0011}$	3.0
2	17 October 2013	$0.1554^{+0.0008}_{-0.0009}$	5.2
3	14 August 2015	$0.1545^{+0.0003}_{-0.0003}$	1.7
4	26 September 2015	WBP 5694-9819	150	$0.1575^{+0.0007}_{-0.0008}$	4.9

tor chips as the target produces the optimal light curve in each bin; we note that this is the same reference star used by [Huitson et al. \(2017\)](#) in their analysis of this target’s transmission spectrum.

3.3. Stability of relative wavelength solutions

In our data reduction and spectral extraction pipeline, we fit for each stellar spectrum a 6th-order polynomial wavelength solution, which converts the pixel coordinates of the spectrum to wavelength units. The deviation between the wavelength solutions for our target and chosen reference star is no more than $\sim 5 \text{ \AA}$ for a given reference feature. A time-linear drift in the wavelength solution can be observed over the course of the night, and is calibrated using a set of reference lines in the stellar spectra, including $H\beta$, $H\alpha$, Na-D , and O_2 .

Magellan-Baade is equipped with an atmospheric dispersion compensator (ADC), which corrects for the effects of differential atmospheric refraction as the airmass of our target changes over the course of the night. Differential refraction would have the effect of introducing a non-linear time dependence into the stellar wavelength solutions. We test the accuracy of the ADC correction by measuring the distance between Na and $H\alpha$ in the target and reference spectra as a function of time, which we calculate to vary by no more than 1 \AA over the course of the night. Furthermore, the difference between the measured distance for the target and reference stellar spectra changes by no more than 0.5 \AA , and less than 0.1 \AA on nights with higher quality data. Since the long-term change in this measurement is much lower than the frame-to-frame scatter (up to 5 \AA), we reason that any effect on the measured transit depth will be negligible.

3.4. Fitting transit models

We model our integrated and spectroscopic light curves using the analytic models introduced by [Mandel & Agol \(2002\)](#), and marginalize over the parameter space with a Markov-Chain Monte Carlo (MCMC) algorithm previously detailed in [Rackham et al. \(2017\)](#). As in previous works, we fit for the limb-darkening coefficients in order to account for any biases that might arise due to our imperfect knowledge of the intensity profile of the photosphere ([Espinoza](#)

& [Jordán 2015](#)). As WASP-4b has been studied extensively through photometric observations, we hold the orbital parameters fixed to the mean values in Table 2.

The fitted parameters include the planet-to-star radius ratio (R_p/R_s), two coefficients for the first-order polynomial used to fit the baseline trend (see Section 3.4.1), two parameters for a quadratic limb-darkening law which are sampled according to the method detailed in [Kipping \(2013\)](#), and noise parameters corresponding to uncorrelated (“white”) and correlated (“red”) noise in the light curve. The mid-transit time is fitted using the white light curve, then fixed for the individual bins. Broad, uniform prior distributions are used for all of the fitted parameters. The likelihood parameter is calculated through the wavelet method detailed in [Carter & Winn \(2009\)](#) and implemented in Python¹, and each bin is fitted independently of the rest.

The fitted white light curves for each night are shown in Figure 1; note that two of the transits exhibit spot-crossing features which have been excluded from the fit (see Section 4.2). The binned light curves are presented in Figures A1 through A4.

The primary output of our fitting routine is a transmission spectrum for 19 independently fitted continuum bins, each approximately 20 nm in width; this bin width was chosen to be as small as possible while maintaining signal-to-noise ratios of a few hundred in each bin. We also fit three narrower bins centered on possible absorption features, including Na D , $H\alpha$, and the K I doublet (767/770 nm). The overlaid and combined transmission spectra are presented in Figure 2, covering wavelengths from 450–900 nm in the first three nights, and 570–900 nm for the fourth night. We have no data from 790–805 nm and 835–850 nm as these correspond to the location of the detector chip gaps in the target and reference star spectra. Additionally, the bluest bin of the second transit is found to deviate by $> 3\sigma$ from the neighboring values in a manner which is highly dependent on the choice of systematics model; this bin is excluded from our results.

¹ <https://github.com/nespinoza/flicker-noise>

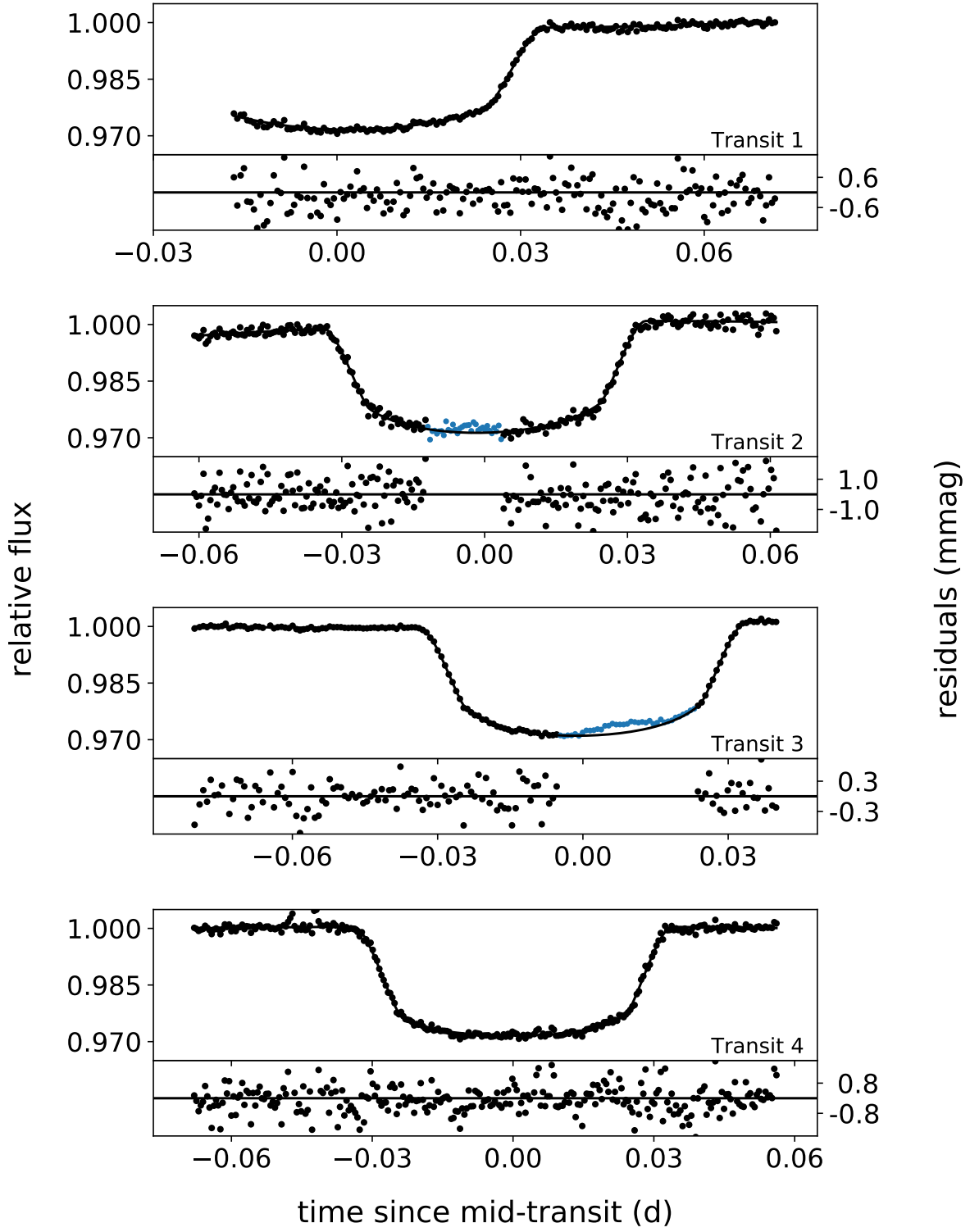


Figure 1. White light curves with transit models and residuals (1σ values marked). Two of the light curves featured spot-crossing events (blue) which are not included in the fit. The long-term trend is modeled by a second-order polynomial.

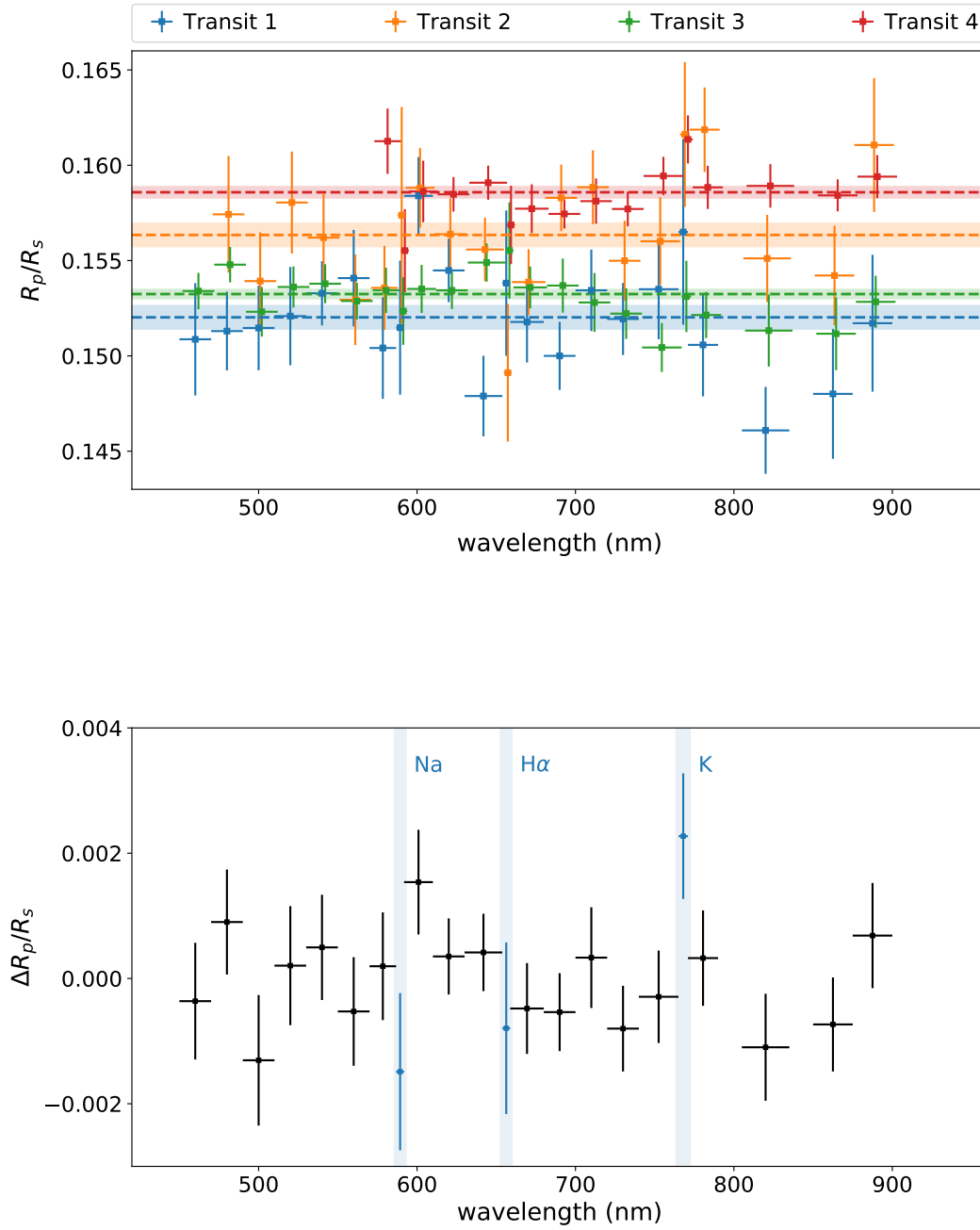


Figure 2. (Top) Transmission spectra from each night, with no corrections or offsets applied. The dashed lines and shaded regions represent weighted average values with $\pm 1\sigma$ uncertainties. (Bottom) Combined transmission spectrum from four nights, with the weighted average value subtracted from each night. Narrow bins centered on potential atomic features are highlighted, and the remaining (continuum) bins are black. The spectrum of Transit 3 is corrected for the effect of an occulted spot as described in Section 3.5 before it is incorporated into the combined result.

Table 2. Relevant previously measured properties of WASP-4 and its companion, with 1σ uncertainties.

Parameter	Value	Reference
WASP-4		
$R_s (R_\odot)$	$0.873^{+0.036}_{-0.027}$	Gillon et al. (2009)
[Fe/H]	$-0.03^{+0.09}_{-0.09}$...
$\log(g)$ (cgs)	$4.487^{+0.019}_{-0.015}$...
T_s (K)	5540^{+55}_{-55}	Maxted et al. (2011)
WASP-4b		
P (d)	1.33823204	Hoyer et al. (2013)
R_p/R_s	$0.15445^{+0.00025}_{-0.00025}$...
i (deg)	$88.52^{+0.39}_{-0.26}$...
a/R_s	$5.463^{+0.025}_{-0.020}$...
e	≈ 0	Beer et al. (2011)
$R_p (R_J)$	$1.395^{+0.022}_{-0.022}$	Hoyer et al. (2013)
$M_p (M_J)$	$1.237^{+0.021}_{-0.021}$	Winn et al. (2009)

3.4.1. Systematics model & common mode correction

We find that the trends which remain in the target light curves after dividing by the reference star light curves are adequately described by a second-order polynomial in time. However, any effort to fit the second-order trend independently in each bin risks degeneracy between the transit depth and the systematics model. To circumvent this degeneracy, we fit a transit model alongside a quadratic polynomial in time to our white light curves, calculate the residuals, then employ a common mode correction by scaling and subtracting the residuals from each binned light curve (e.g., Sing et al. 2013; Nikolov et al. 2015; Huitson et al. 2017). We then fit the corrected binned light curves with a linear systematics model, which is not as degenerate with transit depth.

An additional parameter of our systematics model is the scale of the red noise which we use to calculate the likelihood parameter following Carter & Winn (2009). This consideration results in a $\sim 15\%$ increase in our uncertainties, which may be unnecessary if our noise is uncorrelated. While the ratios in Table 1 suggest a large amount of systematic noise, this does not necessarily imply *correlated* noise; the systematic noise may still be captured by the white noise parameter. Nevertheless, we conservatively choose to apply the red noise correction.

3.5. Correcting for occulted spots

The light curves of Transits 2 and 3 feature prominent spot occultation features which we exclude from the light curve model fit. While doing so allows us to fit the apparent transit depth without simultaneously modeling the spot, it does *not* eliminate the impact of the occulted spot on the result-

ing transit spectrum. Since the spot is cooler than the rest of the photosphere, it breaks the transit model's assumption that the photosphere is azimuthally symmetric and amplifies the fitted transit depth in a wavelength-dependent manner (e.g., Deming et al. 2013). In previous works we have referred to this as the Transit Light Source Effect (TLSE); here we derive an approximate correction (ignoring limb-darkening) for the TLSE due to a single spot.

Assuming the photosphere to have specific flux $F_{\lambda,\text{phot}}$ and radius R_s , the spot to have specific flux $F_{\lambda,\text{spot}}$ and radius R_{spot} , and the planet to have apparent radius R_p and true radius $R_{p,0}$, then the observed relative transit depth will be

$$\left(\frac{R_p}{R_s}\right)^2 = \frac{F_{\lambda,\text{phot}} R_{p,0}^2}{F_{\lambda,\text{phot}}(R_s^2 - R_{\text{spot}}^2) + F_{\lambda,\text{spot}} R_{\text{spot}}^2} \quad (1)$$

By re-arranging this equation, we derive a corrective factor for the transit depth²:

$$\epsilon_\lambda \equiv \left(\frac{R_p}{R_{p,0}}\right)^2 = \left[1 - \left(1 - \frac{F_{\lambda,\text{spot}}}{F_{\lambda,\text{phot}}}\right) \left(\frac{R_{\text{spot}}}{R_s}\right)^2\right]^{-1} \quad (2)$$

Later, in Section 4.2, we find the feature in Transit 3 to be best described by a model of a single spot with radius $R_{\text{spot}}/R_s = 0.27^{+0.03}_{-0.02}$ and a temperature contrast of $T_s - T_{\text{spot}} = 100 \pm 5$ K. While this feature could in principle be due to a mixture of occulted spots and faculae, our data are insufficient to constrain more complex models, so we use the parameters of the single spot model for the purposes of correcting for the TLSE in this spectrum. We interpolate a PHOENIX (Husser et al. 2013) stellar photosphere model grid onto the values in Table 2 to compute $F_{\lambda,\text{phot}}$ and $F_{\lambda,\text{spot}}$ at their respective temperatures, then calculate ϵ_λ from Equation 2.

In Figure 3 we plot $\epsilon^{1/2}$, which is the corrective factor for the planet *radius*, as well as the original and corrected spectra for Transit 3. This factor is given in Table B1, binned to match our final spectrum. This corrective factor is not applied to the spectrum of Transit 3 in the top panel of Figure 2, but is applied before creating the combined spectrum in the bottom panel. We do not derive a similar correction for Transit 2 because (i) it is unclear (see Section 4.2) whether this feature is best modeled by a small or large spot and (ii) the data for this night are of poor quality and thus have a comparatively low weight in the combined spectrum.

Finally, it is worth noting that this effect is just as prominent for spots outside of the transit chord. Equation 2 can be further generalized by replacing $(R_{\text{spot}}/R_s)^2$ with F_{het} , the areal spot covering fraction of the photosphere, and using an average spot temperature to compute $F_{\lambda,\text{spot}}$

² See also Equation 2 of Rackham et al. (2018).

3.6. Combining data from separate nights

The dashed lines in Figure 2 illustrate the weighted average value of R_p/R_s from each night; not all of the nights agree to within 2σ . This is indicative of systematic variation in the fitted transit depth between the four nights. We consider the possible causes of this variation:

1. Variations in the spot-covering fraction of the star due to magnetic activity would cause the baseline brightness of the star to change, while the amount of light blocked by the planet remains the same. The variation in R_p/R_s between each night is as large as $\sim 2\%$; however, the 38 broad-band transits observed by Hoyer et al. (2013) do not reveal such a large variation in R_p/R_s , so stellar activity alone likely does not account for the inconsistency.
2. The application of a common-mode correction results in a loss of information about the absolute transit depth of the target, so we might not expect our observations across each night to agree.
3. Transits 1 and 3 demonstrate $< 2\sigma$ agreement between their mean R_p/R_s values. Meanwhile, transits 2 and 4 suffered from systematic sources of noise as demonstrated by the large scatter-to-photon noise ratios for these nights in Table 1. This systematic noise may have contributed to the elevated transit depths for the two nights.

Regardless of the cause of these variations, we are primarily interested in the relative transit depths; therefore, we subtract the value of R_p/R_s denoted by the dashed lines from the individual spectra then calculate the average values of $\Delta R_p/R_s$ in each bin, weighted by the uncertainties for each night. The combined transmission spectrum is presented in the bottom panel of Figure 2.

3.7. Assessment of data quality

Several of our nights suffered from sub-optimal data quality, as demonstrated by the large values of scatter-to-photon noise in Table 1 and the comparatively large uncertainties in the transit depths for Transits 1 and 2. Here, we will discuss the key factors governing the data quality for each night.

- **Transit 1:** Observations for this night did not commence until shortly after ingress; this introduces significant uncertainty in the determination of the baseline flux. Furthermore, the observing strategy was optimized to minimize readout noise and avoid non-linearity, so we opted for the slowest readout mode and integration times of ~ 15 s. As a result, only 20% of the observing time was spent gathering light.

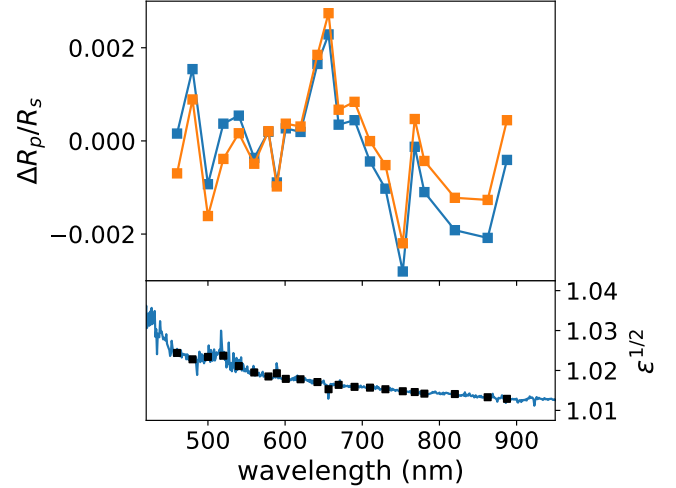


Figure 3. (Top) The spectrum of Transit 3 before (blue) and after (orange) correcting for the large occulted spot. Error bars are excluded for visibility and both spectra have been mean-subtracted. (Bottom) The radius contamination factor due to the spot, binned to 1 nm. The black squares are binned to match the spectrum above.

- **Transit 2:** On this night, the moon was in full phase and at an angle of 53° , resulting in a significant increase in sky brightness compared to the other nights. Similar to Transit 1, our exposure strategy only allowed us to utilize 19% of our time for integration. The resulting transmission spectrum has the lowest S/N of the three nights, and the largest ratio of scatter-to-photon noise.
- **Transit 3:** During this night, we decided to alter our observing strategy, opting instead for the fastest read-out mode and longer integration times. Nevertheless, we find no evidence for non-linearity in our spectra, and the read noise level remains well below the other sources of noise. This allowed us to integrate for 68% of our observing time. The light curves and spectrum from this night have the highest S/N of all of the nights. However, our lack of data after ~ 10 minutes following egress hampers our ability to accurately measure the baseline brightness.
- **Transit 4:** For this observation, we cut a new mask and experimented with a narrower slit width. We later found that narrow slits complicate the spectral extraction procedure as width of the spectral profile is only barely contained within the slit. This introduced aliasing effects and a greater amount of uncertainty in the spectral extraction algorithm, resulting in a large scatter-to-photon noise ratio. During this observation, we utilized 47% of the observing time for integration.

Our observations of WASP-4b have yielded useful insights into the optimal use of IMACS for transit spectroscopy. We

offer the following advice for future observers with similar setups:

1. Always opt for the fastest readout mode; the read noise will be inconsequential for any achievable transit spectroscopy goal.
2. Narrower slits do not offer much benefit unless the science goal specifically requires high spectral resolution. Wide slits (as large as 10" or wider) allow for a better determination of the sky brightness and a more accurate fit to the spectral profile.
3. The IMACS detector can accommodate approximately 35,000 ADU before non-linearity becomes non-negligible at precisions required for transit spectroscopy, so exposure times should be set with this as the target maximum pixel value so as to minimize the total readout overhead.
4. A moderate defocus (up to $\sim 2''$, depending on the sky background) will allow for longer exposure times before non-linearity is reached, thereby increasing the efficiency of the observations. If using wide slits as advised above, this will come at no additional cost to spectral resolution.

4. ANALYSIS

4.1. Atmospheric retrieval

To interpret the combined transmission spectrum in Figure 2, we employ a Bayesian atmospheric retrieval code based on PyMultiNest (Buchner et al. 2014) which we originally introduced in Espinoza et al. (2018). Following Heng & Kitzmann (2017), our atmosphere model includes two regions: an optically thick base region (which could also be interpreted as a cloud layer) with radius $(R_p/R_s)_0$ and reference pressure P_0 , and an isothermal optically thin region above with temperature T . The components of the optically thin region may include either or both of (i) a set of atomic and molecular species and (ii) a scattering haze defined by a cross-section power law $\sigma = a\sigma_0 (\lambda/\lambda_0)^\gamma$, where $\sigma_0 = 5.31 \times 10^{-27} \text{ cm}^2$ is the Rayleigh scattering cross-section of H_2 at the reference wavelength $\lambda_0 = 350 \text{ nm}$ (MacDonald & Madhusudhan 2017). We constrain γ to be between 0 (uniform opacity) and -4 (Rayleigh scattering); this assumption allows for a better constraint on a .

As we have previously argued in Rackham et al. (2017, 2018), it is important that analyses of transmission spectra account for the Transit Light Source Effect (TLSE) which can be introduced by heterogeneous features (spots and faculae) on the stellar surface. We incorporate a three-parameter model for the stellar photosphere into our retrieval code, fitting this model simultaneously with that of the planet's atmosphere. Section 4.1.1 discusses this model in more detail.

The parameters and priors for each component of the model are listed in Table 3. In this paper, we present the results of three atmospheric models, each of which is combined with a photosphere model (+1 free parameter, F_{het}). Motivated by the $\sim 2\sigma$ offset of the K I narrow bin from the mean planet radius in Figure 2, we choose to test two models with potassium and one without. Our models are:

1. A uniform opacity model corresponding to a cloud layer at the top of the atmosphere. (1+1 free parameters)
2. Includes both a cloud layer at $(R_p/R_s)_0$ and potassium. (3+1 free parameters)
3. Includes a cloud layer, potassium, and a scattering haze. (5+1 free parameters)

For each fitted model, the code outputs a Bayesian log-evidence parameter, $\ln(Z)$; two different models can be compared through their logarithmic Bayes' factors, $\ln(B_m) = \ln(Z_2) - \ln(Z_1)$. In the remaining sections, $\ln(B_m)$ for each model is referenced with respect to the model for a uniform opacity atmosphere with no photosphere component (i.e. a flat line).

4.1.1. Stellar heterogeneity

Assessing the level of TLSE contamination is critical to correctly interpreting high-precision transmission spectra of transiting exoplanets (Apai et al. 2018). This is particularly important for investigating features that could have a planetary or stellar origin, such as the He 10,833 Å line (Spake et al. 2018) or the Na and K alkali lines. We have previously demonstrated this effect in the spectrum of GJ 1214b (Rackham et al. 2017), which has a featureless near-IR spectrum but a visible spectrum that is apparently offset below the near-IR transit depths. In this study, we found that for reasonable assumptions about the star's activity, the planetary transmission spectrum of GJ 1214b can be shown to be flat in the visible as well. We also detect the signature of stellar activity in one of our *Magellan*/IMACS transit data sets for WASP-19b (Espinoza et al. 2018).

WASP-4 is of slightly later spectral type than the Sun, and multiple spot-crossing events have been observed during previous transit observations of WASP-4b (Sanchis-Ojeda et al. 2011) as well as in two of our data sets. Furthermore, four years of photometric monitoring data from the ASAS-SN database (Shappee et al. 2014; Kochanek et al. 2017) suggest a moderate 1% V-band variability amplitude, which is mostly likely caused by transient heterogeneous surface features rotating in and out of view. For these reasons, we caution that any analysis of the transmission spectrum of WASP-4b which ignores stellar contamination may be overly optimistic.

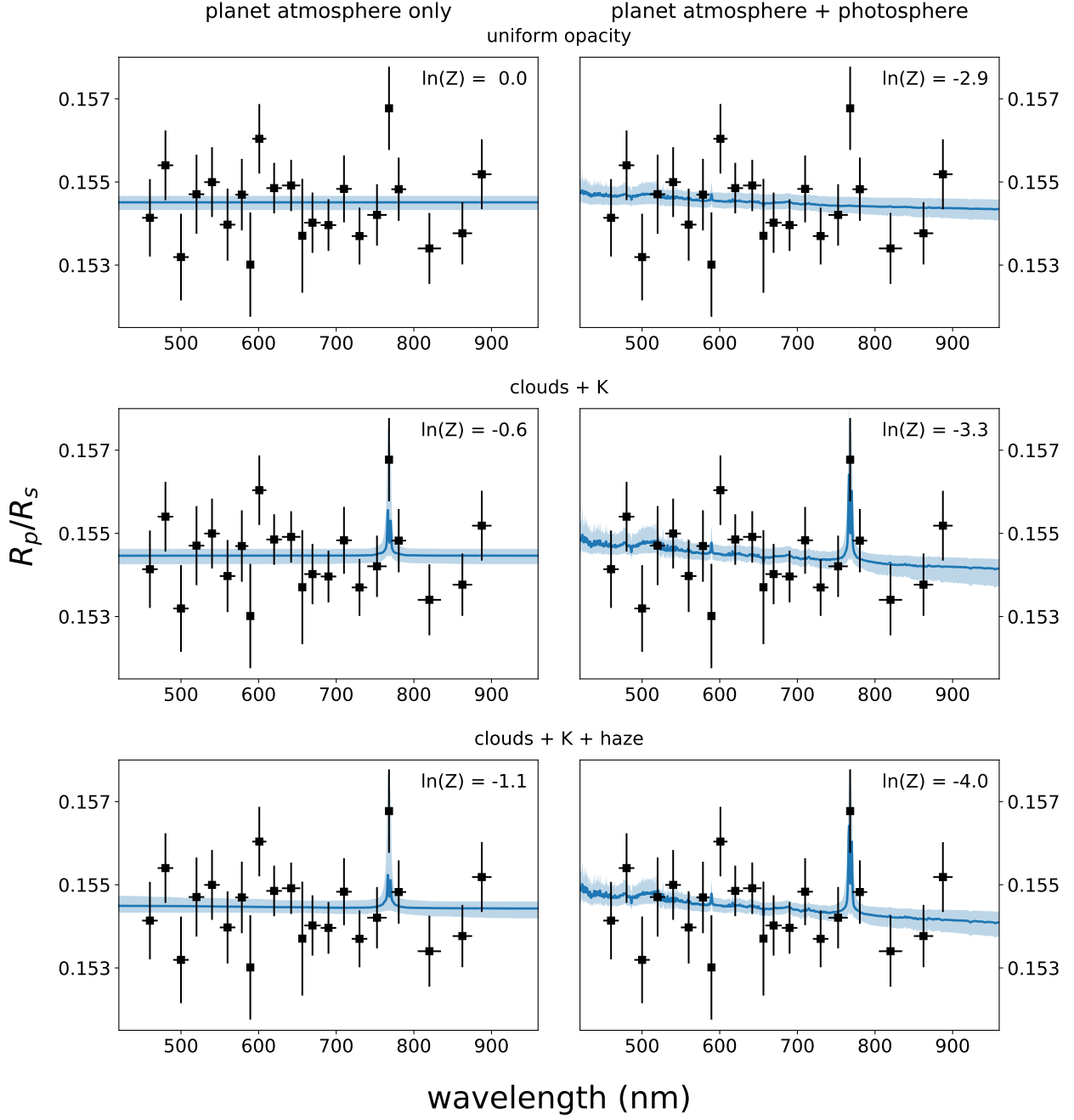


Figure 4. Comparison of models for the observed transmission spectrum of WASP-4b; three separate models for the planetary atmosphere are presented alone (left) or combined with a model for the heterogeneous stellar photosphere (right). The atmosphere models, described in more detail in Section 4.1, include (top) a uniform opacity atmosphere, (middle) a cloud deck and K absorption feature, and (bottom) a cloud deck, K absorption feature, and scattering haze. The shaded region represents the 67% confidence interval around the mean model fit. Bayes factors are given for each model relative to the flat line model (upper left). A constant offset of +0.1545 is added to the y-axis before fitting these models; as described in Section 3.6, the combined spectrum gives only relative values $\Delta R_p/R_s$.

Table 3. Parameters for our combined photosphere and atmosphere models. Not all parameters are incorporated in all models; rather, Bayesian log-evidences are used to compare separate models.

Model component	Parameter	Units	Description	Prior distribution
Base	$(R_p/R_s)_0$	-	Radius corresponding to the top of the cloud layer or $\tau \gg 1$	Uniform(0.1,0.2)
	P_0^\dagger	bar	Reference pressure at $(R_p/R_s)_0$	Log-uniform($10^{-6}, 10^6$)
	T^\dagger	K	Average temperature of the optically-thin region	Uniform(1000,2400)
Atomic features	X	-	Mixing ratio of species X	Log-uniform($10^{-30}, 1$)
Haze	a	-	Amplitude of the haze cross-section power law	Log-uniform($10^{-20}, 10^{10}$)
	γ	-	Index of the haze cross-section power law	Uniform(-4,0)
Stellar photosphere	T_{occ}	K	Average temperature of the transit chord [‡]	Fixed(5540)
	T_{het}	K	Average temperature of the heterogeneous surface features	Fixed(4200)
	F_{het}	-	Fraction of the unocculted photosphere covered by spots	Uniform(0,1)

[†] The temperature and reference pressure are not modeled in the case of a uniform opacity atmosphere.

[‡] This excludes any occulted features which can be identified in the light curve.

In the retrieval code described above, we characterize the heterogeneity with three parameters: T_{occ} , the effective temperature of the transit chord, T_{het} , the mean effective temperature of the heterogeneous features, and F_{het} , the fraction of the unocculted photosphere which is covered in spots. We then compute the effect on the observed transit depth following Section 3.5 and Equation 2, but replacing the area ratio $(R_{spot}/R_s)^2$ with a covering fraction F_{het} .

We test more complex parameterizations which include multiple types of unocculted heterogeneities (e.g., Zhang et al. 2018), but determine from the evidences that they are not warranted by the data.

Given the quality of our data, we find it necessary to fix some parameters of the heterogeneity model to reasonable values. In Table 3 we fix the spot temperature to match typical spots on the Sun, and the transit chord temperature to the measured effective temperature of the photosphere.

F_{het} is allowed to vary over all possible values, and serves as a simple measure of the level of stellar contamination in the transmission spectrum. However, we note that F_{het} represents an *enhancement* over the effect of the large occulted spot in Transit 3, which has already been corrected for in Section 3.5.

4.1.2. Retrieval results

In Figure 4 and Table 4 we present the best-fit spectra and Bayes factors for our models, as well as the posterior values for F_{het} when it is fitted; the other parameters do not converge enough from their prior distributions to be of interest.

The evidences weakly favor the haze-free models versus the hazy model with $\ln(B_m) \sim -1.1$, which is analogous to a 2σ favorability in the frequentist framework (e.g., Benneke & Seager 2013). Even when the haze is present, the amplitude a of the haze opacity is small; the posterior distributions do not converge well enough from the log-uniform prior to offer a meaningful upper limit, but the mode of the distribution for a lies between 10^{-15} and 10^{-10} , which is lower than the prior would suggest.

A uniform opacity atmosphere is also weakly favored versus an atmosphere with weak potassium abundance ($\ln(B_m) \sim -0.5$). For comparison, haze-free atmospheres with sodium are disfavored by $\ln(B_m) \sim -1.5$. This indicates that potassium is more evident in this spectrum than other species, but that the data neither justify nor exclude its inclusion in the model. The abundance of potassium, if it is present, is a strong function of the reference pressure: in low-pressure atmospheres, the abundance of potassium must be higher to produce the same absorption signal as it would under higher pressures. Therefore we cannot offer any meaningful constraints on the mixing ratio without prior knowledge of the pressure profile of the atmosphere, which is not constrained even to within a few orders of magnitude.

Table 4. Bayes factors for the 3x2 models in Figure 4, including atmosphere-only (A) and combined atmosphere-photosphere (A+P) models. $\ln(B_m)$ is calculated relative to the flat line model. Median values of F_{het} are given with 67% confidence intervals for the models which include the photosphere.

Model	$\ln(B_m)_A$	$\ln(B_m)_{A+P}$	F_{het}
uniform opacity	0.0	-2.9	$0.016^{+0.018}_{-0.011}$
clouds + K	-0.6	-3.3	$0.029^{+0.026}_{-0.019}$
clouds + K + haze	-1.1	-4.0	$0.030^{+0.025}_{-0.019}$

Models which include the photosphere are disfavored on average by $\ln(B_m) \sim -2.8$. While it is quite likely that $F_{het} > 0$, our data do not permit us to distinguish this effect beyond what we have already corrected for in Section 3.5. Models with a heterogeneous component are consistent with spot fractions $F_{het} < 5\%$. Using the code described in Rackham et al. (2018) to model the stellar variability as a function of spot covering fraction, we find that the fitted covering fractions are consistent with the 1% V-band variability observed by ASAS-SN.

The effect of even a low spot covering fraction is apparent in Figure 4, and is degenerate with the effect of a scattering haze. For this reason, we recommend the joint modeling of the photosphere and atmosphere in future analyses of optical transmission spectra, and the use of Bayesian evidences or likelihood parameters for model selection (see Pinhas et al., submitted to MNRAS).

4.2. Starspots

In Figure 1 we have identified spot-crossing features in the light curves of Transits 2 and 3. We have excluded these points from our analysis of the transmission spectrum. In this section, we fit combined spot and transit models to determine the size, location, and temperature of the spots. We use SPOTROD (Béky et al. 2014) to produce light curves of transits with a single spot-crossing event, and PyMultiNest to sample the parameter space and to calculate the Bayesian log-evidence $\ln(Z)$ for different models. In addition to the usual transit parameters, SPOTROD models the Cartesian position and radius of the spot (in units of stellar radii) as well as the average spot-to-stellar flux contrast. The details of the spot fitting code are discussed in Espinoza et al. (2018).

We use a second-order polynomial to detrend the white light curve from each night, then fit the combined spot and star model. We limit the uniform prior on spot size to $R_{spot}/R_s < 0.08$ or $0 < R_{spot}/R_s < 1$ to probe for small and large spots. Using the effective temperature in Table 2 as a

Table 5. Best-fit parameters for the spot-crossing events in the light curves of Transits 2 and 3. The data for Transit 2 are consistent with both a small (2a) and large spot (2b). Median and 1σ confidence intervals are reported.

	R_{spot}/R_s	f_{spot}/f_s	$T_s - T_{\text{spot}}$ (K)	T_{spot} (K)
Transit 2a	$0.05^{+0.01}_{-0.01}$	$0.49^{+0.20}_{-0.28}$	880 ± 360	4670 ± 360
Transit 2b	$0.15^{+0.11}_{-0.07}$	$0.79^{+0.11}_{-0.18}$	410 ± 280	5140 ± 280
Transit 3	$0.27^{+0.03}_{-0.02}$	$0.91^{+0.01}_{-0.01}$	100 ± 5	5442 ± 50

prior, the code then constructs and fits models for the flux contrast as a function of wavelength.

The optimal model parameters for each spot feature are presented in Table 5. The precision of the Transit 2 data only permits us to fit the spot feature and spectrum in a single, white-light bin. For this night, we find that our data are consistent with one- and two-spot solutions, as well as large and small spots. We opt for the simpler single-spot models and present the fit for both a single large and single cool spot. The data for Transit 3 allow us to rule out every model except a single large spot ($\Delta \ln(Z) > 2$). For this transit we are able to fit the spot spectrum to the same bins as our transmission spectrum, as shown in Figure 5.

The large spot model for Transit 3 seems inconsistent with smaller individual spots observed on the Sun; nevertheless, such wide features have been observed in transits before (e.g., Espinoza et al. 2018). While our evidences do not favor a two-spot model, it is possible that the large, low-contrast spot is in fact a more complex active region consisting of multiple spots and/or faculae, as has been observed on the Sun. Regardless of the actual structure of the feature, its average contrast is well-described by a single spot with $T_s - T_{\text{spot}} = 100$ K, and we choose this model to calculate the correction in Section 3.5.

4.2.1. Spot sizes and temperatures

The results presented in Table 5 for Transit 3 suggest that WASP-4 can host large spots with low temperature contrasts (~ 100 K). However, the temperature we have chosen to model the spots in Table 3 reflects more Sun-like photospheric activity, with small and cool spots. In short, we do not believe that the spots discovered in our transit light curves are necessarily representative of the spots present elsewhere in the photosphere or below our detection threshold. We offer these arguments as justification:

1. Sanchis-Ojeda et al. (2011) find evidence for a persistent small spot in multiple transit observations of WASP-4b. In this case, they find a spot with radius $R_{\text{spot}}/R_s \approx 0.05$ and temperature contrast $T_s - T_{\text{spot}} \approx$

600 K, suggesting that more Sun-like spots do exist on WASP-4.

2. Cool spots have previously been discovered through Doppler imaging of active G and K dwarfs (e.g., O’Neal et al. 1998, 2004). However, this method of spot detection is subject to biases regarding spot size and temperature.
3. While the precise radius and temperature of the spot observed during Transit 2 are largely unconstrained, we can say with at least 1σ confidence that the spot is smaller and cooler than the spot observed during Transit 3. Therefore, even our data suggest that small spots may be present in the photosphere of WASP-4.

It is nevertheless possible that the distribution of spot sizes on WASP-4 favors larger and hot spots, in which case the spectral contamination signal would be less severe. However, our use of the contamination model is not meant to accurately account for the actual surface heterogeneity, but rather to demonstrate that this aspect should not be ignored in determining the planet’s transmission spectrum.

5. COMPARISON TO *Gemini*/GMOS TRANSMISSION SPECTRUM

Huitson et al. (2017) (hereafter H17) have previously published an optical transmission spectrum of WASP-4b from four combined transit observations with *Gemini*/GMOS, and find that their data favor a uniform opacity model. In this section, we assess the agreement of their results with our data from IMACS. In Figure 6 we compare the red and blue GMOS spectra with the combined spectrum from this work, and find that the slopes of our spectra are generally in agreement.

5.1. Absorption features

The H17 data reveal tentative evidence for Na absorption at 586 nm. Our data reveal no evidence for such a feature, but we concede that the quality of our data in the blue is not sufficient to definitively rule out an atomic feature of small optical depth.

H17 exclude bins in their spectra which are coincident with terrestrial telluric O_2 absorption, finding that their correction for differential atmospheric refraction was not reliable in this wavelength range. Magellan is equipped with an ADC so our data do not require this correction (see Section 3.3). As a result, we are able to test for a K I feature in the same region, and achieved ambiguous results as discussed in Section 4.1.

5.2. Re-analysis of H17 Spectrum

Given the discussion in Section 4.1.1, we believe it worthwhile to re-analyze the reduced H17 spectrum using our own

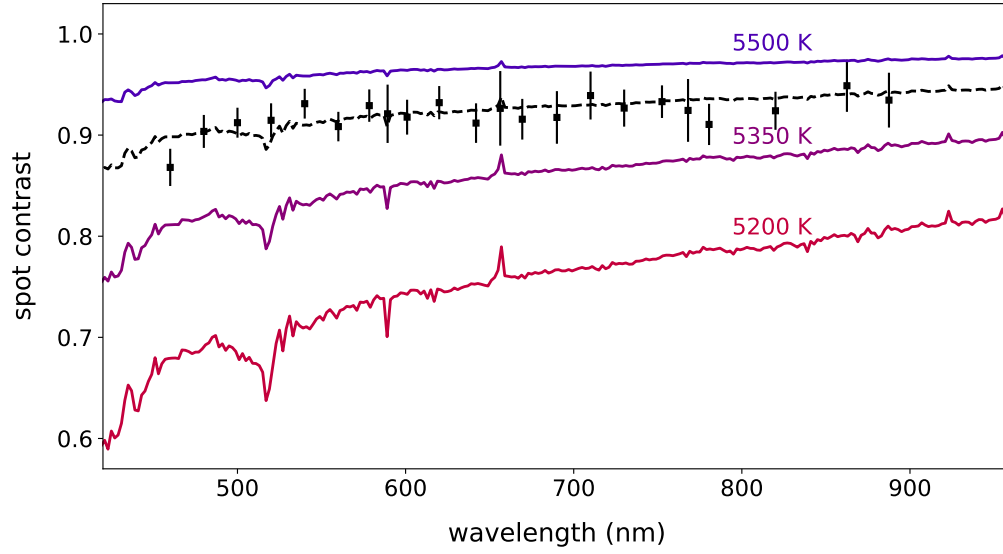


Figure 5. Best-fit model for the occulted spot in the light curve of Transit 3. The data are best described by a single large, warm spot with a radius of $0.27 R_s$ and a temperature contrast of 100 K.

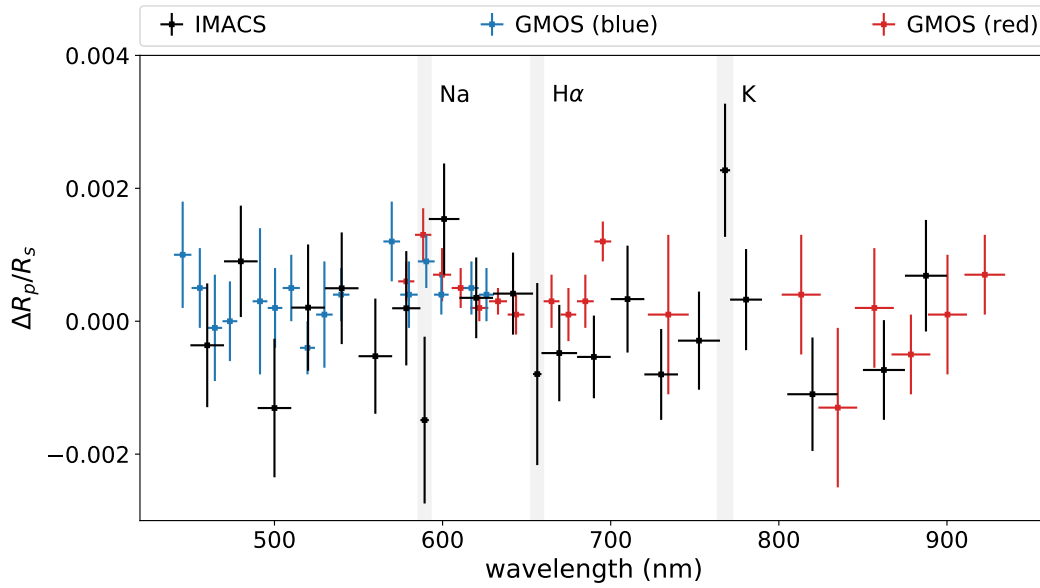


Figure 6. A comparison of our data to the optical transmission spectrum of WASP-4b as observed by *Gemini*/GMOS (Huitson et al. 2017, Table 4). The wavelengths of potential atomic features are highlighted.

retrieval code which incorporates TLSE contamination. We follow the same methodology as Section 4.1, but include sodium instead of potassium. We find that a featureless planet atmosphere is moderately favored over a hazy atmosphere ($\ln(B_m) \sim -1.5$), but only weakly over a haze-free model with sodium ($\ln(B_m) \sim -0.6$); these factors are generally consistent with H17’s conclusions. As with our data, models which include the stellar photosphere are disfavored at $\ln(B_m) \sim -3.8$. However, we note that H17 observed their red and blue bins during separate transits, which may mask the slope introduced by TLSE contamination.

6. CONCLUSIONS

We have extracted a combined optical transmission spectrum from four transit observations of WASP-4b using *Magellan*/IMACS, and use a MultiNest-based retrieval code to test different atmospheric models for the data. We find that the data disfavor an atmosphere with scattering hazes by $\ln(B_m) \sim -1.1$, or $\sim 2\sigma$ in the frequentist framework.

The presence of potassium is only weakly disfavored by the evidences, while by comparison the presence of sodium is moderately disfavored. However, [Huitson et al. \(2017\)](#) find some evidence for Na absorption from more precise data in that wavelength range. Our combined results suggest that the atmosphere of WASP-4b is largely opaque to visible light, but might yet contain a narrow absorption region for Na/K above the optically thick layer.

Furthermore, we are able to fit the size and contrast of the star spots occulted by the planet during Transits 2 and 3. Assuming a single spot model, the quality of the data from Transit 3 allows us to tightly constrain the spot size and contrast, which suggest a spot that is much larger and warmer than is typical for spots on the Sun. We use this spot model to correct for the wavelength-dependent effect on the transmission spectrum from Transit 3 before averaging the individual nights’ spectra. More complex models which include several spots and faculae could be consistent with the data as well, but are not justified by the evidences.

More and higher quality observations of this target must be conducted to confirm these results. In particular, high precision space-based observations could confirm or constrain the suspected absorption by sodium and potassium. However, we note that the presence and strength of a scattering haze is particularly degenerate with the presence of spots and faculae on the star. Since WASP-4 is known to be variable, the stellar photosphere and planet atmosphere should be modeled simultaneously in future analyses of this planet’s transmission spectrum.

The results reported herein benefited from collaborations and/or information exchange within NASA’s Nexus for Exoplanet System Science (NExSS) research coordina-

tion network sponsored by NASA’s Science Mission Directorate. This paper includes data gathered with the 6.5 meter Magellan-Baade Telescope located at Las Campanas Observatory, Chile. This research has made use of NASA’s Astrophysics Data System, and the Python modules Matplotlib, Numpy, and SciPy.

A.B. acknowledges support from the NASA Earth and Space Science Fellowship Program under grant No. 80NSS-C17K0470. B.R. acknowledges support from the National Science Foundation Graduate Research Fellowship Program under grant No. DGE-1143953.

REFERENCES

- Apai, D., Rackham, B. V., Giampapa, M. S., et al. 2018, ArXiv e-prints. <https://arxiv.org/abs/1803.08708>
- Beerer, I. M., Knutson, H. A., Burrows, A., et al. 2011, *ApJ*, 727, 23, doi: [10.1088/0004-637X/727/1/23](https://doi.org/10.1088/0004-637X/727/1/23)
- Béky, B., Kipping, D. M., & Holman, M. J. 2014, *MNRAS*, 442, 3686, doi: [10.1093/mnras/stu1061](https://doi.org/10.1093/mnras/stu1061)
- Benneke, B., & Seager, S. 2013, *ApJ*, 778, 153, doi: [10.1088/0004-637X/778/2/153](https://doi.org/10.1088/0004-637X/778/2/153)
- Brown, T. M. 2001, *ApJ*, 553, 1006, doi: [10.1086/320950](https://doi.org/10.1086/320950)
- Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, *A&A*, 564, A125, doi: [10.1051/0004-6361/201322971](https://doi.org/10.1051/0004-6361/201322971)
- Carter, J. A., & Winn, J. N. 2009, *ApJ*, 704, 51, doi: [10.1088/0004-637X/704/1/51](https://doi.org/10.1088/0004-637X/704/1/51)
- Deming, D., Wilkins, A., McCullough, P., et al. 2013, *ApJ*, 774, 95, doi: [10.1088/0004-637X/774/2/95](https://doi.org/10.1088/0004-637X/774/2/95)
- Dressler, A., Bigelow, B., Hare, T., et al. 2011, *PASP*, 123, 288, doi: [10.1086/658908](https://doi.org/10.1086/658908)
- Espinoza, N., & Jordán, A. 2015, *MNRAS*, 450, 1879, doi: [10.1093/mnras/stv744](https://doi.org/10.1093/mnras/stv744)
- Espinoza, N., Rackham, B., Jordán, A., et al. 2018, *MNRAS*, under review
- Gillon, M., Smalley, B., Hebb, L., et al. 2009, *A&A*, 496, 259, doi: [10.1051/0004-6361/200810929](https://doi.org/10.1051/0004-6361/200810929)
- Heng, K., & Kitzmann, D. 2017, *MNRAS*, 470, 2972, doi: [10.1093/mnras/stx1453](https://doi.org/10.1093/mnras/stx1453)
- Hoyer, S., López-Morales, M., Rojo, P., et al. 2013, *MNRAS*, 434, 46, doi: [10.1093/mnras/stt962](https://doi.org/10.1093/mnras/stt962)
- Huitson, C. M., Désert, J. M., Bean, J. L., et al. 2017, *AJ*, 154, 95, doi: [10.3847/1538-3881/aa7f72](https://doi.org/10.3847/1538-3881/aa7f72)
- Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, *A&A*, 553, A6, doi: [10.1051/0004-6361/201219058](https://doi.org/10.1051/0004-6361/201219058)
- Jordán, A., Espinoza, N., Rabus, M., et al. 2013, *ApJ*, 778, 184, doi: [10.1088/0004-637X/778/2/184](https://doi.org/10.1088/0004-637X/778/2/184)
- Kipping, D. M. 2013, *MNRAS*, 435, 2152, doi: [10.1093/mnras/stt1435](https://doi.org/10.1093/mnras/stt1435)
- Kochanek, C. S., Shappee, B. J., Stanek, K. Z., et al. 2017, *PASP*, 129, 104502, doi: [10.1088/1538-3873/aa80d9](https://doi.org/10.1088/1538-3873/aa80d9)
- MacDonald, R. J., & Madhusudhan, N. 2017, *MNRAS*, 469, 1979, doi: [10.1093/mnras/stx804](https://doi.org/10.1093/mnras/stx804)
- Mandel, K., & Agol, E. 2002, *ApJL*, 580, L171, doi: [10.1086/345520](https://doi.org/10.1086/345520)
- Maxted, P. F. L., Koen, C., & Smalley, B. 2011, *MNRAS*, 418, 1039, doi: [10.1111/j.1365-2966.2011.19554.x](https://doi.org/10.1111/j.1365-2966.2011.19554.x)
- McCullough, P. R., Crouzet, N., Deming, D., & Madhusudhan, N. 2014, *ApJ*, 791, 55, doi: [10.1088/0004-637X/791/1/55](https://doi.org/10.1088/0004-637X/791/1/55)
- Nikolov, N., Henning, T., Koppenhoefer, J., et al. 2012, *A&A*, 539, A159, doi: [10.1051/0004-6361/201118336](https://doi.org/10.1051/0004-6361/201118336)
- Nikolov, N., Sing, D. K., Burrows, A. S., et al. 2015, *MNRAS*, 447, 463, doi: [10.1093/mnras/stu2433](https://doi.org/10.1093/mnras/stu2433)
- O’Neal, D., Neff, J. E., & Saar, S. H. 1998, *ApJ*, 507, 919, doi: [10.1086/306340](https://doi.org/10.1086/306340)
- O’Neal, D., Neff, J. E., Saar, S. H., & Cuntz, M. 2004, *AJ*, 128, 1802, doi: [10.1086/423438](https://doi.org/10.1086/423438)
- Pont, F., Knutson, H., Gilliland, R. L., Moutou, C., & Charbonneau, D. 2008, *MNRAS*, 385, 109, doi: [10.1111/j.1365-2966.2008.12852.x](https://doi.org/10.1111/j.1365-2966.2008.12852.x)
- Pont, F., Sing, D. K., Gibson, N. P., et al. 2013, *MNRAS*, 432, 2917, doi: [10.1093/mnras/stt651](https://doi.org/10.1093/mnras/stt651)
- Rackham, B., Espinoza, N., Apai, D., et al. 2017, *ApJ*, 834, 151, doi: [10.3847/1538-4357/aa4f6c](https://doi.org/10.3847/1538-4357/aa4f6c)
- Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, *ApJ*, 853, 122, doi: [10.3847/1538-4357/aaa08c](https://doi.org/10.3847/1538-4357/aaa08c)
- Ranjan, S., Charbonneau, D., Désert, J.-M., et al. 2014, *ApJ*, 785, 148, doi: [10.1088/0004-637X/785/2/148](https://doi.org/10.1088/0004-637X/785/2/148)
- Sanchis-Ojeda, R., Winn, J. N., Holman, M. J., et al. 2011, *ApJ*, 733, 127, doi: [10.1088/0004-637X/733/2/127](https://doi.org/10.1088/0004-637X/733/2/127)
- Seager, S., & Sasselov, D. D. 2000, *ApJ*, 537, 916, doi: [10.1086/309088](https://doi.org/10.1086/309088)
- Shappee, B. J., Prieto, J. L., Grupe, D., et al. 2014, *ApJ*, 788, 48, doi: [10.1088/0004-637X/788/1/48](https://doi.org/10.1088/0004-637X/788/1/48)
- Sing, D. K., Pont, F., Aigrain, S., et al. 2011, *MNRAS*, 416, 1443, doi: [10.1111/j.1365-2966.2011.19142.x](https://doi.org/10.1111/j.1365-2966.2011.19142.x)
- Sing, D. K., Lecavelier des Etangs, A., Fortney, J. J., et al. 2013, *MNRAS*, 436, 2956, doi: [10.1093/mnras/stt1782](https://doi.org/10.1093/mnras/stt1782)
- Spake, J. J., Sing, D. K., Evans, T. M., et al. 2018, *Nature*, 557, 68, doi: [10.1038/s41586-018-0067-5](https://doi.org/10.1038/s41586-018-0067-5)
- Wilson, D. M., Gillon, M., Hellier, C., et al. 2008, *ApJL*, 675, L113, doi: [10.1086/586735](https://doi.org/10.1086/586735)
- Winn, J. N., Holman, M. J., Carter, J. A., et al. 2009, *AJ*, 137, 3826, doi: [10.1088/0004-6256/137/4/3826](https://doi.org/10.1088/0004-6256/137/4/3826)
- Zhang, Z., Zhou, Y., Rackham, B., & Apai, D. 2018, ArXiv e-prints. <https://arxiv.org/abs/1802.02086>

APPENDIX

A. BINNED LIGHT CURVES

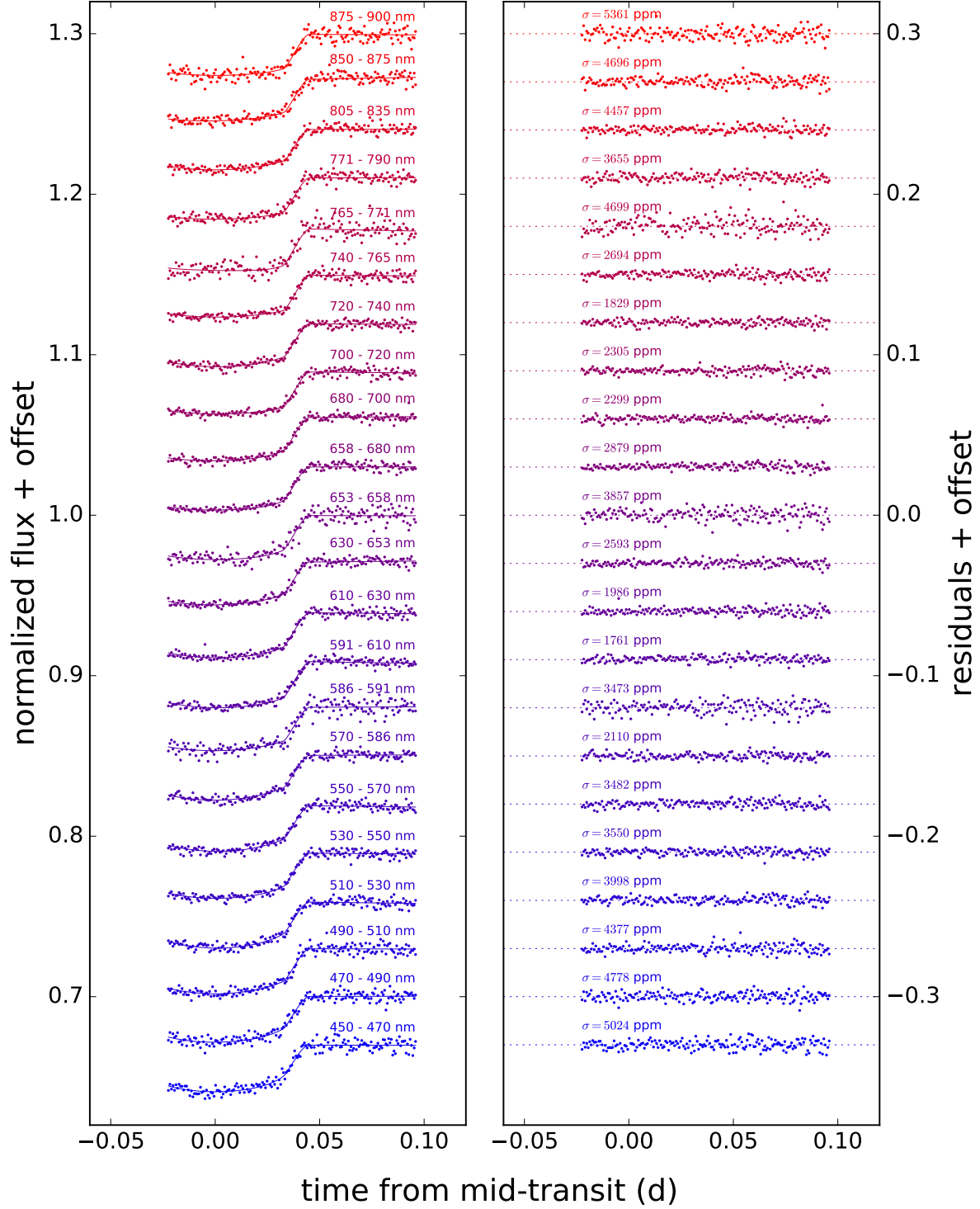


Figure A1. Binned light curve fits from the night of 2013-09-24. The models in the left panel include the linear systematics component.

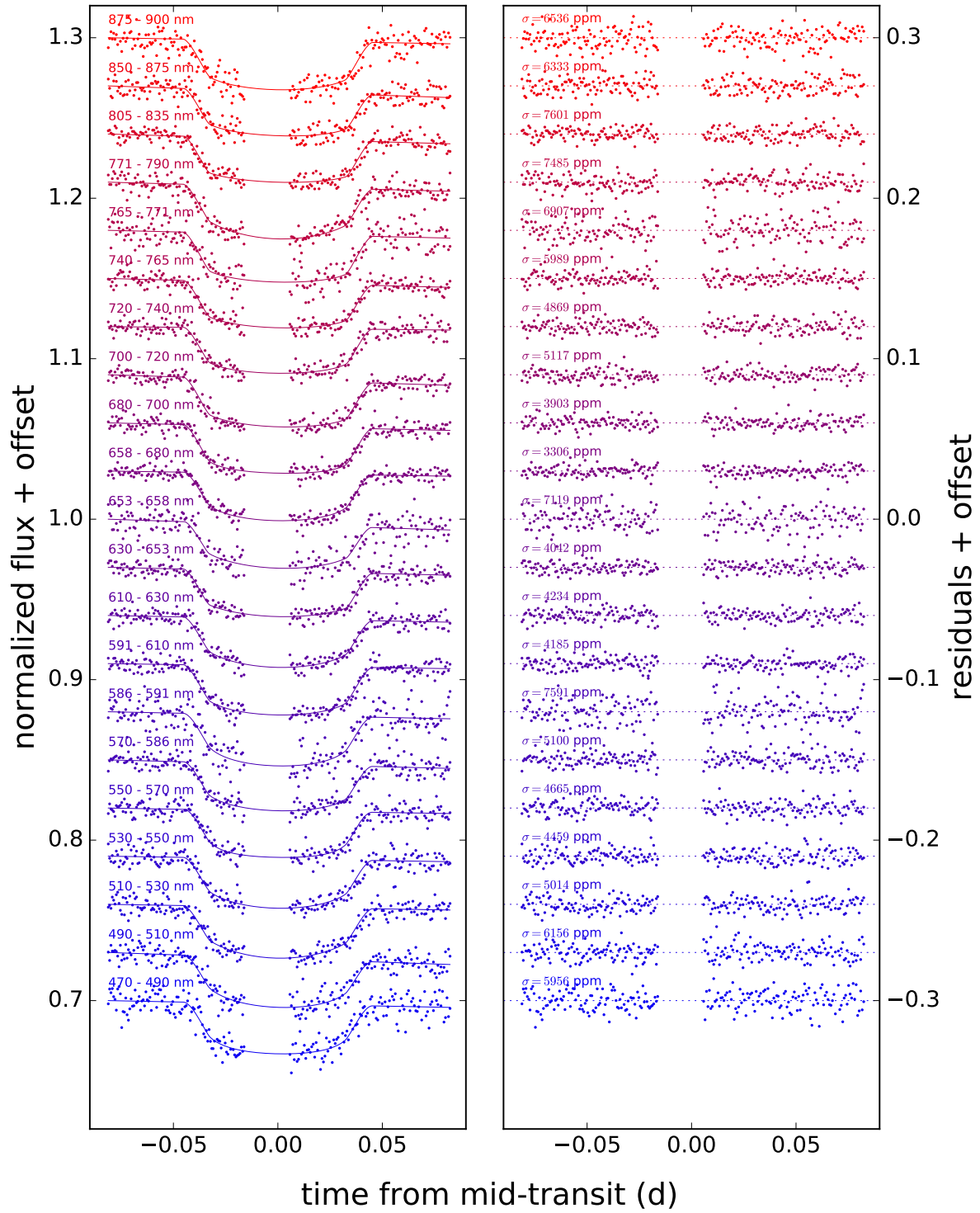


Figure A2. Binned light curve fits from the night of 2013-10-17. The models in the left panel include the linear systematics component. The 450 - 470 nm bin is removed as a $> 3\sigma$ outlier, and a potential spot-crossing event is excluded from the fit.

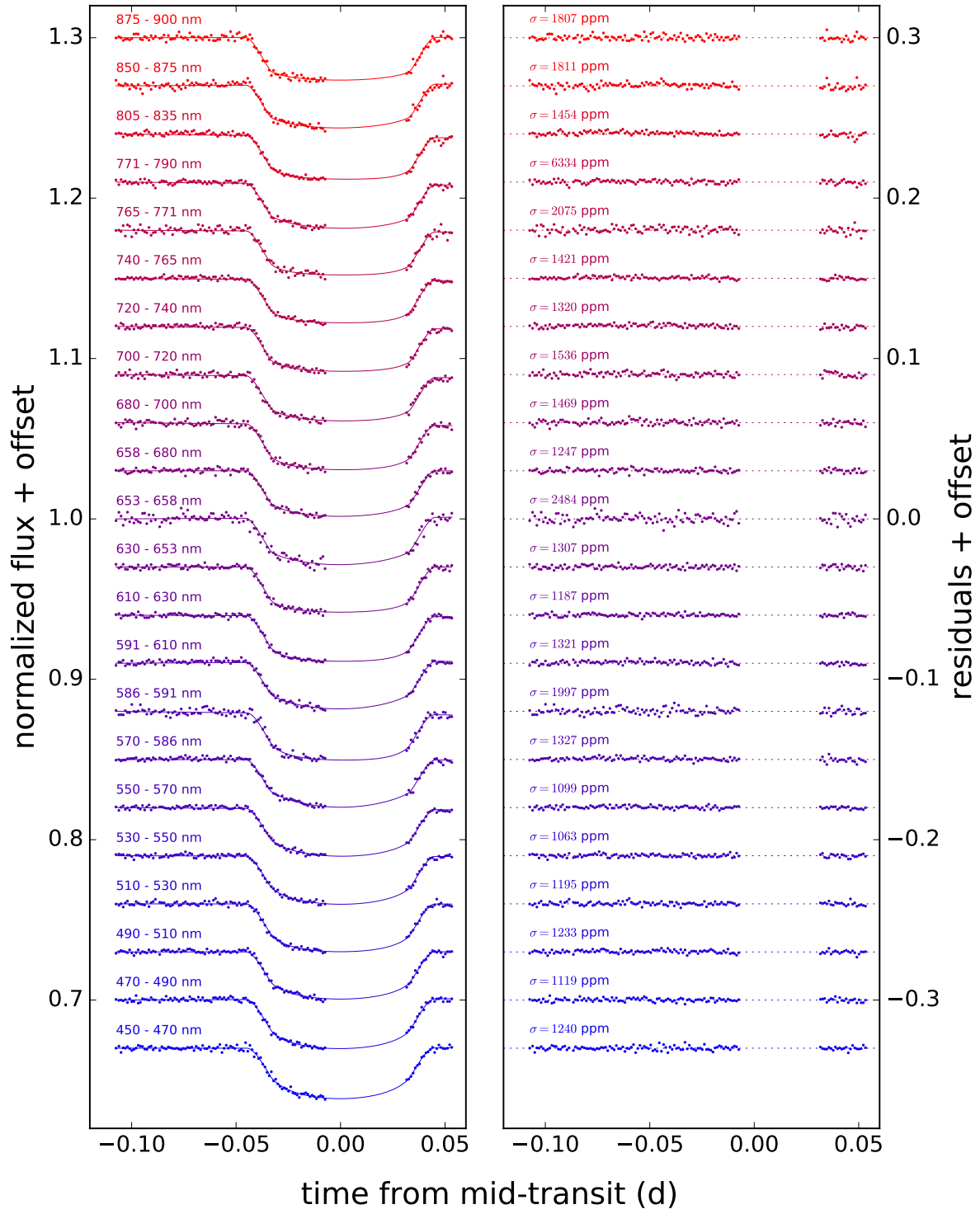


Figure A3. Binned light curve fits from the night of 2015-08-14. The models in the left panel include the linear systematics component. The spot-crossing event is excluded.

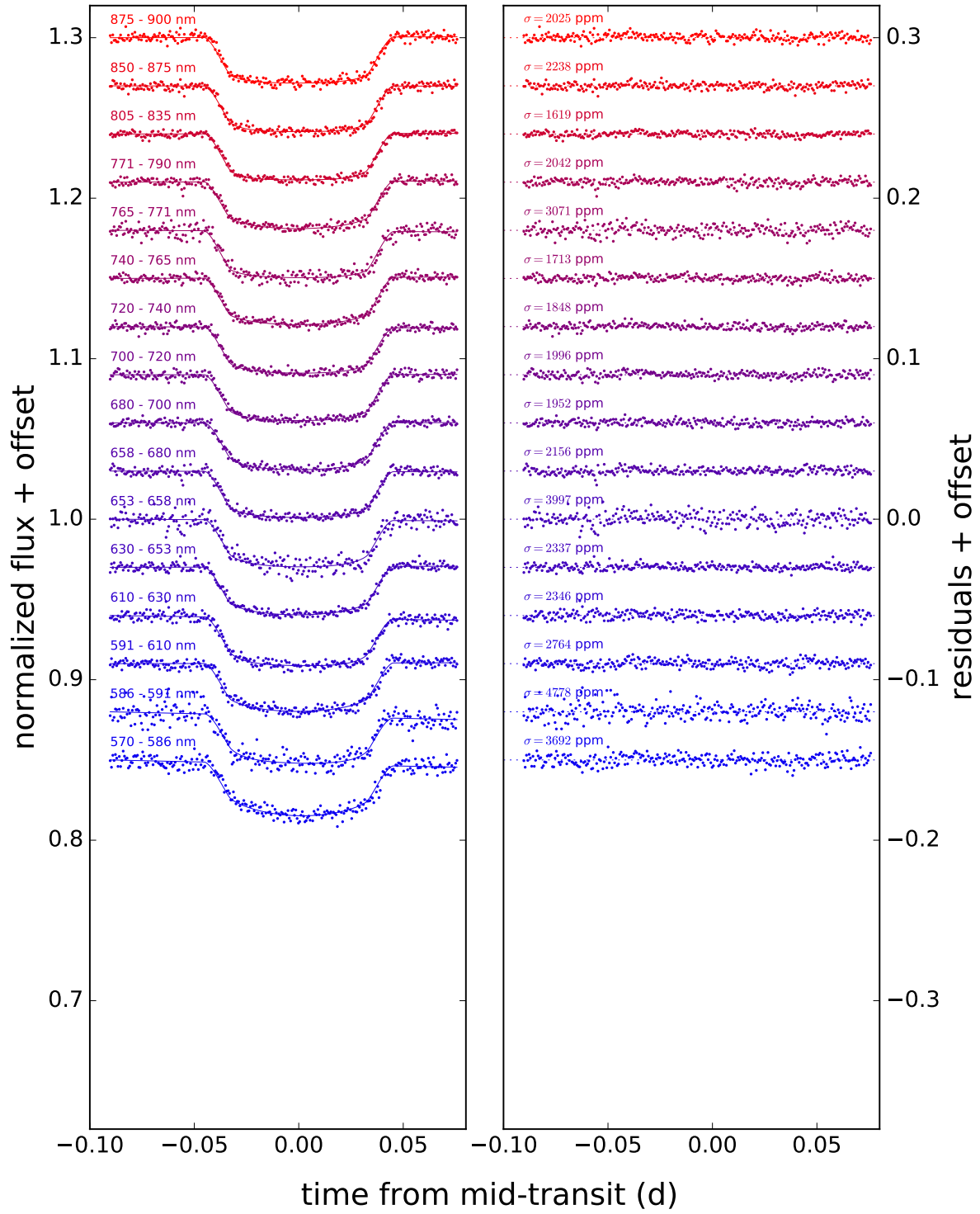


Figure A4. Binned light curve fits from the night of 2015-09-26. The models in the left panel include the linear systematics component. Six of the bins are excluded for lack of data due to the narrower filter.

B. TABULATED TRANSMISSION SPECTRA

Table B1. Data for the combined and individual transmission spectra shown in Figure 2 with 1σ uncertainties. The second column is the weighted mean of the mean-subtracted values from Transits 1-4, where the Transit 3 values have first been corrected for the presence of an occulted spot. The last column is the multiplicative effect on the fitted radii for Transit 3 due to the occulted spot, where the true radius $R_{p,0} = R_p / \epsilon^{1/2}$.

Bin (nm)	$\Delta R_p / R_s$	Transit 1	Transit 2	Transit 3	Transit 4	$\epsilon^{1/2}$
450.0 - 470.0	-0.0004 ± 0.0009	0.1509 ± 0.0029		0.1534 ± 0.0010		1.0121
470.0 - 490.0	0.0009 ± 0.0008	0.1513 ± 0.0021	0.1574 ± 0.0031	0.1548 ± 0.0009		1.0113
490.0 - 510.0	-0.0013 ± 0.0010	0.1515 ± 0.0022	0.1539 ± 0.0026	0.1523 ± 0.0013		1.0116
510.0 - 530.0	0.0002 ± 0.0010	0.1521 ± 0.0026	0.1580 ± 0.0027	0.1536 ± 0.0011		1.0118
530.0 - 550.0	0.0005 ± 0.0008	0.1533 ± 0.0017	0.1562 ± 0.0023	0.1538 ± 0.0010		1.0105
550.0 - 570.0	-0.0005 ± 0.0009	0.1541 ± 0.0025	0.1529 ± 0.0024	0.1529 ± 0.0010		1.0097
570.0 - 586.8	0.0002 ± 0.0009	0.1504 ± 0.0027	0.1536 ± 0.0022	0.1534 ± 0.0012	0.1613 ± 0.0017	1.0092
586.8 - 591.8 [†]	-0.0015 ± 0.0013	0.1515 ± 0.0035	0.1574 ± 0.0057	0.1524 ± 0.0018	0.1555 ± 0.0022	1.0096
591.8 - 610.0	0.0015 ± 0.0008	0.1584 ± 0.0020	0.1588 ± 0.0021	0.1535 ± 0.0013	0.1586 ± 0.0016	1.0089
610.0 - 630.0	0.0004 ± 0.0006	0.1545 ± 0.0017	0.1564 ± 0.0019	0.1534 ± 0.0010	0.1585 ± 0.0009	1.0089
630.0 - 653.8	0.0004 ± 0.0006	0.1479 ± 0.0021	0.1556 ± 0.0017	0.1549 ± 0.0010	0.1591 ± 0.0009	1.0085
653.8 - 658.8 [†]	-0.0008 ± 0.0014	0.1538 ± 0.0038	0.1491 ± 0.0036	0.1555 ± 0.0025	0.1569 ± 0.0021	1.0076
658.8 - 680.0	-0.0005 ± 0.0007	0.1518 ± 0.0021	0.1539 ± 0.0017	0.1536 ± 0.0011	0.1577 ± 0.0013	1.0082
680.0 - 700.0	-0.0005 ± 0.0006	0.1500 ± 0.0018	0.1583 ± 0.0017	0.1537 ± 0.0014	0.1574 ± 0.0008	1.0079
700.0 - 720.0	0.0003 ± 0.0008	0.1534 ± 0.0021	0.1588 ± 0.0019	0.1528 ± 0.0015	0.1581 ± 0.0012	1.0078
720.0 - 740.0	-0.0008 ± 0.0007	0.1519 ± 0.0019	0.1550 ± 0.0021	0.1522 ± 0.0013	0.1577 ± 0.0009	1.0076
740.0 - 765.0	-0.0003 ± 0.0007	0.1535 ± 0.0026	0.1560 ± 0.0023	0.1504 ± 0.0013	0.1594 ± 0.0010	1.0074
765.0 - 771.0 [†]	0.0023 ± 0.0010	0.1565 ± 0.0049	0.1616 ± 0.0038	0.1531 ± 0.0019	0.1614 ± 0.0013	1.0073
771.0 - 790.0	0.0003 ± 0.0008	0.1506 ± 0.0027	0.1619 ± 0.0022	0.1521 ± 0.0012	0.1588 ± 0.0011	1.0071
805.0 - 835.0	-0.0011 ± 0.0009	0.1461 ± 0.0023	0.1551 ± 0.0023	0.1513 ± 0.0019	0.1589 ± 0.0011	1.0070
850.0 - 875.0	-0.0007 ± 0.0007	0.1480 ± 0.0034	0.1542 ± 0.0026	0.1512 ± 0.0019	0.1584 ± 0.0008	1.0066
875.0 - 900.0	0.0007 ± 0.0008	0.1517 ± 0.0036	0.1611 ± 0.0035	0.1528 ± 0.0014	0.1594 ± 0.0011	1.0064

[†] Denotes bins centered on Na, H α , and K absorption lines.