

GLOBAL CLIMATE AND ATMOSPHERIC COMPOSITION OF THE ULTRA-HOT JUPITER WASP-103b FROM HST AND Spitzer PHASE CURVE OBSERVATIONS

LAURA KREIDBERG,^{1,2} MICHAEL R. LINE,³ VIVIEN PARMENTIER,⁴ KEVIN B. STEVENSON,⁵ TOM LOUDEN,⁶
MICKAËL BONNEFOY,⁷ JACQUELINE K. FAHERTY,⁸ GREGORY W. HENRY,⁹ KEIVAN STASSUN,¹⁰ JACOB L. BEAN,¹¹
JONATHAN J. FORTNEY,¹² ADAM P. SHOWMAN,¹³ JEAN-MICHEL DÉSERT,¹⁴ AND JACOB ARCANGELI¹⁴

¹ Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138

² Harvard Society of Fellows, 78 Mount Auburn Street, Cambridge, MA 02138

³ School of Earth & Space Exploration, Arizona State University, Tempe AZ 85287, USA

⁴ Aix Marseille Univ, CNRS, LAM, Laboratoire d'Astrophysique de Marseille, Marseille, France

⁵ Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

⁶ Department of Physics, University of Warwick, Coventry, CV4 7AL, UK

⁷ Univ. Grenoble Alpes, IPAG, F-38000 Grenoble, France. CNRS, IPAG, F-38000 Grenoble, France

⁸ American Museum of Natural History, Department of Astrophysics, Central Park West at 79th Street, New York, NY 10034, USA

⁹ Center of Excellence in Information Systems, Tennessee State University, Nashville, TN 37209, USA

¹⁰ Vanderbilt University, Dept. of Physics and Astronomy, 6301 Stevenson Center Ln, Nashville TN, 37235, USA

¹¹ Department of Astronomy & Astrophysics, University of Chicago, 5640 S. Ellis Avenue, Chicago, IL 60637, USA

¹² Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064

¹³ Department of Planetary Sciences and Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA

¹⁴ Anton Pannekoek Institute for Astronomy, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

ABSTRACT

We present thermal phase curve measurements for the hot Jupiter WASP-103b observed with *Hubble*/WFC3 and *Spitzer*/IRAC. The dayside planet-to-star flux ratio is $0.151 \pm 0.015\%$, $0.446 \pm 0.038\%$, and $0.569 \pm 0.014\%$ in bandpasses centered at 1.4, 3.6, and $4.5\mu\text{m}$, respectively. The phase curves have large amplitudes and negligible hotspot offsets, indicative of poor heat redistribution to the nightside. We fit the phase variation with a range of climate maps and find they all provide reasonable fits to the data, but have large degeneracies in temperature (up to 1000 K for a given latitude/longitude). The phase-resolved spectra are consistent with blackbodies in the WFC3 bandpass, with brightness temperatures ranging from 2000 K on the nightside to 3000 K on the dayside. However, the dayside spectrum has a strong feature in the *Spitzer* bands, which we attribute to CO emission and a thermal inversion. The emission feature is not present on the nightside, suggesting changes in thermal structure with longitude. We retrieved the atmospheric composition with a 1D self-consistent modeling technique and found the atmosphere is moderately metal-enriched ($[\text{M}/\text{H}] = 23_{-13}^{+29} \times \text{solar}$), and the carbon-to-oxygen ratio is below 0.9 at 3σ confidence. In contrast to cooler hot Jupiters, we do not detect significant water absorption, which we attribute to partial H_2O dissociation and additional opacity from H- and trace amounts of hydrides/oxides. We compare the phase curves to 3D general circulation models and find that atmospheric drag is required to explain the small hotspot offsets, but that these models overpredict the phase curve amplitude. We also compare the WASP-103b spectra to brown dwarfs and young directly imaged companions and find these more massive and self-luminous objects have significantly larger water features, indicating that surface gravity and irradiation environment play an important role in shaping the spectra of hot Jupiters. These results highlight the 3D structure of exoplanet atmospheres and illustrate the importance of phase curve observations for understanding their complex chemistry and physics.

Keywords: planets and satellites: individual (WASP-103b), planets and satellites: atmospheres

Corresponding author: Laura Kreidberg
laura.kreidberg@cfa.harvard.edu

1. INTRODUCTION

Planets are round, rotating, and irradiated on one hemisphere at a time – all of which contribute to rich spatial structure in their climate and atmospheric composition. Short period, tidally locked planets are an extreme example, with one hot, continuously illuminated side. This asymmetry is expected to produce large gradients in temperature, chemistry, and cloud coverage with longitude (Showman et al. 2009; Kataria et al. 2016; Parmentier et al. 2016), and provides an opportunity to learn about atmospheric dynamics in a very different regime from the Solar System planets.

Exoplanets are so distant that they are not spatially resolved from their host stars, but it is still possible to reveal inhomogeneities in their atmospheres. One approach is to measure a phase curve, which consists of continuous monitoring of the planet-to-star flux ratio over a complete orbital revolution of the planet. This observation is sensitive to different longitudes at each orbital phase of the planet. The first phase curve of an exoplanet was observed with *Spitzer* for the hot Jupiter HD 189733b by Knutson et al. (2007), followed by additional *Spitzer* observations for about a dozen more systems (cataloged in Parmentier & Crossfield 2017). The observations revealed large day-night temperature contrasts (in excess of 300 K), and eastward shifted peak brightness due to heat circulation, as predicted by 3D models (Showman & Guillot 2002). These infrared measurements were complemented by optical phase curves from *Kepler* that show evidence for reflected light from patchy and possibly variable dayside clouds with a range of compositions (Borucki et al. 2009; Demory et al. 2013; Hu et al. 2015; Armstrong et al. 2016; Parmentier et al. 2016). A spectroscopic phase curve was observed for WASP-43b with *Hubble*/Wide Field Camera 3 (*HST*/WFC3) in the near-infrared, which provided the first phase-resolved measurements of an exoplanet’s water abundance and thermal structure (Stevenson et al. 2014b, 2017).

In this paper we present spectroscopic phase curve observations of the hot Jupiter WASP-103b, measured with *HST*/WFC3 and Spitzer/IRAC. This planet is an ideal target for phase curve observations, with an orbital period of just 22 hours and an equilibrium temperature of 2500 K (Gillon et al. 2014). Previous observations of WASP-103b’s atmosphere revealed a blackbody-like dayside emission spectrum, with possible evidence for a K_S -band emission feature (Cartier et al. 2017; Delrez et al. 2018). The optical transmission spectrum shows evidence for sodium and potassium absorption features that are consistent with expectations for a cloud-free atmosphere (Lendl et al. 2017).

WASP-103b is an archetype of the class of ultra-hot Jupiters, with orbital periods of about one day and day-side temperatures typically > 2000 K. These very hot planets were initially predicted to have inverted temperature pressure profiles due to strong optical absorption by TiO/VO in the upper atmospheres (Hubeny et al. 2003; Fortney et al. 2008); however, observations of their emission spectra have been inconclusive on their thermal structure and composition. In the near-infrared, where water is the dominant absorber, some spectra show water absorption features, some show emission features, and some are consistent with blackbody models (Madhusudhan et al. 2011; Crossfield et al. 2012; Stevenson et al. 2014a; Haynes et al. 2015; Evans et al. 2016; Beatty et al. 2017a,b; Sheppard et al. 2017; Arcangeli et al. 2018; Mansfield et al. submitted). A variety of explanations have been proposed for these results, including low metallicity or high carbon-to-oxygen compositions, dayside clouds, and finely tuned isothermal temperature pressure profiles. Recently, Arcangeli et al. (2018) showed that water dissociation and H- opacity on the hot dayside play an important role in the atmospheres of these ultra-hot planets and may be responsible for some of the blackbody-like near-IR spectra. In this work, we put these results in context by investigating the global thermal structure and composition of the ultra-hot Jupiter WASP-103b.

The structure of the paper is the following: in § 2 we describe the observations and data reduction. § 3 details the models fit to the phase curves. In § 4 we discuss results, including the phase curve amplitudes and hotspot offsets, the phase-resolved spectra, and estimates of the planet’s climate. In § 6 and 7, we compare the observations to general circulation model (GCM) predictions and spectra from similar temperature stars and directly imaged companions. § 8 concludes.

2. OBSERVATIONS AND DATA REDUCTION

We observed two full-orbit phase curves of WASP-103b with *HST*/WFC3 and one each with *Spitzer*/IRAC at 3.6 and 4.5 μm (from HST Program 14050 and Spitzer Program 11099, PI: L. Kreidberg). We also reduced two *HST*/WFC3 secondary eclipse observations of WASP-103b from HST Program 13660 (PI: M. Zhao).

2.1. *HST*/WFC3

The *HST* phase curve observations consisted of two visits on 26–27 February and 2–3 August 2015. Each visit was 15 orbits in duration and spanned 23 hours. The last half of orbit 15 in each visit was used for a gyro bias update and does not produce useable science data. We took a direct image of the star with the F126N

filter at the beginning of each orbit to determine the wavelength solution zero-point. The remainder of the orbit consisted of time-series spectroscopy with the G141 grism ($1.1 - 1.7 \mu\text{m}$) and the 256 x 256 pixel subarray. We used the SPARS10/NSAMP = 15 read-out mode, which has an exposure time of 103 seconds. To optimize the duty cycle of the observations, we used the spatial scan observing mode with a scan rate of 0.03 arcsec/s, alternating between forward and backward scanning on the detector. The scan height was 25 pixels and the peak counts were 3.5×10^4 photoelectrons per pixel. We collected a total of 18 spatial scan exposures per orbit. The two eclipse observations from Program 13660 had a similar observing setup (described in detail in [Cartier et al. 2017](#)).

We reduced the data from both programs using a custom pipeline developed for past analyses of WFC3 data (for details see [Kreidberg et al. 2014a,b, 2015](#)). Briefly, we use the optimal extraction algorithm of [Horne \(1986\)](#) to extract each up-the-ramp sample (or “stripe”) separately. The stripes are then summed to create the final spectrum. For each stripe, the extraction window is 24 pixels high and centered on the stripe midpoint. We estimate the background from the median of a region of the detector that is uncontaminated by the target spectrum (rows 5-50). The typical background counts are low (10-15 photoelectrons per pixel, roughly 0.03% of the peak counts from the target star). We note that the extracted spectrum includes flux from a nearby companion star, which is separated from WASP-103 by less than two pixels (0.2”;[Wöllert & Brandner 2015](#)). We account for this contamination later in the analysis.

2.2. Spitzer

We also obtained *Spitzer*/IRAC observations with 3.6 and $4.5 \mu\text{m}$ photometric filters (referred to as Channel 1 and Channel 2, respectively). The observations had the following setup. Each phase curve observation consisted of 30 hours of time series photometry, beginning three hours prior to one secondary eclipse and ending three hours after a second eclipse. We read out the full array and used 12 s exposures to maximize the duty cycle without saturating the detector. To minimize the intrapixel effect (variations in flux caused by imprecise pointing), we did not dither and also used PCRS peak-up¹ to improve the pointing accuracy. We began each observation with a 30-minute position settling period, followed by three Astronomical Observation Requests (AORs) of equal duration. At the beginning of each

AOR, the telescope was repointed to position the target in the “sweet spot” of the detector, where the response is fairly uniform over the pixel.

The data were reduced with the POET pipeline ([Stevenson et al. 2012; Cubillos et al. 2013](#)). We performed aperture photometry with an aperture size of 2.75 pixels (chosen from a grid of apertures between 2 and 4 pixels to minimize the residual noise in the light curve fits). We masked pixels that were flagged in the bad pixel mask provided in the ancillary data for the observations. The target centroid was determined with a two-dimensional Gaussian fit. The target remains centered near the sweet spot for the entire AOR in each observation, with a maximum drift of 0.1 pixels. We estimated and subtracted the background from an annulus with a radius of 7 to 15 pixels from the centroid. The contaminating flux from the nearby star is within the same pixel as the target, so we included it in the photometry and corrected it in the light curve fits.

2.3. Photometric Monitoring

To assess how stellar activity might impact the phase curve observations, we monitored WASP-103’s photometric variability over 158 nights during 2014 - 2016 with the Tennessee State University Celestron 14-inch (C14) automated imaging telescope (AIT), located at Fairborn Observatory in southern Arizona ([Henry 1999](#)). The observations of WASP-103 were made in the Cousins R passband with an SBIG STL-1001E CCD camera. Each observation consisted of 4–10 consecutive exposures on WASP-103 along with several dozen comparison stars in the same field. The individual consecutive frames were co-added and reduced to differential magnitudes (i.e., WASP-103 minus the mean brightness of the six best comparison stars). The nightly observations were corrected for bias, flat-fielding, pier-side offset, and differential atmospheric extinction.

The photometric analyses are summarized for each observing season in [Table 1](#). The standard deviation of WASP-103’s brightness in each season is given in column 4. The mean of the three standard deviations is 0.0058 mag, which is very comparable to the mean standard deviation of the six comparison stars (0.0055 mag). To maximize the possibility of detecting WASP-103’s rotation, we normalized the photometry such that each observing season has the same mean, thereby removing any long-term variability in WASP-103 and/or the comparison stars ([Figure 1](#), top panel). To estimate the stellar rotation period, we performed a periodogram analysis of the normalized data set based on least-squares fitting of sine curves. The resulting frequency spectrum and the phase curve computed with the best period are shown

¹ http://irsa.ipac.caltech.edu/data/SPITZER/docs/irac/pcrs_obs.shtml

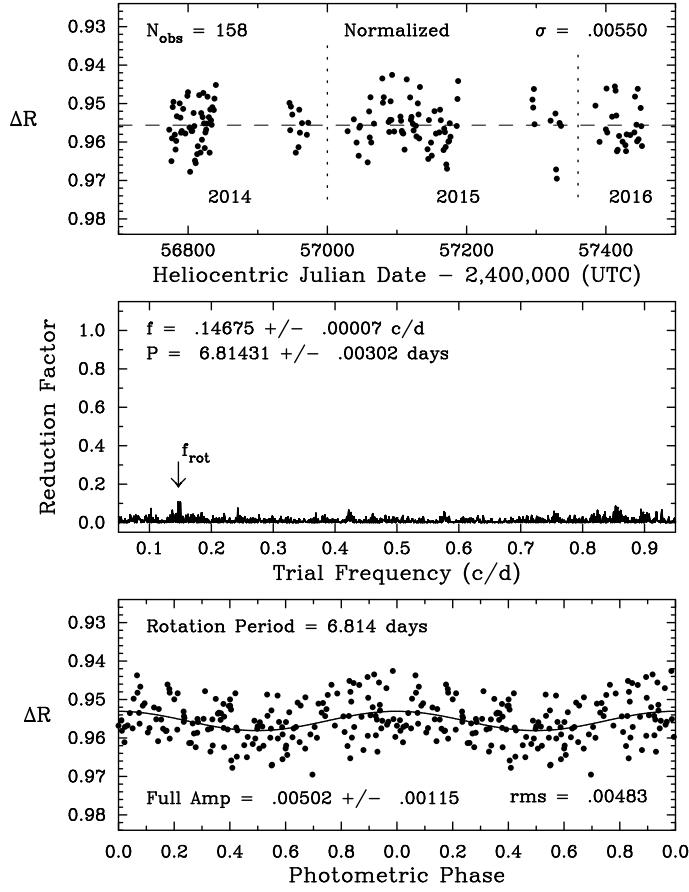


Figure 1. *Top:* The normalized nightly Cousins R band photometric dataset for WASP-103, acquired with the C14 automated imaging telescope at Fairborn Observatory. Vertical dashed lines denote separate observing seasons. Gaps are due to target visibility and the Arizona monsoon season (July - September). *Middle:* The frequency spectrum of the normalized dataset suggests low-amplitude variability with a period of 6.814 days. *Bottom:* The normalized dataset phased to the 6.814-day period, which we interpret as rotational modulation of a star spot or spots. A least-squares sine fit to the 6.814-day rotation period gives a peak-to-peak amplitude of just 0.005 mag.

in the middle and lower panels of Figure 1 respectively. The best-fit period is 6.814 days, which agrees closely with the estimated stellar rotation period of 6.855 days (based on the projected stellar rotation velocity reported in Gillon et al. 2014). The peak-to-peak variability amplitude is 0.005 mag. Based on the formalism in Zellem et al. (2017), we calculate that this variability will bias the measured eclipse depth by $\lesssim 10$ parts per million (ppm) from epoch to epoch, which is well below the photon-limited precision of our measurements.

3. LIGHT CURVE FITS

We fit a two-component model to the light curves. One component models the astrophysical signal (the

Table 1. Photometric Observations of WASP-103

Observing	Date Range	Sigma	Seasonal Mean
		(mag)	(mag)
2014	59	56722–56972	0.0057 0.9546 ± 0.0007
2015	73	57028–57335	0.0062 0.9549 ± 0.0007
2016	26	57385–57451	0.0055 0.9485 ± 0.0011

planet’s thermal phase variation and transit), and the other component models the systematic noise introduced by time-dependent changes in instrument performance. For each light curve, we fit the physical and systematic components simultaneously, such that the total observed flux as a function of time is given by $F(t) = F_{\text{physical}}(t) \times F_{\text{sys}}(t)$. For the *HST* data, where we observed two phase curves and two additional eclipses, we constrain the physical parameters to be the same for all visits, but allow some of the systematics parameters to vary (for more details see § 3.2.1). We fit the WFC3 band-integrated “white” light curve, as well as spectroscopic light curves created from 10 wavelength bins uniformly spaced at $0.05 \mu\text{m}$ intervals between 1.15 and $1.65 \mu\text{m}$.

3.1. Astrophysical Signal

We assume the measured astrophysical signal F_{physical} has the following form:

$$F_{\text{physical}}(\lambda, t) = T(\lambda, t) + c(\lambda, t) \times F_p/F_s(\lambda, t) \quad (1)$$

where λ is wavelength, $T(\lambda, t)$ is the transit model (the fraction of the stellar disk that is visible at time t), $F_p/F_s(\lambda, t)$ is the disk-integrated planet-to-star flux, and c is a correction factor for companion star dilution and the planet’s tidal distortion.

We calculated the transit model $T(t)$ with the **batman** package (Kreidberg 2015). Many of the physical parameters are tightly constrained by Southworth et al. (2015), so we fixed the orbital period, time of inferior conjunction, orbital inclination, and ratio of semi-major axis to stellar radius to the previously published values ($P = 0.925545613$ day, $t_0 = 2456836.2964455$ BJD_{TDB}, $i = 87.3^\circ$, and $a/R_s = 2.999$). As a test, we fit for these parameters with the *Spitzer* Channel 2 light curve, which has the best phase coverage and least systematic noise of the three data sets. We found that the transit parameters are consistent with the Southworth et al. (2015) results, so we proceeded with the remainder of the analysis holding those parameters fixed. The free

parameters for the transit model were the transit depth r_p and a linear limb darkening parameter u . More complex limb darkening laws with additional free parameters were not merited according to the Bayesian Information Criterion (BIC).

We calculated the planet-to-star flux F_p/F_s in two different ways. First, we fit a sinusoid with a period equal to the planet's orbital period. The free parameters were the sine curve amplitude and phase offset. For the second approach, we used the `spiderman` package (Louden & Kreidberg 2017) to model F_p/F_s . This package allows users to input a climate map (temperature or brightness as a function of latitude and longitude), and generate the corresponding flux ratio for an observation at time t . In our fit, we calculated the stellar flux with a NextGen model (Allard et al. 2012) interpolated to an effective temperature of 6110 K (Gillon et al. 2014), solar metallicity, and $\log g$ of 4.2 (in cgs units). For the planet flux, we tested three different maps: a two-temperature map, with a uniform dayside temperature T_d and a uniform nightside temperature T_n ; a spherical harmonic maps of degree two (with four free parameters); and the physically-motivated kinematic model from Zhang & Showman (2017), which has just three free parameters (the nightside temperature T_n , the change in temperature from day-to-night side Δ_T , and the ratio of radiative to advective timescales ξ). In all cases, we assumed that the planet is tidally locked, such that each orbital revolution corresponds to one complete rotation on its spin axis.

We scaled the planet-to-star flux by a correction factor c to account for dilution from the companion star and ellipsoidal variability due to the planet's tidal distortion. The correction factor took the form:

$$c(\lambda, t) = [1 + \alpha(\lambda)]A(t) \quad (2)$$

where $\alpha(\lambda)$ is the additional fractional flux from the companion star and $A(t)$ is the sky-projected area of the planet. We estimated $\alpha(\lambda)$ based on the best fit spectral energy distribution from Cartier et al. (2017). The companion star contribution ranges from 10% at $1.1\mu\text{m}$ to 16% at $4.5\mu\text{m}$. The uncertainty on the companion star flux contribution is less than 1%, which introduces negligible error in the estimated planet-to-star-flux compared to the photon noise. We calculated $A(t)$ using the analytic formula from Leconte et al. (2011a), equation B.9, which computes the projected area of a triaxial ellipsoid. We estimated the ellipsoid properties using Table B.3 of Leconte et al. (2011b), assuming the planet radius is $1.5R_{\text{Jup}}$ and age is 5 Gyr. The predicted ellipsoidal variability is shown in Figure 2. At quadrature, the projected area is 8% larger than at phase zero (mid-transit). Us-

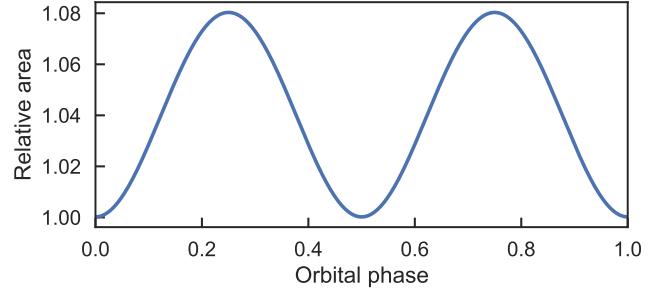


Figure 2. Projected area of the planet as a function of orbital phase, normalized to unity at phase zero. The area variation was predicted analytically using the model from Leconte et al. (2011a).

ing the analytic expression from Loeb & Gaudi (2003), we estimated the effect of Doppler beaming and found that it contributes less than 10 ppm to the measured flux.

3.2. Systematics

Both the *HST* and *Spitzer* phase curves have systematic noise caused by variations in the sensitivity of the instrument over time. For the *HST*/WFC3 data, the dominant systematic is an orbit-long exponential trend due to charge traps filling up over successive exposures (Long et al. 2015; Zhou et al. 2017). For *Spitzer* the primary source of noise is the intrapixel sensitivity effect. The detector's pixels do not have uniform sensitivity, so slight changes in telescope pointing cause the recorded flux to vary. In Figure 3, we show the raw light curves before systematic noise was removed. The systematics have comparable amplitude to the thermal phase variation signal, so they must be carefully corrected to recover the underlying planet-to-star flux.

3.2.1. *HST* Systematics

We fit the WFC3 systematics using an analytic model of the form:

$$F_{\text{sys}}(t) = (c S(t) + v_1 t_v + v_2 t_v^2)(1 - \exp(-a t_{\text{orb}} - b)) \quad (3)$$

where t_v is time elapsed since the first exposure in a visit and t_{orb} is time since the first exposure in an orbit. $S(t)$ is a scale factor equal to 1 for exposures with spatial scanning in the forward direction and s for reverse scans, to account for the upstream-downstream effect (McCullough & MacKenty 2012). The orbit-long ramp parameters are consistent for all the visits, so we constrained a , b , and s to have the same value for all visits in the final fit. The visit-long trends differ from visit to visit, so c , v_1 , and v_2 were allowed to vary between visits.

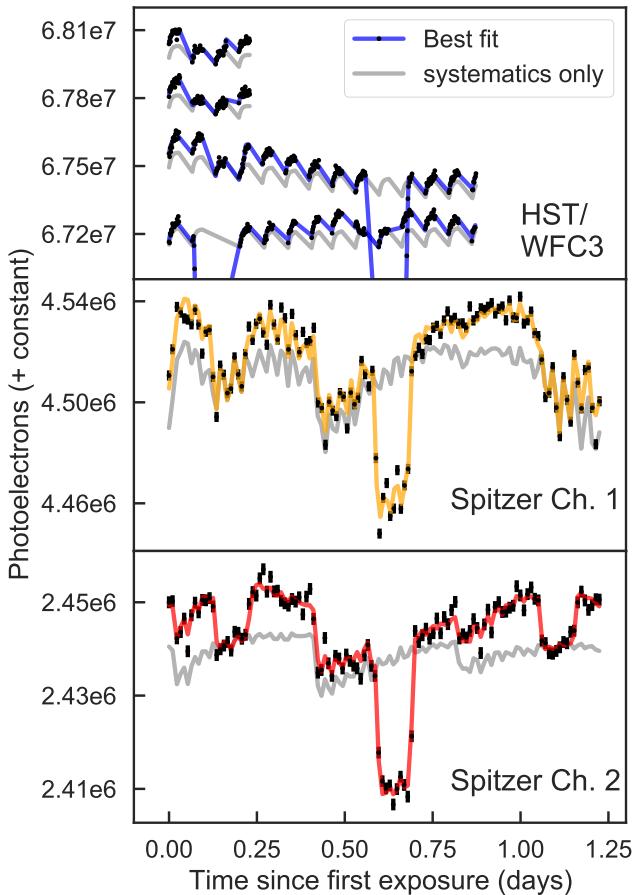


Figure 3. Raw light curves for WASP-103b observed with *HST*/WFC3 (top panel) and *Spitzer*/IRAC (bottom two panels). The data points are indicated with black dots. The *HST* data are unbinned, and the Spitzer data are binned in time segments of 15 minutes with error bars indicating the bin standard deviation. The colored lines show the best fit models, which include the astrophysical signal and instrument systematics. The gray lines indicate the contribution from the instrument systematics alone (which would be observed for a source with constant brightness and no planet). For visual clarity, we corrected the *HST* data for the upstream-downstream effect, separated the four visits by adding a flux offset, and zoomed in on the phase variation, so the transits are not displayed in the panel.

We fixed v_2 to zero for the two secondary eclipse observations from Program 13360, since the visit-long trend for shorter observations is fit well by a linear slope.

Some segments of the data exhibit stronger systematics than others, so we exclude these data in our final analysis. We drop the first orbit from every visit and the first exposure from every orbit (following common practice; see e.g. Kreidberg et al. 2014a). We also discard exposures from the last half of orbit 15 from the phase

curve observations, which were taken in staring mode to enable a gyro bias update. Since we observed two phase curves, we have complete orbital phase coverage of the planet despite discarding some data.

3.2.2. Spitzer Systematics

We fit the *Spitzer* systematics with the POET pipeline, which uses the BLISS mapping technique (Stevenson et al. 2012). This approach creates a map of the intrapixel sensitivity while simultaneously fitting for other systematics and the physical parameters of the system. The sensitivity map is determined by bilinear interpolation over a grid of knots centered on the stellar flux. Each knot’s sensitivity is calculated from the residuals to the light curve fit: the higher the flux values for data points near a given knot, the higher the detector sensitivity is at that position. To avoid overfitting, we chose the grid scale such that bilinear interpolation performed better than nearest neighbor interpolation. For the $3.6\mu\text{m}$ data, the grid scale was 0.008 pixel in both x and y . For $4.5\mu\text{m}$, the scale was 0.022 pixel. In addition to the intrapixel sensitivity variation, we fit the data for a linear trend in time. We tested a quadratic trend but did not find significant evidence for the additional model complexity based on the BIC.

3.3. Best Fits and Uncertainties

To determine the best fits, we performed a least-squares χ^2 minimization for each wavelength and model. For a subset of these cases, we also performed a Markov chain Monte Carlo (MCMC) analysis to estimate parameter uncertainties. We used emcee (Foreman-Mackey et al. 2013) to fit the *HST*/WFC3 light curves and differential evolution Monte Carlo for the *Spitzer* fits (Braak 2006). We ran the MCMC until convergence according to the Gelman-Rubin statistic. We initialized the MCMC chain on the best fit parameters and discarded the first 10% of the chain as burn-in. MCMC techniques only produce robust uncertainties when the noise is normally distributed and white, so to account for correlated noise in the $3.6\mu\text{m}$ light curve (described in § 3.4) we fit the wavelet model from Carter & Winn (2009) simultaneously with the other model parameters. We used a Haar wavelet and let the power spectral density of the red noise vary, following Diamond-Lowe et al. (2014). In our final fit, the noise power spectrum $1/f^\gamma$ had $\gamma = 1.1 \pm 0.1$.

3.4. Goodness of Fit

We performed several tests of the quality of the light curve fits. First we predicted the level of scatter in the light curves based on photon noise alone, then compared this value to the root-mean-square (rms) of the fit

residuals. For the spherical harmonics fit to the phase variation, the Spitzer 4.5 μm light curve rms reaches the expected photon noise limit (637 versus 640 ppm). The 3.6 μm light curve has significantly larger rms (767 versus 470 ppm), due to time-correlated red noise (discussed below). The expected photon-limited rms for the WFC3 spectroscopic light curves ranges from 430 - 530 ppm, and the measured rms was typically within 5% of expectations for all spectroscopic channels. For the WFC3 band-integrated white light curve, the rms was slightly larger than predicted (172 versus 122 ppm). There are a number of possible origins for this discrepancy, including imperfect background subtraction, variation in the position of the spatial scan on the detector, loss of flux outside the extraction window, and imperfect modeling of the astrophysical signal or instrument systematics. However, our analysis of the composition and thermal structure of the planet's atmosphere is based on the wavelength-dependent data (where we reach the photon limit), so we do not further explore the small excess noise in the white light curve.

In addition to calculating the fit rms compared to the photon noise, we also tested for the presence of red noise based on whether the rms decreases as expected when the light curve is binned in time. If the noise is white (uncorrelated in time), the residuals are expected to decrease by a factor of \sqrt{N} , where N is the number of points in a bin. Figure 4 shows the binned residuals compared to expectations for white noise. The *HST*/WFC3 and *Spitzer* Channel 2 light curves agree well with expectations, whereas *Spitzer* Channel 1 shows higher noise than expected as bin size increases. This test confirms the presence of time-correlated noise in the Channel 1 light curve that can be seen by eye in the residuals in Figure 5. Both *Spitzer* channels use the same detector, but Channel 1 data are more susceptible to time-correlated noise because the point spread function is more undersampled at shorter wavelengths, making intrapixel sensitivity variations more pronounced.

4. RESULTS

The fitted light curves are shown in Figure 5. This figure shows results from the spherical harmonics model for the thermal phase variation and has instrument systematics removed. Broadly speaking, the phase curves show large dayside planet-to-star flux values, ranging from $0.151 \pm 0.015\%$, $0.446 \pm 0.38\%$, and $0.569 \pm 0.014\%$ in the WFC3 white light curve, and *Spitzer* 3.6, and 4.5 μm bandpasses, respectively. The planet flux changes significantly with orbital phase in all three of the data sets, suggesting a strong gradient from dayside to night-

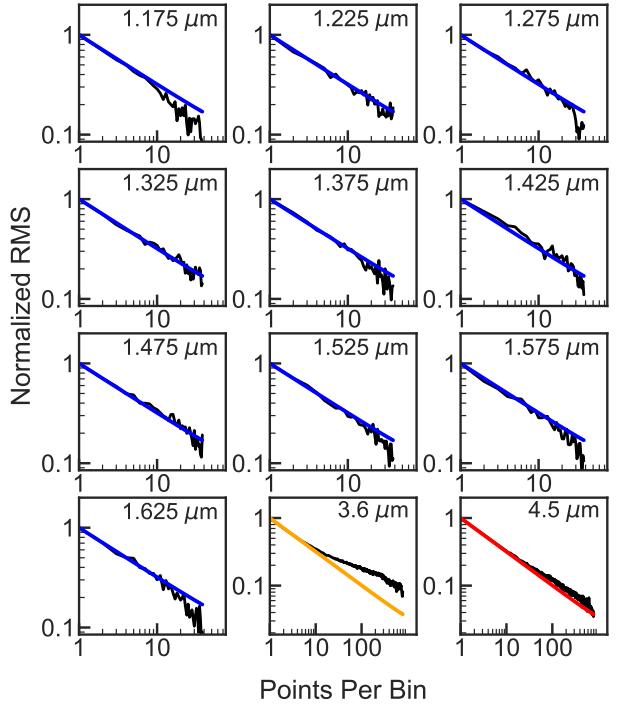


Figure 4. Root mean square (rms) variability in the light curves as a function of bin size (black lines) compared to the expected rms from photon noise (colored lines). The central wavelength of the light curve is indicated in the upper right corner of each panel. With the exception of the *Spitzer* 3.6 μm channel, the rms for the light curves bins down in agreement with predictions from the photon noise.

side temperature, and peak brightness occurs near phase 0.5. In this section, we quantitatively characterize the phase curve shape, split the data into phase resolved spectra, evaluate different temperature maps, compare with previous observations of the dayside thermal emission spectrum, and report the transmission spectrum.

4.1. Phase Curve Amplitudes and Hotspot Offsets

The shape of a phase curve can be summarized with two parameters: the amplitude of thermal phase variation (minimum to maximum brightness, divided by the secondary eclipse depth) and the location of peak brightness (typically called a “hotspot offset” and measured in degrees eastward of the substellar point). Table 2 lists the estimated amplitudes and hotspot offsets (median and 1σ credible interval) for the band-integrated WFC3 phase curve and both *Spitzer* channels. The estimates are from the sine curve model for the thermal phase variation, which fits directly for the amplitude and offset.

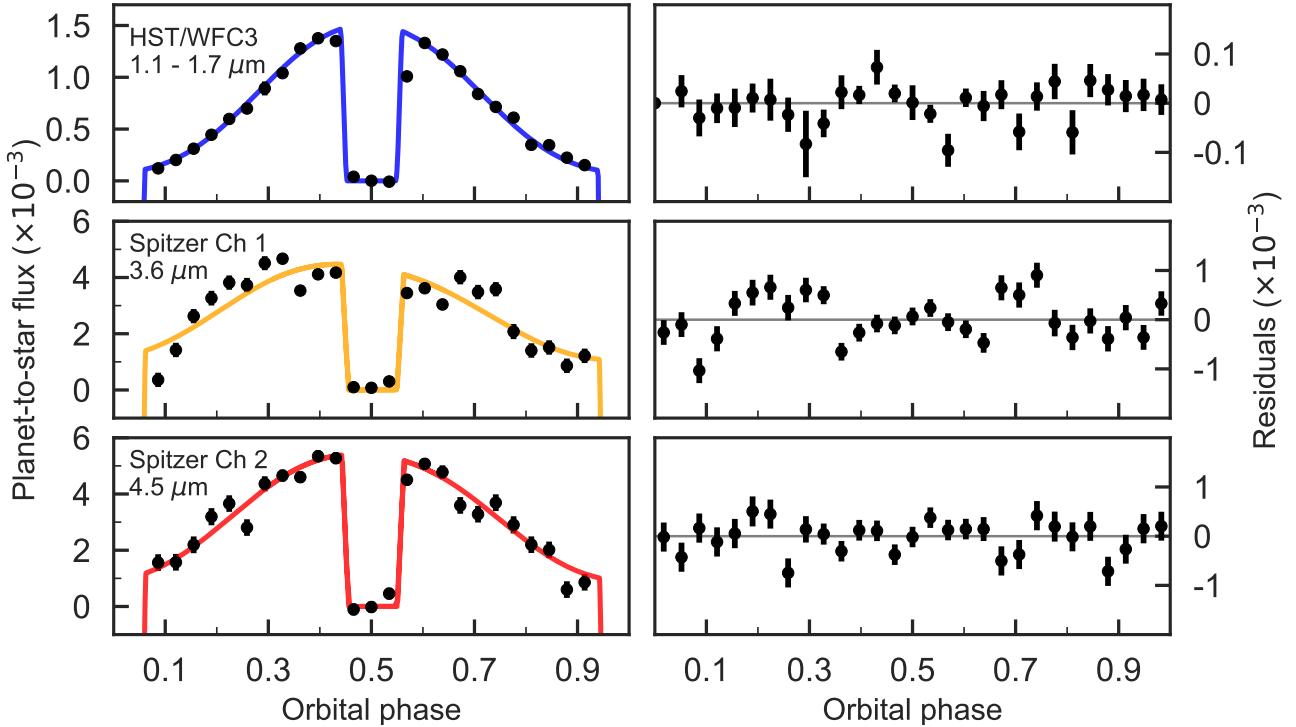


Figure 5. WASP-103b phase curve observations from *HST*/WFC3 (top) and *Spitzer*/IRAC (middle and bottom). For clarity, the data are phase-folded on the planet’s orbital period and binned in 30 uniformly spaced bins between 0 and 1. The left column shows the phase curves with systematic noise removed (black points) compared to the best fit spherical harmonics model (colored lines). The error bars denote 1σ uncertainties (in some cases, the errors are smaller than the data points). We include the transits in the fit, but they are not displayed in this figure. The right-hand column shows the binned residuals for the best-fit curve.

For all three phase curves, the hotspot offset is consistent with zero degrees, which could indicate a small ratio of radiative to advective timescales (the incident flux is reradiated to space faster than it is advected around to the nightside). Fast radiative timescales are predicted at high temperatures, and small hotspot offsets are also observed for other very hot Jupiters (Perez-Becker & Showman 2013; Komacek & Showman 2016; Komacek et al. 2017). The phase curve amplitudes are large (near 0.8 - 0.9), as expected for an atmosphere with inefficient heat redistribution. In § 6 we compare these results to expectations from 3D GCMs.

4.2. Phase-Resolved Spectra

We used the best-fit phase curves (with systematics removed) to generate phase-resolved emission spectra. Since the `spiderman` thermal phase variation models fit the temperature of the planet rather than the eclipse depth directly, we estimated the dayside emission spectrum as follows. We used `spiderman`’s `eclipse_depth` method to calculate the average planet-to-star flux for the best-fit model during secondary eclipse. To esti-

mate uncertainties, we took the standard deviation of the residuals of the in-eclipse data points, then added this value in quadrature to the standard deviation of the residuals of the out-of-eclipse data. This quadrature sum accounts for the uncertainty in the baseline flux. To account for red noise in the *Spitzer* 3.6 μm light curve, we use the approach of Pont et al. (2006) to determine the red noise contribution on the timescale of the eclipse. We add the estimated red noise in quadrature, which increases the uncertainty on planet-to-star flux by a factor of 2.5.

For the other orbital phases, we binned the light curve (with systematics removed) in eight intervals of about 0.1 in orbital phase, with endpoints at phases 0.06, 0.15, 0.25, 0.35, 0.44 and 0.56, 0.65, 0.75, 0.85, 0.94. These endpoints were chosen to ensure that there is no contribution from in-transit or in-eclipse data. In each phase bin, we estimated the planet-to-star flux from the mean value of the data points in the bin. To estimate the uncertainty, we took the standard deviation of the points in the bin and added it in quadrature to the standard

Table 2. Phase Curve Properties

Bandpass	Source	Amplitude	Offset (Degrees)
WFC3	data	0.91 ± 0.02	-0.3 ± 0.1
	nominal GCM	0.89	15.32
	[M/H] = 0.5 GCM	0.84	19.64
	$\tau_{\text{drag}4}$ GCM	0.97	2.34
	$\tau_{\text{drag}3}$ GCM	0.99	0.18
Spitzer 3.6 μm	data	0.86 ± 0.13	2.0 ± 0.7
	nominal GCM	0.78	9.19
	[M/H] = 0.5 GCM	0.72	12.79
	$\tau_{\text{drag}4}$ GCM	0.86	0.90
	$\tau_{\text{drag}3}$ GCM	0.97	0.18
Spitzer 4.5 μm	data	0.83 ± 0.05	1.0 ± 0.4
	nominal GCM	0.79	8.11
	[M/H] = 0.5 GCM	0.73	11.35
	$\tau_{\text{drag}4}$ GCM	0.85	0.90
	$\tau_{\text{drag}3}$ GCM	0.93	0.18

deviation of the data points during secondary eclipse (phase 0.46 – 0.54), to account for the uncertainty in baseline stellar flux. For the 3.6 μm data, we also add red noise on the timescale of a phase bin, following Pont et al. (2006). The phase-resolved emission spectra are shown in Figure 6 and listed in Table 3.

To test that the phase-resolved planet-to-star flux values are robust to different approaches for fitting the phase curves, we compared the estimated planet-to-star flux for all four of the thermal phase variation models (sinusoid, kinematic, spherical harmonics, and two temperature). We found that the choice of model does not significantly change the estimated planet-to-star flux ratios. Since the systematic noise is not strongly correlated with the astrophysical signal, the systematics-divided data are nearly identical for all the models. This point is illustrated in Figure 7 for the broadband WFC3 light curve. For all phase bins, the estimated planet-to-star flux agrees to better than one sigma for all the models.

4.3. Climate

We fit three different models to characterize the planet’s climate: a two-temperature map, the physically-motivated kinematic model of Zhang & Showman

(2017), and a spherical harmonic map. We also fit the thermal phase variation with a sinusoid, which can be inverted to map the climate (Cowan & Agol 2008; Cowan & Fujii 2017). All of the models provide reasonable fits to the data, with χ^2_ν near unity, but they yield significantly different temperature maps. Table 4 lists the best fit minimum and maximum temperatures, as well as the mean day- and nightside temperatures. We also list the information criterion (BIC) values for the fits (a ΔBIC value greater than 10 constitutes strong evidence against a given model; Kass & Raftery 1995).

The spherical harmonics map generally fits the data the best. It has a lower BIC value than all the other models for the broadband WFC3 and Spitzer 3.6 μm phase curves. For the 4.5 μm phase curve, the two temperature map provides the best fit, but it only lowers the BIC value by 2 relative to the spherical harmonics map, which is not a statistically significant improvement (Kass & Raftery 1995). The worst fit model is gene FIXME

The spherical harmonics model yields a mean day-side temperature near 2700 K, whereas the nightside is closer to 1800 K. The other models produce more extreme day-night temperature gradients. These differences arise because the kinematic model allows a steep temperature gradient on the dayside, so the substellar point is much hotter than the terminator, whereas the two temperature map imposes a constant dayside temperature. The intermediate temperature gradient in the spherical harmonics model fits the data the best. However, this model may not be physically realistic: on the nightside, it produces much higher temperatures at the poles than at the equator, contrary to predictions from GCMs (which we discuss in § 6).

4.4. Comparison with Previous Observations

We compared our results to the day-side emission spectrum reported by Cartier et al. (2017), which is based on a Gaussian process analysis of two secondary eclipses from *HST*/WFC3. The shape of their spectrum is consistent with what we find, but their eclipse depths are 125 ppm smaller on average (a difference of about 10%). A likely explanation for this difference is that the Cartier et al. (2017) analysis does not include the planet’s thermal phase variation, so that the Gaussian process models it as an instrument systematic. If the phase variation is absorbed into the systematic model, the measured eclipse depths would be biased low by about 100 ppm (see Figure 7). Our estimated uncertainties are a factor of four smaller than those reported in Cartier et al. (2017), which is consistent with photon-limited expectations (our data set includes two additional eclipses, a

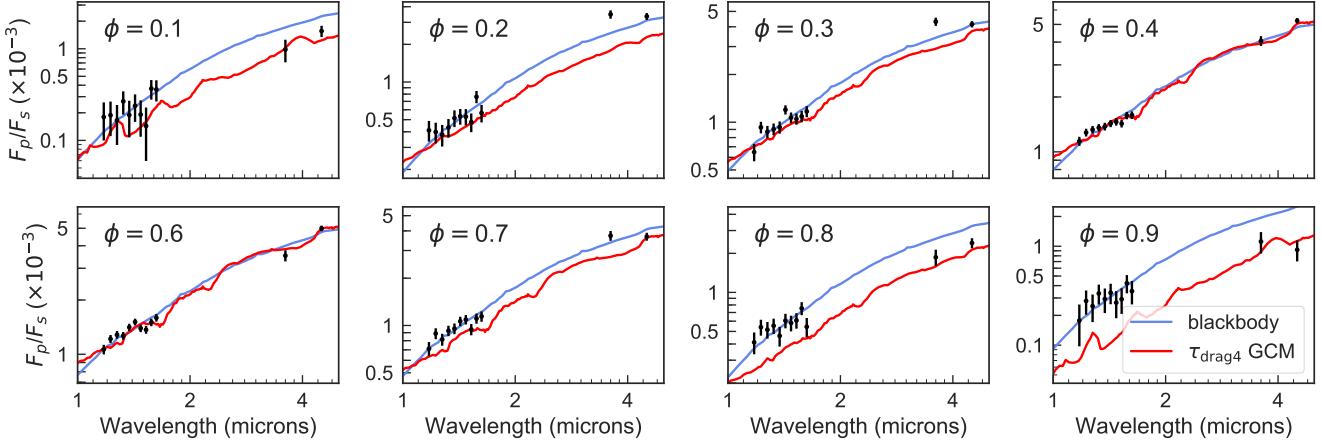


Figure 6. Phase-resolved planet-to-star flux ratios (points) compared to the best fit blackbody (blue line) and the GCM with τ_{drag4} . Phases $\phi = 0.5$ and 0.0 correspond to times of secondary eclipse and transit, when the substellar and antistellar points are respectively aimed at Earth. The blackbody model is fit to the *HST/WFC3* data only.

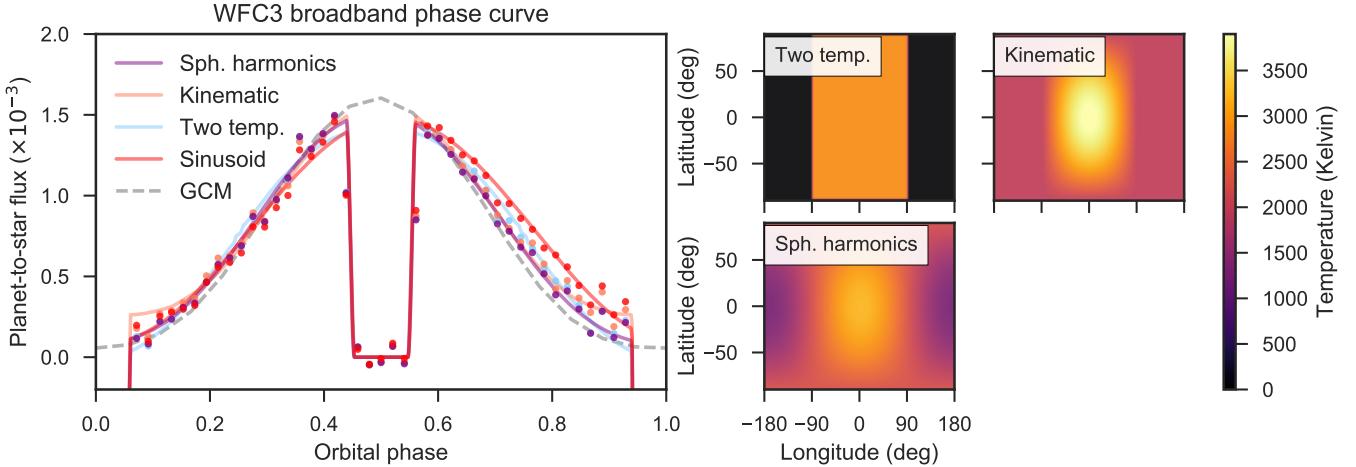


Figure 7. Left: Fits to the broadband WFC3 phase curve compared to a GCM. The colored lines correspond to different temperature maps fit to the data, and the dashed gray line is from the τ_{drag4} GCM. We also show the measured planet-to-star flux for each map (points), which is model-dependent due to slight degeneracies with the instrument systematic model. **Right:** Temperature maps from the best fit models and the GCM at a pressure of 0.01 bar.

factor of five longer out-of-eclipse baseline, and has 60% larger wavelength bins).

We also compared our dayside spectrum to the z' and K_S -band secondary eclipse depths reported in Delrez et al. (2018). The z' ($0.9 \mu\text{m}$) eclipse is 1.0σ lower than our best fit blackbody spectrum (described in § 5), and the K_S ($2.1 \mu\text{m}$) measurement is higher than the model by 2.5σ . Since these results are consistent with (but less precise than) the WFC3 data, we do not include them in our analysis of the atmospheric composition, but we encourage additional measurements in the K_S band to confidently determine whether an emission feature is present at those wavelengths.

4.5. Transmission Spectrum

Each phase curve observation includes a transit of WASP-103b. To measure the wavelength-dependent transit depths (the transmission spectrum), we select a subset of each phase curve. The subset includes the transit and additional baseline on either side, such that the total light curve has twice the duration of the transit. Over this short duration, there is negligible curvature in the light curve due to the planet’s thermal phase variation. We fit the data with a transit model, which has free parameters for the planet-to-star radius and a linear limb darkening parameter. The orbital parameters (inclination, a/R_s , period, and time of central transit)

Table 3. Phase-Resolved Emission Spectra¹

λ	Dilution	$\phi = 0.1$	$\phi = 0.2$	$\phi = 0.3$	$\phi = 0.4$	$\phi = 0.5$	$\phi = 0.6$	$\phi = 0.7$	$\phi = 0.8$	$\phi = 0.9$
1.175	0.10	179 ± 79	411 ± 77	647 ± 80	1143 ± 65	1259 ± 47	1063 ± 64	710 ± 73	412 ± 78	177 ± 79
1.225	0.11	188 ± 76	398 ± 74	928 ± 77	1276 ± 62	1448 ± 46	1216 ± 62	888 ± 71	539 ± 75	280 ± 76
1.275	0.11	166 ± 76	379 ± 74	869 ± 77	1323 ± 62	1480 ± 46	1282 ± 62	814 ± 71	515 ± 75	247 ± 76
1.325	0.11	266 ± 75	432 ± 73	904 ± 76	1357 ± 62	1498 ± 45	1267 ± 61	925 ± 70	552 ± 74	333 ± 75
1.375	0.12	189 ± 81	514 ± 78	928 ± 82	1376 ± 66	1611 ± 48	1411 ± 65	954 ± 75	461 ± 79	292 ± 81
1.425	0.13	238 ± 79	532 ± 76	1198 ± 79	1431 ± 64	1718 ± 47	1511 ± 64	1063 ± 73	605 ± 77	338 ± 79
1.475	0.14	191 ± 81	527 ± 79	1068 ± 82	1460 ± 66	1667 ± 48	1392 ± 66	1090 ± 75	580 ± 80	268 ± 81
1.525	0.14	143 ± 84	478 ± 81	1048 ± 85	1429 ± 69	1623 ± 50	1367 ± 68	943 ± 77	607 ± 82	291 ± 84
1.575	0.15	367 ± 88	761 ± 85	1088 ± 89	1581 ± 72	1749 ± 52	1503 ± 71	1107 ± 81	754 ± 86	422 ± 88
1.625	0.16	359 ± 93	565 ± 90	1169 ± 94	1590 ± 76	1843 ± 56	1593 ± 75	1142 ± 86	542 ± 91	351 ± 93
3.6	0.17	1148 ± 333	3633 ± 330	4487 ± 320	4107 ± 314	4458 ± 383	3495 ± 315	3720 ± 330	2102 ± 331	1431 ± 334
4.5	0.16	1586 ± 220	3468 ± 214	4188 ± 190	5334 ± 179	5686 ± 138	4904 ± 183	3626 ± 214	2575 ± 217	1057 ± 221

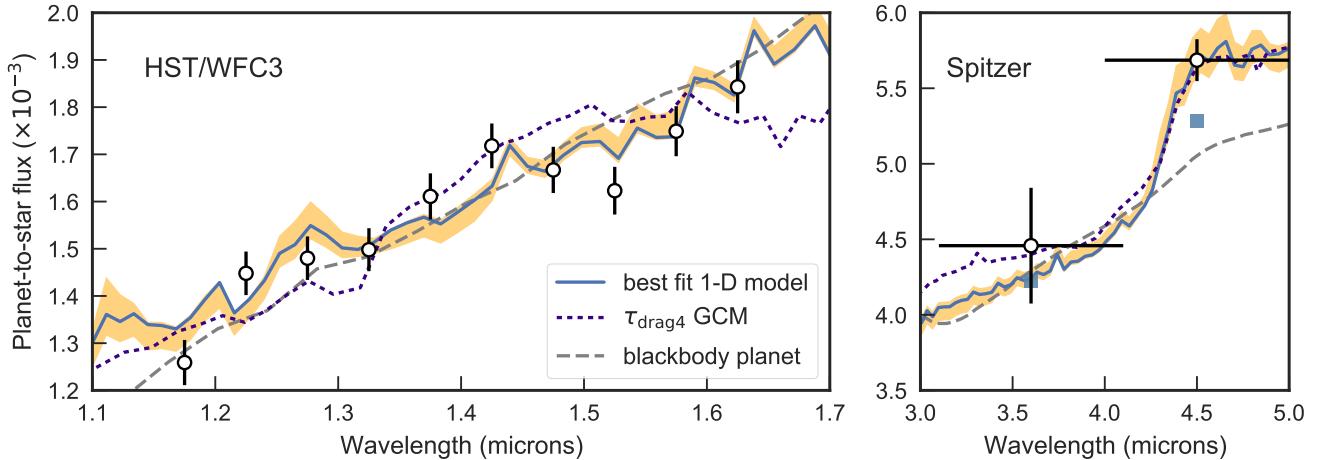
¹The planet-to-star flux in each phase bin ϕ is in units of ppm.

Figure 8. Dayside planet-to-star flux (points) compared to a blackbody model for the planet (dashed gray line) and the best fit 1D model (blue line, with 1σ uncertainty shaded in orange) for the HST/WFC3 data (left) and Spitzer (right). The 1D model was fit to all the data simultaneously, but the blackbody was fit to the WFC3 data only. The best fit 1D model binned at the resolution of the Spitzer data is indicated by blue squares (right).

were fixed at previously published values listed in § 3. We fit for the instrument systematics using the same model as for the phase curve fits, except that we modeled the visit-long systematic as a linear trend in time (which is sufficient for the shorter duration). The advantage of fitting the transits separately from the full phase curves is that the resultant transit depths are not dependent on how the phase variation is modeled.

We show the measured transmission spectrum in Figure 9. The spectrum is biased by flux contamination from the planet's nightside; to correct for nightside emission, we estimate the average nightside planet-to-star flux ratio and subtract it from the transit depth. We calculate planet-to-star flux using a NextGen spectrum for the star (interpolated to $T_{\text{eff}} = 6110$ K), a blackbody for the planet, and a planet-to-star radius ratio of 0.1146. We assume a nightside temperature of 1700

Table 4. Model Comparison

Data	Model	T_{\min}	T_{\max}	\bar{T}_{night}	\bar{T}_{day}	Δ_{BIC}
WFC3	Sph. Harmonics	1227	3237	1822	2636	0
	Kinematic	1977	3953	1977	2769	14
	Two Temp.	0	2879	0	2879	42
	Sinusoid	–	–	–	–	17
Ch 1	Sph. Harmonics	1269	3391	1912	2741	0
	Kinematic	1932	3630	1975	2614	34
	Two Temp.	1418	2990	1418	2990	11
	Sinusoid	–	–	–	–	25
Ch 2	Sph. Harmonics	888	3714	1729	2864	2
	Kinematic	1614	3931	1621	2544	15
	Two Temp.	1344	3241	1344	3241	0
	Sinusoid	–	–	–	–	22

K, which is near the median of the nightside temperatures estimated from the phase variation models (Table 4). We also show the uncorrected transit depths. The corrected and uncorrected transit depths are listed in Table 5. For the uncorrected data, there is an offset between the *HST* and *Spitzer* data of more than five atmospheric scale heights. The uncorrected spectrum is inconsistent with a flat line at 3.2σ confidence, whereas the nightside-corrected spectrum is consistent within 1σ . The corrected spectrum is also consistent with predictions from the τ_{drag4} GCM, which shows water features in the WFC3 bandpass. Future observations with higher precision could reveal these features.

5. ATMOSPHERIC COMPOSITION AND THERMAL STRUCTURE

We characterized the planet’s atmospheric composition by fitting 1D models to the phase-resolved emission spectra. First, we modeled the planet flux as a simple blackbody to estimate the dayside brightness temperature and test for significant absorption or emission features. We then performed a more sophisticated grid-based retrieval to estimate the atmospheric metallicity, carbon-to-oxygen ratio, and thermal structure.

5.1. Blackbody Fits

A blackbody is the simplest model for the planet’s thermal emission and provides a useful first evaluation of the atmospheric properties. To calculate the best fit blackbody model, we assumed a planet-to-star radius ratio of 0.1146 and used a NextGen stellar spectrum

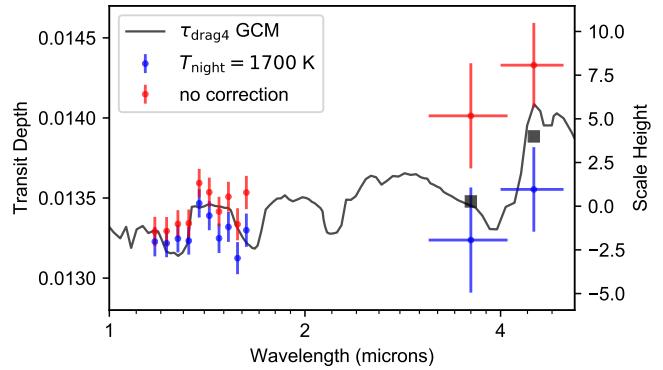


Figure 9. The transmission spectrum of WASP-103b, corrected for planet nightside emission at 1700 K (blue points) and uncorrected (red points). The dark gray line is the model transmission spectrum from the τ_{drag4} GCM. Atmospheric scale height H is shown on the right y-axis, where $H = 5.5 \times 10^6$ m (assuming a mean molecular weight of 2.3 atomic mass units, surface gravity $g = 15.9$ m/s 2 , and a temperature $T = 2410$ K).

Table 5. WASP-103b Transmission Spectrum

Wavelength (micron)	Transit Depth (%) ($T_{\text{night}} = 0$ K)	Transit Depth (%) ($T_{\text{night}} = 1700$ K)
1.175	1.329049 ± 0.000092	0.013228 ± 0.000092
1.225	1.329408 ± 0.000087	0.013217 ± 0.000087
1.275	1.333837 ± 0.000086	0.013246 ± 0.000086
1.325	1.334184 ± 0.000087	0.013233 ± 0.000087
1.375	1.359366 ± 0.000089	0.013467 ± 0.000089
1.425	1.353629 ± 0.000091	0.013390 ± 0.000091
1.475	1.341570 ± 0.000092	0.013249 ± 0.000092
1.525	1.350816 ± 0.000093	0.013319 ± 0.000093
1.575	1.333617 ± 0.000100	0.013125 ± 0.000100
1.625	1.353438 ± 0.000104	0.013299 ± 0.000104
3.6	1.401322 ± 0.000328	0.013238 ± 0.000328
4.5	1.432949 ± 0.000264	0.013554 ± 0.000264

interpolated to an effective temperature of 6110 K (Allard et al. 2012). The best-fit blackbodies are shown in Figures 6 and 8. This model performs well for some orbital phases, with χ_{ν}^2 ranging from 0.9 to 1.1 for phases $\phi = 0.1, 0.7$ and 0.8 . At other phases the blackbody model is a poorer fit, particularly to the *Spitzer* data. For phases 0.2 – 0.6 , the *Spitzer* eclipse depths are signif-

Table 6. Phase-resolved Brightness Temperatures

Orbital Phase	T_b	T_b	T_b	χ^2_ν	¹
	WFC3	Ch. 1	Ch. 2	(9 dof)	
0.06 – 0.15	1917 ± 43	1674 ± 177	1677 ± 102	0.7	
0.15 – 0.25	2254 ± 25	2821 ± 136	2486 ± 87	0.8	
0.25 – 0.35	2650 ± 17	3170 ± 128	2772 ± 74	1.4	
0.35 – 0.44	2905 ± 11	3016 ± 128	3216 ± 68	1.1	
0.44 – 0.56	3011 ± 8	3158 ± 153	3350 ± 53	2.4	
0.56 – 0.65	2884 ± 11	2764 ± 133	3051 ± 70	1.7	
0.65 – 0.75	2634 ± 16	2857 ± 135	2550 ± 85	1.1	
0.75 – 0.85	2312 ± 23	2153 ± 151	2117 ± 91	1.2	
0.85 – 0.94	2025 ± 34	1825 ± 167	1414 ± 113	0.5	

¹ χ^2_ν values are for the fits to the WFC3 data only.

icantly larger than predicted for the best fit blackbody, suggesting the presence of an emission feature (due to CO, as we discuss in § 5.2). On the nightside (phase 0.9), the *Spitzer* 4.5 μm eclipse depth is *smaller* than predicted, suggesting an absorption feature. These differences indicate changes in thermal structure with longitude in the atmosphere.

We also fit blackbody models to the WFC3 data only, and found that they provide a good fit at all orbital phases. The χ^2_ν values range from 0.5 to 2.4, with a median of 0.9. The best fit brightness temperatures range from 1917 ± 43 Kelvin on the nightside to 3011 ± 8 Kelvin on the dayside. We report the best fit temperatures with uncertainties in Table 6. The dayside spectrum is the poorest fit by the blackbody model. This finding is not a surprise because the dayside has higher signal-to-noise than the other orbital phases, thanks to the two secondary eclipse observations from Cartier et al. (2017). The more sophisticated grid-based retrieval (described in the next section) provides a better fit to the dayside.

5.2. Grid-Based Retrieval

To infer abundances from the dayside spectrum, we use a self-consistent grid-based method (ScCHIMERA) similar to that employed in Arcangeli et al. (2018); Mansfield et al. (submitted). We generated a grid from one-dimensional forward models of the atmosphere over a broad range of metallicities (M/H), carbon-to-oxygen ratios (C/O), and stellar irradiation (f). The f parameter is a scaling factor for the stellar flux at the top of the atmosphere, where $f = 1$ corresponds to full heat

redistribution and $f = 2$ is equivalent to only allowing the dayside to re-radiate.

At each point in the grid, we compute forward models to determine self-consistent, radiative-convective temperature-pressure (T-P) profiles. We determine the molecular abundances in each atmospheric layer assuming thermochemical equilibrium (calculated with the NASA CEA routine; Gordon & McBride 1994). We include opacity from the major absorbers expected for a hot Jupiter atmosphere, including H₂O, CO, CO₂, TiO, VO, FeH, and H₂-H₂ CIA. Notably, in contrast to most prior atmospheric retrievals for the hottest planets, we also included opacity from H-, which is an important absorber at temperatures above 2500 K (Arcangeli et al. 2018; Parmentier et al. in prep.). Using these opacities and T-P profiles, we calculated thermal emission spectra over the full grid using the CHIMERA retrieval suite (described in Line et al. 2013, 2014). We then explored the grid with an MCMC chain using the emcee package (Foreman-Mackey et al. 2013) and interpolated in the grid to calculate the likelihood at each model step. The priors were uniform over the ranges $0.2 \leq f \leq 2.8$, $-1 \leq \log \text{M/H} \leq 2.5$, and $-1 \leq \log \text{C/O} \leq 0.95$. We fit this model to the dayside and nightside emission spectra (phases $\phi = 0.5$ and 0.1).

5.2.1. Dayside Spectrum

The main characteristics of the dayside emission spectrum are: (1) it is blackbody-like at WFC3 wavelengths, and (2) in the *Spitzer* bands, the planet-to-star flux is significantly higher than predicted for the best fit blackbody, indicating an emission feature. The best-fit spectrum reproduces these data fairly well, with $\chi^2_\nu = 1.77$ (for 9 degrees of freedom). The largest contribution to the χ^2 value is the 4.5 μm eclipse depth, which is larger than the best fit model prediction by 2.9σ . When the 4.5 μm point is removed, the fit has $\chi^2_\nu = 1.17$ (8 degrees of freedom). The best fit model has a moderately enhanced metallicity (23× solar), carbon-to-oxygen equal to 0.76, poor heat redistribution, and a thermal inversion (temperature increasing with altitude).

Figure 10 shows the opacity contributions of key absorbers for the best fit model. In the optical (which we do not observe directly), there is strong absorption by TiO, VO, and FeH. In the near-infrared, H₂O, H-, and hydrides/oxides all contribute to the opacity, leading to nearly constant opacity over the WFC3 wavelength range. In cooler atmospheres, water is the dominant absorber over this bandpass (e.g. Kreidberg et al. 2014b; Line et al. 2016), but in WASP-103b, H₂O is partially dissociated in the photosphere, leading to a drop in abundance by a factor of ∼ 10 (see Figure 11). Water

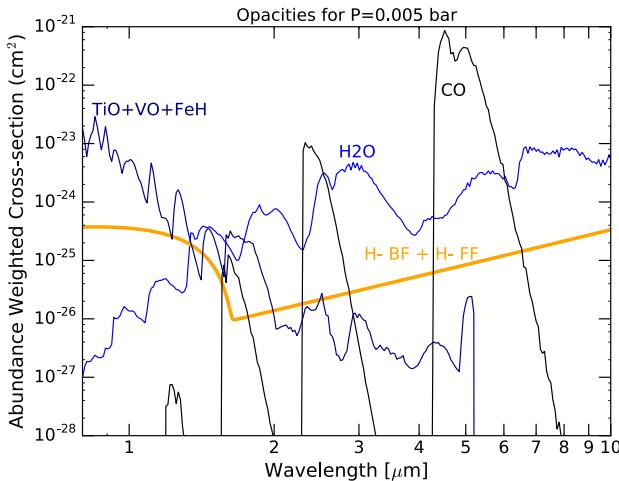


Figure 10. Abundance weighted absorption cross-sections illustrating the important opacity sources at the photospheric pressure and temperature (5 mbar, 3036 K). The strong CO feature at $4.5\mu\text{m}$ contributes to the high planet-to-star flux at that wavelength. Water, hydrides/oxides, and the H- bound-free opacities all play a role in shaping the *HST/WFC3* spectrum.

also has intrinsically weaker features at high temperature (e.g. Tinetti et al. 2012). On top of this, there is significant H- opacity from single H atoms bound with free electrons, which fills in the opacity at wavelengths shorter than $1.5\mu\text{m}$. Finally, the sharp vertical gradient in water abundance results in water becoming optically thick over a very narrow range in pressure, where temperature is nearly constant. Taken together, all these factors add up to produce a nearly featureless spectrum from $1.1 - 1.7\mu\text{m}$. Finally, in the infrared the dominant absorber is CO, which produces the emission feature at *Spitzer* wavelengths.

Figure 11 shows a summary of the temperature-pressure profile and abundances for the best fit model. The T-P profile is inverted, with temperature increasing from 2800 to 3500 K over the pressure range $10^{-2} - 10^{-3}$ bar. The thermal inversion is probably driven by absorption of optical light by oxides and hydrides in the upper atmosphere and the absence of cooling by water molecules (which have dissociated). The observations are sensitive to pressures of $\sim 0.01 - 0.001$ bar, which spans the tropopause, where temperature begins to increase and the water abundance drops by more than an order of magnitude.

In Figure 12, we show the posterior distributions from the grid retrieval. We infer a range in metallicity of $23^{+29}_{-13}\times$ solar, somewhat higher than expected based on Jupiter's metal enrichment ($3 - 5\times$ solar; Wong et al. 2004) and the trend toward decreasing metallicity with

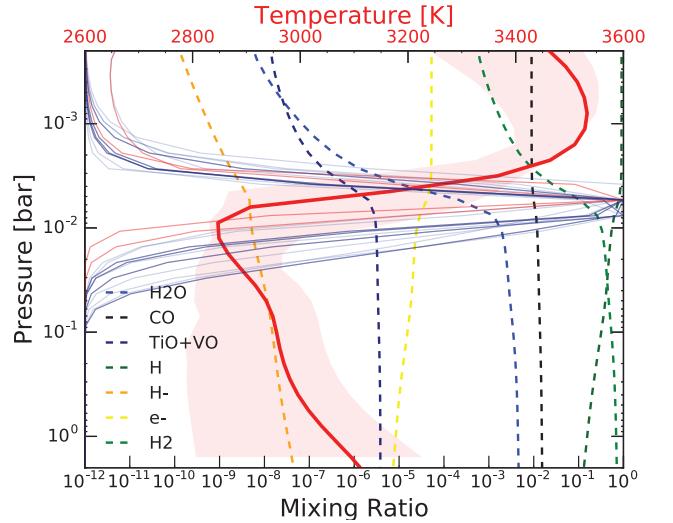


Figure 11. Summary of the 1D self-consistent model atmosphere fits to the dayside emission spectrum. The temperature-pressure profile (top axis) is indicated by the 1-sigma spread of 500 randomly drawn T-P profiles from the posterior (light red) and a representative fit, with $f=0.4$, $[\text{M}/\text{H}]=1.5$, and $\text{C}/\text{O}=0.7$ (dark red). The normalized thermal emission contribution functions for the *Spitzer* points are shown in solid red, the WFC3 in-water band ($1.33 - 1.48\mu\text{m}$) in dark blue, and WFC3 out-of-water-band in light blue. The observations probe between ~ 0.01 and 0.001 bar, just above the tropopause region of the atmosphere where the temperature is increasing. The dashed curves are thermochemical equilibrium mixing ratios for important absorbers computed along the representative fit's self-consistent T-P profile. Note the rapid dissociation of water above the ~ 10 mbar level where the inversion begins.

increasing planet mass observed for the Solar System and exoplanets (e.g. Kreidberg et al. 2014b). However, planet population synthesis models predict some scatter in atmospheric metallicity. Planets near WASP-103b's mass ($1.5 M_{\text{Jup}}$) are expected to have metallicities ranging from roughly $1 - 10\times$ solar (Fortney et al. 2013; Mordasini et al. 2016). Our result for WASP-103b lies on the upper end of this range, and may be indicative of intrinsic scatter in the mass-atmospheric metallicity relation.

The retrieved C/O is consistent with solar, with a 1σ confidence interval of $0.54 - 0.85$. We infer an upper limit on C/O of 0.9 at 3σ confidence, driven by the fact that the atmospheric chemistry is expected to change dramatically when C/O exceeds unity. For a carbon-rich composition, the equilibrium abundance of methane relative to CO increases by orders of magnitude compared to an oxygen-rich composition (e.g. Madhusudhan et al. 2011). Our *Spitzer* eclipse depths are sensitive to the relative abundance of these species, so we can confidently

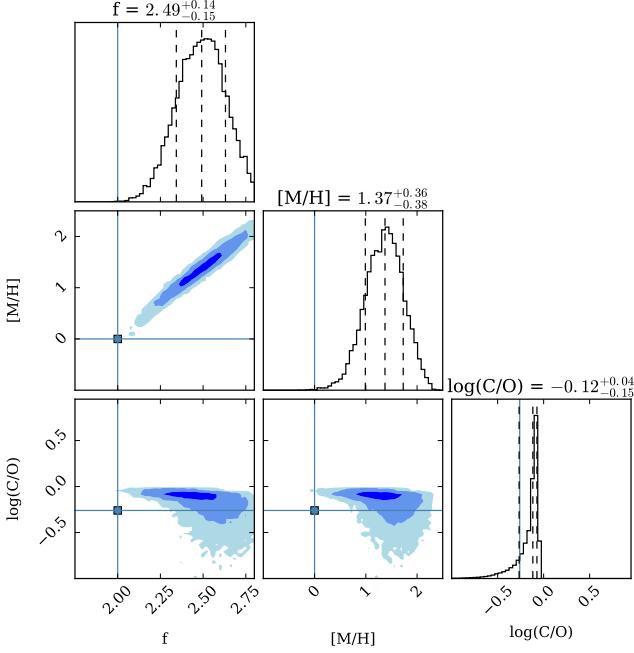


Figure 12. Posterior distributions for WASP-103b’s atmospheric heat redistribution, metallicity, and C/O, from a grid-based fit to the dayside emission spectrum. The histograms on the diagonal show the marginalized distribution of each parameter, with dashed lines indicating the median and surrounding 68% credible interval. The blue lines correspond to solar metallicity (1) and C/O (0.54). The 2D histograms mark the 1, 2, and 3 σ credible regions in dark, medium, and light blue, respectively.

rule out a carbon-rich composition despite the lack of spectrally resolved features (assuming the atmosphere is in chemical equilibrium).

We infer a heat redistribution $f = 2.49_{-0.15}^{+0.14}$. An f parameter of unity represents isotropic heat distribution, whereas $f = 2$ corresponds to dayside emission only. We estimate $f > 2$, indicating a thermal inversion and likely inefficient transport of heat to the nightside. The heat redistribution is strongly correlated with atmospheric metallicity because increasing metallicity shifts the T-P profile to lower pressures, resulting in hotter temperatures at a given pressure level (equivalent to less efficient heat redistribution).

We note that an important caveat for our results is that the best fit model is not a perfect fit to the data (with $\chi^2_\nu = 1.77$), so the uncertainties produced by the MCMC may be underestimated. Furthermore, the inferred C/O and metallicity are highly sensitive to the planet-to-star flux at *Spitzer* 4.5 μm , which is the worst fit data point. To fit this data point, the model favors super-solar metallicities and C/O, which drive up the CO abundance (the dominant absorber at 4.5 μm).

Given that the result hinges on a single photometric band and there are no spectrally resolved molecular features to constrain the temperature-pressure profile or which absorbers are present, we caution against over-interpreting these results until wider spectral coverage is available.

5.2.2. Nightside Spectrum

We also fit the nightside spectrum (phase 0.1) with the grid-based retrieval. The best fit spectrum has a non-inverted temperature pressure profile. At 1 σ confidence, the metallicity is $15 - 240 \times$ solar and the C/O is unbounded over the full prior range. The atmospheric composition is consistent with results from the dayside spectrum.

This agreement is an encouraging sanity check; however, there are several model assumptions that may result in artificially tight constraints on the atmospheric properties on the nightside. One challenge in modeling the nightside spectrum is that the physical processes shaping the T-P profile are unknown. Our model assumes a scaled stellar irradiation at the top of the atmosphere, but in reality, the heat source is advection from the dayside. Another caveat is that the model is not self-consistent: the energy leaving the dayside is not constrained to equal the energy entering the nightside. Further work is needed to develop a fully self-consistent 2-D retrieval method.

6. COMPARISON WITH GCMS

To explore the three-dimensional effects of atmospheric dynamics, we ran several GCMs to compare with the measured phase-resolved spectra. (**Vivien:** add text describing GCM). Our nominal GCM was a cloud-free, solar composition atmosphere with TiO/VO opacity and no added drag. We also ran models with enhanced metallicity ($[M/H] = 0.5$), no TiO/VO, and added atmospheric drag. Lorentz drag can occur in the hottest exoplanet atmospheres, which may be partially ionized and coupled to the background magnetic field. We included Rayleigh drag forces in the GCM, which are proportional to the wind speed scaled by a time constant ($F_{\text{drag}} \propto -v/\tau_{\text{drag}}$; Menou & Rauscher 2009). We assumed the drag timescale is uniform throughout the atmosphere and considered two cases: $\tau_{\text{drag}} = 10^3$ and 10^4 s, which we label $\tau_{\text{drag}3}$ and $\tau_{\text{drag}4}$, respectively.

To assess how well the GCM predictions reproduce the data, we calculated the amplitude and hot spot offset for all the models (listed in Table 2). The observed hotspot offsets (0.3 – 2.0 degrees) are best reproduced by the $\tau_{\text{drag}3}$ and $\tau_{\text{drag}4}$ models, which have small offsets due to changes in the wind pattern. In the drag

models, the winds shift from a substellar to an antistar flow rather than an equatorial jet. The nominal and metal-enhanced GCMs both have offsets that are too large (10–20 degrees). On the other hand, the phase curve amplitudes are better matched by the nominal GCM than either of the drag models. The most precise observed amplitude is 0.91 ± 0.02 (in the WFC3 bandpass), which is significantly smaller than the amplitudes predicted by the drag models (0.97, 0.99). As shown in Figure 14, this large amplitude causes the drag models to underestimate the planet-to-star flux away from the substellar point. (FIXME: this is all comparison to WFC3 phase curve; add comparison to Spitzer too). There are several physical processes not included in the GCM that could contribute to a hotter nightside, including shocks, longitude-dependent drag, latent heat released from H₂ recombination.

To check that fast drag timescales of order $10^3 - 10^4$ s are physically realistic, we computed the free electron abundance due to ionized metals for the $\tau_{\text{drag}4}$ GCM. Using the order-of-magnitude expression from Perna et al. (2010) (Equation 12), we estimate a drag timescale at the substellar point of $\tau_{\text{drag}} = 2 \times 10^4 / B^2$ s, where B is the magnetic field in Gauss. We assume a temperature and pressure of 3359 K and 0.11 bar, and that the magnetic field is perpendicular to the flow. To reach a drag timescale of 10^4 s, a magnetic field stronger than ~ 1 Gauss is required. This value is comparable to Jupiter's magnetic field strength of 5–10 Gauss (Bagenal et al. 2004).

We compared the TP profiles from the GCM to cloud condensation curves and the best fit radiative-convective equilibrium models from the 1D retrieval (Figure 13).

We also compared the GCM output to temperature maps retrieved with `spiderman`. Figure 7 shows the 0.01 bar temperature map for the $\tau_{\text{drag}4}$ GCM compared to the best fit models. At this pressure, the GCM has minimum and maximum temperatures of 580 and 3520 K. The temperature gradient from the dayside to the terminator is intermediate between the kinematic and spherical harmonics models. The GCM predicts a cooler nightside than all models except the two-temperature model. We note that the GCM does not provide a perfect fit to the data, whereas multiple `spiderman` maps fit quite well, with large differences in temperature between them (see Table 4). Determining which models (if any) accurately describe the climate will require higher precision phase curves and/or eclipse mapping (e.g. de Wit et al. 2012).

The GCMs also provide insight into what molecules are present in which parts of the atmosphere. As discussed in § 5, water dissociation and H- opacity are

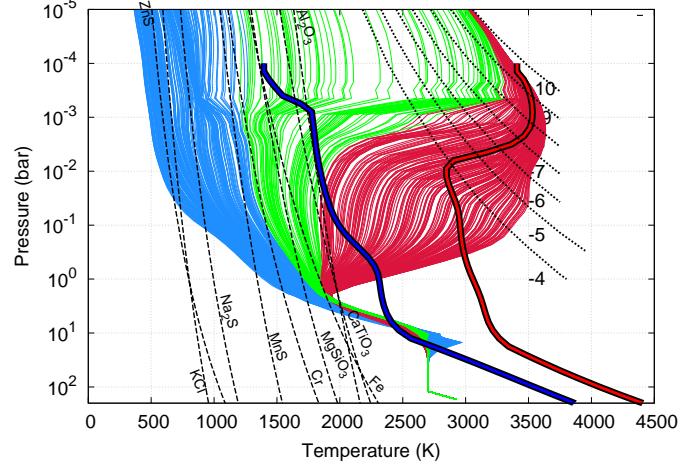


Figure 13. TP profiles from the GCM (thin colored lines) and 1D retrievals (thick lines), compared to condensation curves of potential cloud species. The blue, green, and red lines correspond to TP profiles randomly drawn from the FIXME GCM at orbital phases FIXME. The thick red and blue lines are the best fit 1D TP profiles for the dayside ($\phi = 0.5$) and nightside ($\phi = 0.1$), respectively.

needed to explain the dayside emission spectrum. Figure 14 shows the photospheric abundances of H₂O and H- compared to the predicted temperature for the $\tau_{\text{drag}4}$ GCM. The water abundance drops by ~ 10 at the substellar point, and the H- opacity increases by ~ 100 . By contrast, CO remains intact throughout the atmosphere. Our observations are not precise enough to detect water features on the nightside of the planet (see § 7), but future high precision data may be sensitive to these features, and will help constrain the strength of horizontal transport in the atmosphere (Agúndez et al. 2014).

7. COMPARISON WITH BROWN DWARFS AND DIRECTLY IMAGED COMPANIONS

WASP-103b is so highly irradiated that its photospheric temperature (2000 – 3000 K) is comparable to that of low mass stars. However, the planet's other properties (surface gravity, rotation rate, irradiation) are different. To explore the effects of varying these parameters, we selected spectra from WASP-103b at three orbital phases: dayside ($\phi = 0.5$), quadrature ($\phi = 0.25$), and nightside ($\phi = 0.1$), and compared them to brown dwarfs and young directly imaged companions with comparable brightness temperatures.

We also used three brown dwarfs/low mass stars for comparison. We chose the field sources: 2MASS J13204427+0409045, (1320+0409) an optical L3, 2MASS J04285096-2253227 (0428-2253) an optical L0.5 and 2MASS J00034227-2822410 (0003-2822) an optical M8 (see Filippazzo et al. 2015). We then used all currently

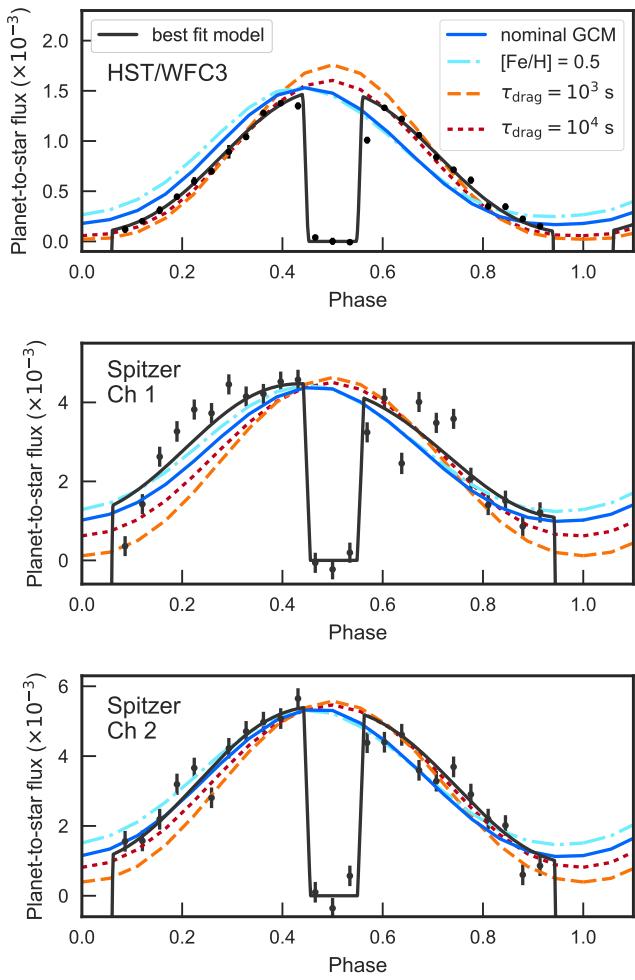


Figure 14. GCM predictions (colored lines) compared to the best fit model for the WFC3 white light, Spitzer 3.6, and Spitzer 4.5 micron phase curves (black lines, top to bottom). The nominal model is solar composition, cloud-free, with TiO/VO opacity and no drag. The models are corrected for the predicted ellipsoidal variability of the planet.

available photometric, astrometric, and spectroscopic data for each source to evaluate fundamental parameters such as mass, T_{eff} and $\log g$ (Filippazzo et al. 2015; Faherty et al. 2016) and create flux-calibrated spectral energy distributions. For 1320+0409 we used SDSS, WISE and 2MASS photometry along with the optical spectrum from Reid et al. (2008), the near infrared spectrum from Bardalez Gagliuffi et al. (2014), and the parallax reported in Faherty et al. (2012). For 0428-2253 we used 2MASS, DENIS and WISE photometry along with the optical spectrum from Kendall et al. (2003) the near infrared spectrum from Bardalez Gagliuffi et al. (2014) and the parallax reported in Dierich et al. (2014), and for 0003-2822 we used 2MASS

and WISE photometry along with the optical spectrum from Cruz et al. (2007) the near infrared spectrum from Cruz et al. (2018), and the parallax reported in Faherty et al. (2010). Coincidentally, both 1320+0409 and 0003-2822 are widely separated (> 2000 AU) companions to K2 and G8 stars respectively. All data was gathered from the Brown Dwarfs in New York City (BDNYC) database (Filippazzo et al. 2015)². At the assumed field ages of each source, 0003-2822 would be above the nuclear burning boundary (star) while 1320+0409 would be below (brown dwarf). 0428-2253 is likely a star but at a slightly younger field age could be a brown dwarf.

The directly imaged spectra are for the sources CD-35 2722, USco 1610-1913B, and TWA 22A and are taken from Wahhaj et al. (2011); Aller et al. (2013); Bonnefoy et al. (2014). They are calibrated in absolute flux using published H-band photometry (Wahhaj et al. 2011; Aller et al. 2013; Bonnefoy et al. 2009) and distances (Gaia Collaboration et al. 2016; Teixeira et al. 2009), a flux-calibrated spectrum of Vega (Mountain et al. 1985; Hayes 1985), and the corresponding filter passbands.

The system properties for all the objects are summarized in Table 7. Figure 16 shows the flux-calibrated spectra (assuming a distance of 10 pc). We compared the spectra over the wavelength range $1.1 - 1.7 \mu\text{m}$. The dominant spectral features over this range are expected to come from water, which has a forest of absorption lines near $1.4 \mu\text{m}$. Spectra for the brown dwarfs and imaged companions have noticeable features in the water band, whereas WASP-103b does not. To quantitatively compare the water feature amplitude for different objects, we define an amplitude $A = (F_{1,2} - F_{3,4})/F_{1,2}$, where $F_{1,2}$ is the weighted mean flux in a wavelength bin $\lambda_1 - \lambda_2$. We calculated the water feature amplitude for two choices of wavelength bins. For the first, A_1 , we considered data in and out of the water band, with $\lambda_{1,2,3,4} = 1.15, 1.3, 1.35, 1.5 \mu\text{m}$. The ground-based direct imaging data do not span this entire wavelength range, so we also define an amplitude A_2 with $\lambda_{1,2,3,4} = 1.2, 1.35, 1.5, 1.65 \mu\text{m}$. The estimated amplitudes and uncertainties are listed in Table 7.

The A_1 and A_2 values are significantly lower for WASP-103b at dayside and quadrature than for brown dwarfs and imaged companions of similar temperature. WASP-103b typically has A_1 and A_2 consistent with zero, indicating no water absorption (in agreement with the analysis in § 4 that showed water is depleted in the photosphere). By contrast, the brown dwarfs and young companions have significant water features, with drops

² The BDNYC Database: <http://database.bdnyc.org/>

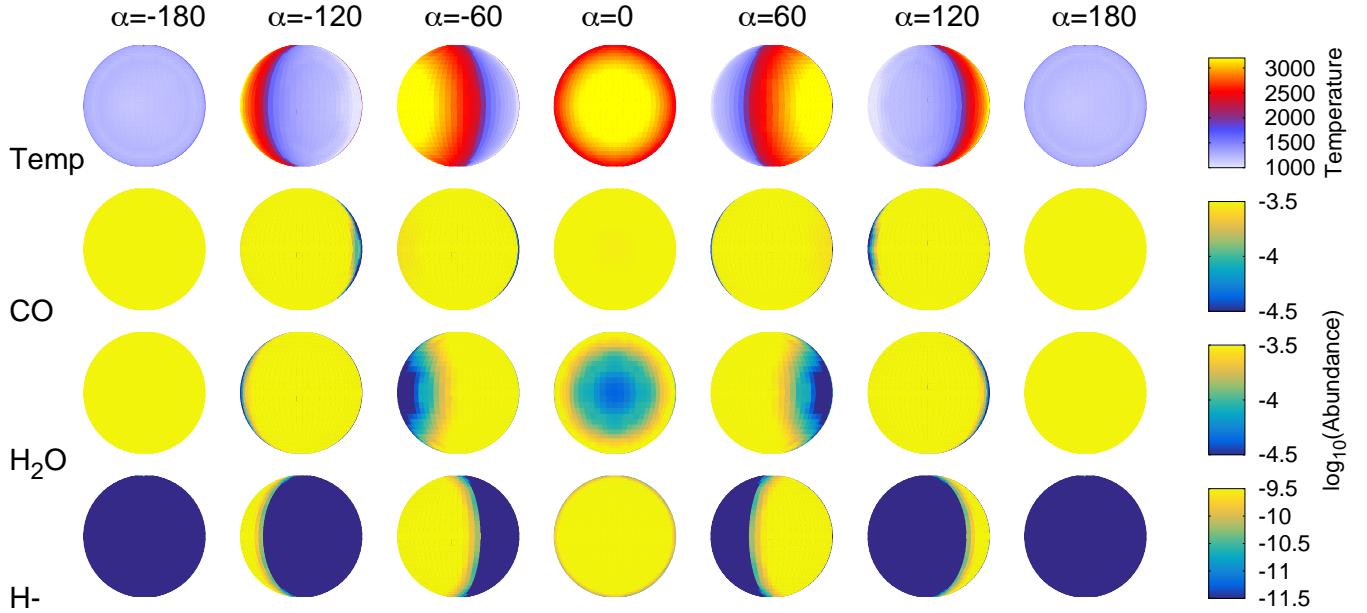


Figure 15. Row (1): photospheric temperatures for the τ_{drag4} GCM for different viewing geometries. The orbital phase $\alpha = 0$ corresponds to secondary eclipse (when the substellar point faces Earth) and $\alpha = 180$ corresponds to transit (when the antistar point faces Earth). FIXME: Row (2) indicates the ratio of pressures probed in and out of the water band. On the dayside, the opacity out of the water band is filled in by absorption from H- and TiO/VO. Rows (3-4) show the abundances of H₂O and H-. At dayside temperatures, water dissociates and the photospheric abundance drops by ~ 10 .

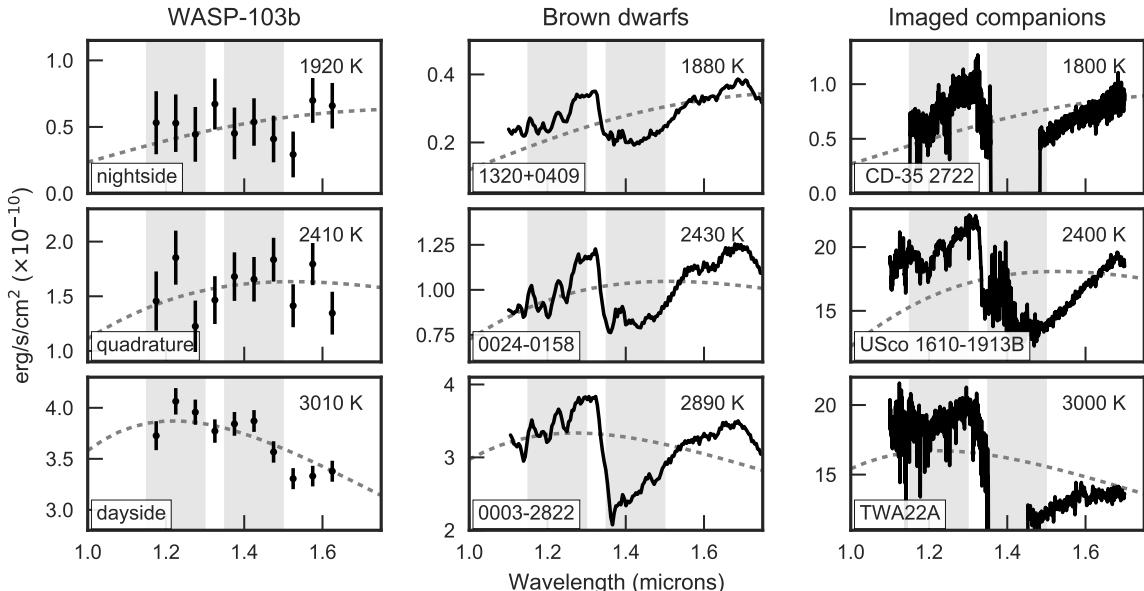


Figure 16. Flux-calibrated spectra for WASP-103b (left column) compared to observed spectra for brown dwarfs (middle) and directly imaged companions (right), assuming a distance of 10 pc. The WASP-103b spectra are from the nightside (phase 0.1, top row), quadrature (phase 0.25, middle) and the dayside (phase 0.5, bottom). Each row shows spectra from objects of comparable temperature. The dotted gray line corresponds to the best fit blackbody. Effective temperatures are listed in the upper right corners. The wavelength bins used to calculate the water feature amplitude A_1 are shaded in gray.

in flux of about 20% in the water band. This is not surprising: stars in the temperature range 2000 – 3000 K are well known to have prominent water features (Kirkpatrick et al. 1993). Based on the grid retrieval of WASP-103b’s atmospheric composition, there are several reasons WASP-103b would exhibit different behavior at the same temperature. WASP-103b is irradiated from above rather than below, changing the shape of the temperature-pressure profile. In addition, WASP-103b also has much lower surface gravity ($\log g = 3.2$ versus 4–5 for stars), which pushes the photosphere to lower pressures, where water dissociates more readily (Arcangeli et al. 2018). These factors are not relevant on the nightside, and 3D models predict that WASP-103b has nightside water absorption features; however, the current data are not precise enough to distinguish between a blackbody spectrum versus water features like those seen in the other objects.

8. SUMMARY

We observed thermal phase curves of the hot Jupiter WASP-103b measured with *HST*/WFC3 time series spectroscopy (1.15 – 1.65 μm) and *Spitzer*/IRAC broadband photometry (3.6 and 4.6 μm bands). Here we summarize our conclusions about the atmosphere based on these measurements.

- The dayside planet-to-star flux is $0.151 \pm 0.015\%$, $0.446 \pm 0.38\%$, and $0.569 \pm 0.014\%$ in the WFC3 bandpass, *Spitzer* 3.6, and *Spitzer* 4.5 μm , respectively. The best fit blackbody to the WFC3 dayside spectrum has a brightness temperature of 3013 ± 8 K, making WASP-103b among the hottest exoplanets ever observed.
- The phase curves have large amplitudes (0.8 – $0.9 \times$ the secondary eclipse depth), and small offsets in peak brightness from the substellar point (consistent with zero degrees at all wavelengths). These characteristics indicate inefficient redistribution of heat to the nightside, as seen in other very hot Jupiters (Komacek et al. 2017).
- We fit the phase variation with the `spiderman` package (Louden & Kreidberg 2017) to evaluate different models of the planet’s climate, including a two-temperature map, a physically-motivated kinematic map, and spherical harmonics. The spherical harmonic temperature map generally provides the best fit to the data; however, all the maps produce reasonable fits (χ^2_ν near unity), and there are large differences in temperature between them (up to 1000 K at a given latitude/longitude).

Breaking the degeneracy between different climate maps will require higher precision phase curves and/or secondary eclipse mapping (e.g. de Wit et al. 2012).

- We calculated phase-resolved spectra in ten orbital phase bins. The *HST*/WFC3 spectra are consistent with blackbody emission from the planet at all orbital phases. The best fit brightness temperatures ranges from 1920 ± 40 K (phase $\phi = 0.1$) to 3010 ± 10 K on the dayside. We attribute the absence of water features at WFC3 wavelengths to (1) H₂O dissociation on the dayside and (2) additional near-IR opacity from H-, TiO/VO and FeH.
- The *Spitzer* data are *not* consistent with a blackbody: there is an infrared emission feature at phases where the substellar point is visible ($\phi = 0.2$ – 0.7), which transitions to an absorption feature on the nightside ($\phi = -0.1$ – 0.1). We attribute these features to CO absorption and changing thermal structure from the dayside to the nightside.
- We characterized the composition with a 1D grid-based retrieval that assumes thermochemical and radiative-convective equilibrium. The atmosphere is moderately metal-enriched ($23_{-13}^{+29} \times$ solar; and $> 1 \times$ solar at 3σ confidence). This value is somewhat higher than what is observed for other gas giants (e.g. Wong et al. 2004; Kreidberg et al. 2014b), but may be indicative of intrinsic scatter in the relationship between atmospheric metallicity and planet mass predicted by theoretical models (Fortney et al. 2013; Mordasini et al. 2016). However, the metallicity is strongly sensitive to the 4.5 μm *Spitzer* eclipse depth, and additional observations would be useful in confirming the metal enhancement. In addition to metallicity, we also infer an upper limit on the carbon-to-oxygen ratio of 0.9 (3σ confidence). This estimate agrees with expectations from planet formation models that pollution from water ice in planetesimals leads to C/O < 1 in gas giant atmospheres (Mordasini et al. 2016; Espinoza et al. 2017). The best fit temperature pressure profile has a thermal inversion from $\sim 10^{-2}$ – 10^{-3} bars due to TiO/VO absorption at high altitudes.
- We ran several 3D GCMs to compare to the data, including a nominal model with a cloud-free, solar composition, a metal-enriched model ([Fe/H] = 0.5), and two models with Lorentz drag. The

Table 7. Source Properties

	Object	T_{eff} (K)	$\log g$ (cgs)	$\text{H}_2\text{O A}_1$	$\text{H}_2\text{O A}_2$
Hot Jupiter	W103b night	1920 ± 40	3.2 ± 0.04	$0.07 \pm 1.8e - 01$	$-0.00 \pm 1.6e - 01$
	W103b quadrature	2410 ± 20	3.2 ± 0.04	$-0.14 \pm 7.8e - 02$	$-0.01 \pm 6.7e - 02$
	W103b dayside	3010 ± 10	3.2 ± 0.04	$0.04 \pm 1.4e - 02$	$0.15 \pm 1.2e - 02$
Brown Dwarf	2MASS J1320+0409	1880 ± 70	5.19 ± 0.16	$0.21 \pm 6.3e - 04$	$-0.06 \pm 5.0e - 04$
	2MASS J0428-2253	2430 ± 80	5.22 ± 0.09	$0.16 \pm 1.2e - 04$	$-0.03 \pm 1.0e - 04$
	2MASS J0003-2822	2890 ± 80	5.18 ± 0.04	$0.26 \pm 1.3e - 04$	$0.10 \pm 1.1e - 04$
Imaged Companion	CD-35 2722 ¹	1800 ± 100	4.5 ± 0.5	—	$0.15 \pm 1.0e - 05$
	USco 1610-1913B ²	2400 ± 150	—	$0.27 \pm 2.0e - 04$	$0.19 \pm 2.2e - 04$
	TWA 22A ³	3000 ± 100	4.5 ± 0.5	—	$0.29 \pm 2.0e - 05$

¹Wahhaj et al. (2011)²Aller et al. (2013)³Bonnefoy et al. (2014)

GCMs with Lorentz drag have different circulation patterns and provide the best match to the small hotspot offsets that are observed, but they underpredict the planet’s nightside flux.

- We compared the spectra of WASP-103b at phases 0.5 (dayside), 0.25 (quadrature), and 0.1 (nightside) to brown dwarfs and directly imaged companions of similar temperature. We quantify the strength of the water feature and find that both brown dwarfs and imaged companions show evidence for water absorption at $1.4 \mu\text{m}$, whereas the WASP-103b dayside and quadrature spectra do not. We attribute the difference to two factors: WASP-103b’s irradiation environment, which changes the temperature pressure profile, and its low surface gravity, which pushes the photosphere to higher altitudes where water dissociates more easily. The WASP-103b nightside spectra have larger uncertainties and are consistent with the water feature amplitudes for other objects; higher precision phase curves are needed to detect water on the nightside.

These results provide a first look at the global composition and thermal structure of WASP-103b. The planet is complex, with changes in temperature profile with longitude and possible gradients in composition from dayside to nightside. These findings highlight the 3D nature of exoplanets and illustrate the importance of phase

curve observations to develop a comprehensive understanding of their atmospheric chemistry and physics.

Support for *HST* program GO-15050 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. Support for *Spitzer* program 11099 was provided by NASA through an award issued by JPL/Caltech. The data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST) and the NASA/IPAC Infrared Science Archive. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. The Infrared Science Archive is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This work also made use of the Python packages SciPy and NumPy (Jones et al. 2001; Van Der Walt et al. 2011). The authors are grateful for helpful conversations with Caroline Morley, Thomas Beatty, Ming Zhao, Kim Star Cartier, Hannah Diamond-Lowe, and Nick Cowan. We also thank the organizers of the 2016 Santa Cruz Kavli Summer Program and the 2017 Ringberg Atmospheres of Disks and Planets meeting for facilitating productive discussion and collaboration. J.L.B. acknowledges support from the David and Lucile Packard Foundation. J.M.D. acknowledges support from the European Research Council (ERC) under the programme Exo-Atmos (grant agreement no. 679633).

REFERENCES

- Agúndez, M., Parmentier, V., Venot, O., Hersant, F., & Selsis, F. 2014, A&A, 564, A73
- Allard, F., Homeier, D., & Freytag, B. 2012, Philosophical Transactions of the Royal Society of London Series A, 370, 2765
- Aller, K. M., Kraus, A. L., Liu, M. C., et al. 2013, ApJ, 773, 63
- Arcangeli, J., Desert, J.-M., Line, M. R., et al. 2018, ArXiv e-prints, arXiv:1801.02489
- Armstrong, D. J., de Mooij, E., Barstow, J., et al. 2016, Nature Astronomy, 1, 0004
- Baggenal, F., Dowling, T. E., & McKinnon, W. B. 2004, Jupiter : the planet, satellites and magnetosphere
- Bardalez Gagliuffi, D. C., Burgasser, A. J., Gelino, C. R., et al. 2014, ApJ, 794, 143
- Beatty, T. G., Madhusudhan, N., Pogge, R., et al. 2017a, AJ, 154, 242
- Beatty, T. G., Madhusudhan, N., Tsiaras, A., et al. 2017b, AJ, 154, 158
- Bonnefoy, M., Chauvin, G., Lagrange, A.-M., et al. 2014, A&A, 562, A127
- Bonnefoy, M., Chauvin, G., Dumas, C., et al. 2009, A&A, 506, 799
- Borucki, W. J., Koch, D., Jenkins, J., et al. 2009, Science, 325, 709
- Braak, C. J. F. T. 2006, Statistics and Computing, 16, 239. <http://dx.doi.org/10.1007/s11222-006-8769-1>
- Carter, J. A., & Winn, J. N. 2009, ApJ, 704, 51
- Cartier, K. M. S., Beatty, T. G., Zhao, M., et al. 2017, AJ, 153, 34
- Cowan, N. B., & Agol, E. 2008, ApJL, 678, L129
- Cowan, N. B., & Fujii, Y. 2017, ArXiv e-prints, arXiv:1704.07832
- Crossfield, I. J. M., Barman, T., Hansen, B. M. S., Tanaka, I., & Kodama, T. 2012, ApJ, 760, 140
- Cruz, K. L., Núñez, A., Burgasser, A. J., et al. 2018, AJ, 155, 34
- Cruz, K. L., Reid, I. N., Kirkpatrick, J. D., et al. 2007, AJ, 133, 439
- Cubillos, P., Harrington, J., Madhusudhan, N., et al. 2013, ApJ, 768, 42
- de Wit, J., Gillon, M., Demory, B.-O., & Seager, S. 2012, A&A, 548, A128
- Delrez, L., Madhusudhan, N., Lendl, M., et al. 2018, MNRAS, 474, 2334
- Demory, B.-O., de Wit, J., Lewis, N., et al. 2013, ApJL, 776, L25
- Diamond-Lowe, H., Stevenson, K. B., Bean, J. L., Line, M. R., & Fortney, J. J. 2014, ApJ, 796, 66
- Dieterich, S. B., Henry, T. J., Jao, W.-C., et al. 2014, AJ, 147, 94
- Espinoza, N., Fortney, J. J., Miguel, Y., Thorngren, D., & Murray-Clay, R. 2017, ApJL, 838, L9
- Evans, T. M., Sing, D. K., Wakeford, H. R., et al. 2016, ApJL, 822, L4
- Faherty, J. K., Burgasser, A. J., West, A. A., et al. 2010, AJ, 139, 176
- Faherty, J. K., Burgasser, A. J., Walter, F. M., et al. 2012, ApJ, 752, 56
- Faherty, J. K., Riedel, A. R., Cruz, K. L., et al. 2016, ApJS, 225, 10
- Filippazzo, J. C., Rice, E. L., Faherty, J., et al. 2015, ApJ, 810, 158
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
- Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, ApJ, 678, 1419
- Fortney, J. J., Mordasini, C., Nettelmann, N., et al. 2013, ApJ, 775, 80
- Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
- Gillon, M., Anderson, D. R., Collier-Cameron, A., et al. 2014, A&A, 562, L3
- Gordon, S., & McBride, B. J. 1994, Computer program for calculation of complex chemical equilibrium compositions and applications, Vol. 1 (National Aeronautics and Space Administration, Office of Management, Scientific and Technical Information Program)
- Hayes, D. S. 1985, in IAU Symposium, Vol. 111, Calibration of Fundamental Stellar Quantities, ed. D. S. Hayes, L. E. Pasinetti, & A. G. D. Philip, 225–249
- Haynes, K., Mandell, A. M., Madhusudhan, N., Deming, D., & Knutson, H. 2015, ApJ, 806, 146
- Henry, G. W. 1999, PASP, 111, 845
- Horne, K. 1986, PASP, 98, 609
- Hu, R., Demory, B.-O., Seager, S., Lewis, N., & Showman, A. P. 2015, ApJ, 802, 51
- Hubeny, I., Burrows, A., & Sudarsky, D. 2003, ApJ, 594, 1011
- Jones, E., Oliphant, T., Peterson, P., & Others. 2001, SciPy: Open source scientific tools for Python, . . <http://www.scipy.org/>
- Kass, R. E., & Raftery, A. E. 1995, Journal of the American Statistical Association, 90, 773. <http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1995.10476572>
- Kataria, T., Sing, D. K., Lewis, N. K., et al. 2016, ApJ, 821, 9

- Kendall, T. R., Mauron, N., Azzopardi, M., & Gigoyan, K. 2003, *A&A*, 403, 929
- Kirkpatrick, J. D., Kelly, D. M., Rieke, G. H., et al. 1993, *ApJ*, 402, 643
- Knutson, H. A., Charbonneau, D., Allen, L. E., et al. 2007, *Nature*, 447, 183
- Komacek, T. D., & Showman, A. P. 2016, *ApJ*, 821, 16
- Komacek, T. D., Showman, A. P., & Tan, X. 2017, *ApJ*, 835, 198
- Kreidberg, L. 2015, *PASP*, 127, 1161
- Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014a, *Nature*, 505, 69
- . 2014b, *ApJL*, 793, L27
- Kreidberg, L., Line, M. R., Bean, J. L., et al. 2015, *ApJ*, 814, 66
- Leconte, J., Lai, D., & Chabrier, G. 2011a, *A&A*, 536, C1
- . 2011b, *A&A*, 528, A41
- Lendl, M., Cubillos, P. E., Hagelberg, J., et al. 2017, *A&A*, 606, A18
- Line, M. R., Knutson, H., Wolf, A. S., & Yung, Y. L. 2014, *ApJ*, 783, 70
- Line, M. R., Wolf, A. S., Zhang, X., et al. 2013, *ApJ*, 775, 137
- Line, M. R., Stevenson, K. B., Bean, J., et al. 2016, *AJ*, 152, 203
- Loeb, A., & Gaudi, B. S. 2003, *ApJL*, 588, L117
- Long, K. S., Baggett, S. M., & MacKenty, J. W. 2015, Persistence in the WFC3 IR Detector: an Improved Model Incorporating the Effects of Exposure Time, Tech. rep.
- Louden, T., & Kreidberg, L. 2017, ArXiv e-prints, arXiv:1711.00494
- Madhusudhan, N., Harrington, J., Stevenson, K. B., et al. 2011, *Nature*, 469, 64
- Mansfield, M., Bean, J. B., Line, M., Paramentier, V., & Kreidberg, L. submitted, Submitted to *ApJ*
- McCullough, P., & MacKenty, J. 2012, Considerations for using Spatial Scans with WFC3, Tech. rep.
- Menou, K., & Rauscher, E. 2009, *ApJ*, 700, 887
- Mordasini, C., van Boekel, R., Mollière, P., Henning, T., & Benneke, B. 2016, *ApJ*, 832, 41
- Mountain, C. M., Selby, M. J., Leggett, S. K., Blackwell, D. E., & Petford, A. D. 1985, *A&A*, 151, 399
- Parmentier, V., & Crossfield, I. 2017, ArXiv e-prints, arXiv:1711.07696
- Parmentier, V., Fortney, J. J., Showman, A. P., Morley, C., & Marley, M. S. 2016, *ApJ*, 828, 22
- Parmentier, V., Line, M. R., Bean, J., et al. in prep, In preparation
- Perez-Becker, D., & Showman, A. P. 2013, *ApJ*, 776, 134
- Perna, R., Menou, K., & Rauscher, E. 2010, *ApJ*, 719, 1421
- Pont, F., Zucker, S., & Queloz, D. 2006, *MNRAS*, 373, 231
- Reid, I. N., Cruz, K. L., Kirkpatrick, J. D., et al. 2008, *AJ*, 136, 1290
- Sheppard, K. B., Mandell, A. M., Tamburo, P., et al. 2017, *ApJL*, 850, L32
- Showman, A. P., Fortney, J. J., Lian, Y., et al. 2009, *ApJ*, 699, 564
- Showman, A. P., & Guillot, T. 2002, *A&A*, 385, 166
- Southworth, J., Mancini, L., Ciceri, S., et al. 2015, *MNRAS*, 447, 711
- Stevenson, K. B., Bean, J. L., Madhusudhan, N., & Harrington, J. 2014a, *ApJ*, 791, 36
- Stevenson, K. B., Harrington, J., Fortney, J. J., et al. 2012, *ApJ*, 754, 136
- Stevenson, K. B., Désert, J.-M., Line, M. R., et al. 2014b, *Science*, 346, 838
- Stevenson, K. B., Line, M. R., Bean, J. L., et al. 2017, *AJ*, 153, 68
- Teixeira, R., Ducourant, C., Chauvin, G., et al. 2009, *A&A*, 503, 281
- Tinetti, G., Tennyson, J., Griffith, C. A., & Waldmann, I. 2012, *Philosophical Transactions of the Royal Society of London Series A*, 370, 2749
- Van Der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science & Engineering, 13, 22
- Wahhaj, Z., Liu, M. C., Biller, B. A., et al. 2011, *ApJ*, 729, 139
- Wöllert, M., & Brandner, W. 2015, *A&A*, 579, A129
- Wong, M. H., Mahaffy, P. R., Atreya, S. K., Niemann, H. B., & Owen, T. C. 2004, *Icarus*, 171, 153
- Zellem, R. T., Swain, M. R., Roudier, G., et al. 2017, *ApJ*, 844, 27
- Zhang, X., & Showman, A. P. 2017, *ApJ*, 836, 73
- Zhou, Y., Apai, D., Lew, B. W. P., & Schneider, G. 2017, *AJ*, 153, 243