ACCESS: Confirmation of No Potassium in the Atmosphere of WASP-31b

Chima D. McGruder[1,14] , Mercedes López-Morales[1] , Néstor Espinoza[2] , Benjamin V. Rackham[3,15] , Dániel Apai[4,5,6] ,
Andrés Jordán[7,8] , David J. Osip[9], Munazza K. Alam[1,14] , Alex Bixel[4,6] , Jonathan J. Fortney[10] , Gregory W. Henry[11] ,
James Kirk[1] , Nikole K. Lewis[12] , Florian Rodler[13] , and Ian C. Weaver[1]

[1] Center for Astrophysics, Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138, USA; chima.mcgruder@cfa.harvard.edu
[2] Space Telescope Science Institute (STScI), 3700 San Martin Dr., Baltimore, MD 21218, USA
[3] Department of Earth, Atmospheric and Planetary Sciences, and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology,
Cambridge, MA 02139, USA
[4] Department of Astronomy/Steward Observatory, The University of Arizona, 933 N. Cherry Ave., Tucson, AZ 85721, USA
[5] Lunar and Planetary Laboratory, The University of Arizona, 1640 E. Univ. Blvd., Tucson, AZ 85721
[6] Earths in Other Solar Systems Team, NASA Nexus for Exoplanet System Science, USA
[7] Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Av. Diagonal las Torres 2640, Peñalolén, Santiago, Chile
[8] Millennium Institute for Astrophysics, Santiago, Chile
[9] Las Campanas Observatory, Carnegie Institution of Washington, Colina el Pino, Casilla 601 La Serena, Chile
[10] Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA
[11] Center of Excellence in Information Systems, Tennessee State University, Nashville, TN 37209, USA
[12] Department of Astronomy and Carl Sagan Institute, Cornell University, 122 Sciences Dr., 14853, Ithaca, NY, USA
[13] European Southern Observatory, Alonso de Cordova 3107, Vitacura, Santiago de Chile, Chile

Abstract

We present a new optical (400–950 nm) transmission spectrum of the hot Jupiter WASP-31b ($M = 0.48\,M_J$; $R = 1.54\,R_J$; $P = 3.41$ days), obtained by combining four transit observations. These transits were observed with IMACS on the Magellan Baade Telescope at Las Campanas Observatory as part of the ACCESS project. We investigate the presence of clouds/hazes in the upper atmosphere of this planet, as well as the contribution of stellar activity on the observed features. In addition, we search for absorption features of the alkali elements Na I and K I, with particular focus on K I, for which there have been two previously published disagreeing results. Observations with Hubble Space Telescope (HST)/STIS detected K I, whereas ground-based low- and high-resolution observations did not. We use equilibrium and nonequilibrium chemistry retrievals to explore the planetary and stellar parameter space of the system with our optical data combined with existing near-IR observations. Our best-fit model is that with a scattering slope consistent with a Rayleigh slope ($\alpha = 5.3^{+2.9}_{-3.1}$), high-altitude clouds at a log cloud top pressure of $-3.6^{+2.7}_{-2.1}$ bars, and possible muted H_2O features. We find that our observations support other ground-based claims of no K I. Clouds are likely why signals like H_2O are extremely muted and Na or K cannot be detected. We then juxtapose our Magellan/IMACS transmission spectrum with existing VLT/FORS2, HST/WFC3, HST/STIS, and Spitzer observations to further constrain the optical-to-infrared atmospheric features of the planet. We find that a steeper scattering slope ($\alpha = 8.3 \pm 1.5$) is anchored by STIS wavelengths blueward of 400 nm and only the original STIS observations show significant potassium signal.

1. Introduction

We are at the forefront of an exoplanet revolution with over 4200 planets discovered to date. Of these planets, over half have their fundamental properties (i.e., orbital parameters, masses, and radii) characterized, which give a basic understanding of their formation and interior composition.[16] However, to obtain a more in-depth understanding of how planets form and evolve, it is necessary to study their atmospheres in detail.

To date, transmission spectroscopy is the most successful approach to probe exoplanets' atmospheres. However, most of the instruments currently used for transmission spectroscopy were not designed with this science in mind, and unconstrained instrumental systematics are a major limitation to data precision. In the case of ground-based observations, effects related to the Earth's atmosphere are also a nontrivial factor. Because each telescope/instrument combination has its own unique systematics, there is no specific prescription to handle all of these limitations, which can lead to different interpretations of the same data based solely on the detrending process. This problem has caused several claims and counterclaims in this field (e.g., Espinoza et al. 2019 versus Sedaghati et al. 2017; Diamond-Lowe et al. 2018 versus Southworth et al. 2017). WASP-31b is one such planet that has had a detection of K I (Sing et al. 2015) that was later refuted (Gibson et al. 2017, 2019). To exacerbate the situation, stellar activity can mimic planetary features (see, e.g., Pont et al. 2013; McCullough et al. 2014; Espinoza et al. 2019), and tellurics might mask signals in ground-based observations (Gibson et al. 2017). Thus, the best approach in understanding the atmospheric transmission spectrum of an exoplanet is to repeatedly observe the planet and robustly detrend the data.

---

[14] NSF Graduate Research Fellow.
[15] 51 Pegasi b Fellow.
[16] See, e.g., https://exoplanets.nasa.gov/exoplanet-catalog/.

WASP-31 is a low-activity (3–5 Gyr), F6V dwarf, with low metallicity ([Fe/H] = −0.19) and an effective temperature of about 6300 K. WASP-31 ($V = 11.7$) is also a visual binary with a $V \sim 15.8$ mag companion approximately $35''$ away. WASP-31b is a transiting low-density ($\rho = 0.129~\rho_J$), inflated ($R = 1.54~R_J$) hot Jupiter with an orbital radius of 0.047 au, a period of 3.4 days, a mass of $0.48~M_J$, and a zero bond albedo equilibrium temperature of 1575 K (Anderson et al. 2011). This planet is a particularly good target for transmission spectroscopy because its lower gravity leads to a larger atmospheric scale height and stronger signals.

WASP-31b's transmission spectrum was observed with the Space Telescope Imaging Spectrograph (STIS) and Wide Field Camera 3 (WFC3) instruments on the Hubble Space Telescope (HST). These observations were combined with Spitzer photometry to construct a high signal-to-noise ratio (S/N) transmission spectrum. Analysis of those data indicated a Rayleigh scattering signature at short wavelengths, a cloud deck at longer wavelengths, and a strong potassium absorption feature (4.2σ confidence, with 78 Å wide bins), but no sodium (Sing et al. 2015). This is mostly consistent with ground-based observations later made with FORS2 on the VLT, aside from the detection of K I (Gibson et al. 2017). Gibson et al. reanalyzed the HST observations using Gaussian processes (GPs) to correct for systematics and found that the potassium detection significance decreased to 2.5σ. Additional high-resolution observations ($R > 80,000$) with the VLT's UV-Visual Echelle Spectrograph (UVES) did not detect sodium or potassium, supporting the idea that HST observations also might be susceptible to spurious signals produced by systematics (Gibson et al. 2019). However, because the detection of the K I feature from the ground can be contaminated by Earth's O_2 telluric features, it is worth further investigating these results.

Here we present four new transit observations of WASP-31b with spectral resolutions $300 < R < 1300$ and wavelength coverage between 0.4 and 0.97 μm, obtained between 2013 and 2019 as part of ACCESS.[17] ACCESS utilizes the multiobject spectrograph (MOS) IMACS (Dressler et al. 2011) on the 6.5 m Magellan Baade Telescope at Las Campanas Observatory, which allows for simultaneous collection of spectra from the exoplanet host star and many comparison stars to enable better systematic corrections. To date, ACCESS has produced transmission spectra of WASP-6b (Jordán et al. 2013), GJ 1214b (Rackham et al. 2017), WASP-4b (Bixel et al. 2019), WASP-19b (Espinoza et al. 2019), and WASP-43b (Weaver et al. 2020). To these we add here optical measurements of WASP-31b, which we combine with previously published HST/WFC3 and Spitzer IR data to produce a new transmission spectrum with coverage between 0.4 and 5.0 μm.

This paper is structured as follows: In Section 2, we present our Magellan/IMACS observations of WASP-31b in addition to photometric monitoring observations to constrain stellar variability. In Section 3, we discuss our transit light-curve extraction process, binning schemes, and detrending techniques. We outline the production of our final transmission spectrum and present a qualitative comparison of this spectrum to data from the literature in Section 4. In Section 5, we introduce the retrievals used to analyze the transmission spectra
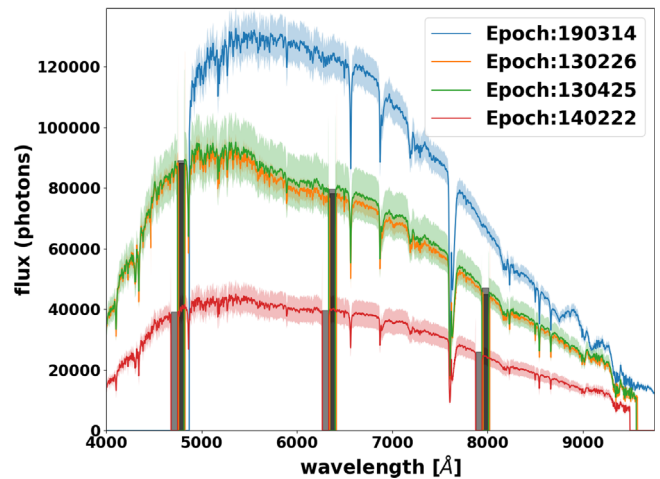

Figure 1. Median extracted spectra of WASP-31 from transits UT130226 (orange), UT130425 (green), UT140222 (red), and UT190314 (blue). The shaded regions of the same color extending past the median lines are the 1σ range of counts extracted for that night. The difference in counts between the $f/4$ observations are because of the exposure times. Different throughputs with the $f/4$ and $f/2$ setups explain the different shape of the spectra from UT190314 and from the $f/4$ observations. Dark shaded regions indicate chip gaps; as can be seen, there are three gaps for $f/4$ observations and usable data starts at 0.48 μm for the $f/2$ mode.

and report our best-fit results compared to best fits obtained from previous data. Interpretations of the retrieval results in the context of WASP-31b's overall atmospheric properties are discussed in Section 6. We conclude and summarize in Section 7.

2. Observations

2.1. General Setup

We observed four transits of WASP-31b on UT dates UT130226, UT130425, UT140222, and UT190314 (UTYYMMDD) with IMACS on the 6.5 m Baade Magellan Telescope at Las Campanas Observatory in Chile. The large time lapse between the first three transits and the last is because transit UT130226 could not be used for the transmission spectrum (as discussed in Appendix B), which we only found to be the case later in the analysis process.

IMACS is a wide-field imager and spectrograph with two $8K \times 8K$ CCD mosaic cameras at the $f/2$ and $f/4$ foci (Dressler et al. 2011). Each camera has a multi-object spectroscopy (MOS) mode, allowing for custom slit mask designs. The $f/2$ camera has a large, $27'.4$ field of view (FOV) and resolving powers up to $R \sim 5000$. The $f/4$ camera has a smaller FOV ($15'.4 \times 15'.4$) but higher spectral resolution (max $R \sim 9000$). The orientation of the $f/4$ camera causes each spectrum to be dispersed over up to four of the eight CCDs, whereas in the $f/2$ mode each spectrum is dispersed over up to two of the CCDs. The gaps between CCDs introduce gaps in the spectra, as shown in Figure 1. Our first three transits (UT130226, UT130425, and UT140222) were observed with the $f/4$ camera, and the last one (UT190314) was observed with the $f/2$ camera.

The large FOV of IMACS allows for simultaneous observations of the target and several comparison stars to correct for common time-series variations. We imposed strict criteria for the comparison star selection to ensure that they are photometrically stable, have similar apparent magnitudes, and

Table 1
Target and Comparison Star Magnitudes and Coordinates from the UCAC4 Catalog (Zacharias et al. 2013)

Star	R.A.	Decl.	B	V	J	K
WASP-31	11:17:45.4	−19:03:17.2	12.4	11.9	10.9	10.7
1	11:17:45.3	−19:07:25.9	11.8	11.1	10.0	9.6
2	11:17:32.0	−19:08:37.0	13.1	12.6	11.6	11.4
3	11:17:38.1	−18:59:19.6	13.7	13.0	11.9	11.5
4	11:17:44.3	−18:58:06.5	14.7	13.8	11.9	11.4
5	11:17:10.8	−19:14:26.0	12.5	11.9	11.0	10.7
6	11:16:35.7	−19:01:35.6	13.2	11.9	9.7	8.9
7	11:18:23.2	−18:58:45.8	13.3	12.8	11.9	11.6

Note. We used comparisons 2–4 for the white-light fit of epoch UT130226, 1–4 for epochs UT130425 and UT140222, and stars 2 and 5–7 for epoch UT190314. Comparison 1 for epoch UT130226 was saturated in most bins and not used for this transit.

have similar spectral types (color distance <1[18]), as described in detail in Rackham et al. (2017). The comparison stars selected for each observation are summarized in Table 1.

We designed masks with slits wide enough to prevent slit losses in typical seeing conditions at Las Campanas Observatory (LCO) but narrow enough to prevent contamination from nearby stars. This resulted in slit widths of $10''$ and $5''$ for the $f/2$ and $f/4$ mode, respectively. The lengths of the slits were $22''$ for $f/2$ and $12''$ for $f/4$ to adequately sample the sky background. We used a 300 line mm^{-1} grating at blaze angle of $4°3$ for the $f/4$ observations and a 300 line mm^{-1} grism at blaze angle of $17°5$ for the $f/2$ observation. These setups provided a wavelength coverage of 0.4–0.95 μm and 0.45–0.97 μm for the $f/4$ and $f/2$ observations, respectively. Finally, to reduce readout time and improve the duty cycle of the observations, we binned the detectors by 2×4 (dispersion $\times$ spatial directions) for the UT130266 transit. Transit UT130266 was the first observation done as part of ACCESS, while we were still experimenting with settings. After this first try we realized that seeing conditions at Las Campanas Observatory are often very good (typically 0.7 arcsec or better), so 2×4 binning undersamples the spectrum's point-spread function (PSF). We decided to use 2×2 binning for all subsequent WASP-31b observations, as an extra precaution to ensure proper sampling of the spectrum's PSF under all seeing conditions. Given the wide slits, the spectral resolution of the observations is seeing limited, with an average resolution of $R \sim 900$ (FWHM dispersion of 8.08 Å).

2.2. IMACS Data Collection

The transits of WASP-31b last 2.6 hr, so to cover the full transits and obtain enough baseline coverage to extract precise transit depths, we monitored the system during each transit epoch for about 5 hr continuous time blocks. We limited observations to epochs when WASP-31 was at air masses below 1.7 during the full observation window to minimize atmospheric refraction effects in the data. Exposure times were adjusted from 27 to 60 s to provide maximum count rates of 25,000–35,000 in analog-to-digital units (ADU; gain = 0.56 e^{-1}/ADU for $f/4$ setup, 1.0 e^{-1}/ADU for $f/2$ setup), which

was well within the CCD's linearity limit (Bixel et al. 2019). More details of specific transit observations are given in Table 2.

For each night we collected bias frames, high-S/N quartz lamp flats, and calibration arcs with HeNeAr lamps, each with the same binning scheme as the science observations. From the biases we found that the levels were essentially constant across the detector, so we adopted a constant bias level on each science frame using the median of their corresponding overscan region. We did not take dark frames because dark counts are negligible for our exposure times (Rackham et al. 2017). We took series of ~20 1 s flat frames before and after the science frames with the same configuration as the science images. The exposure times of the flats were set to give a count rate of about 35,000 ADU or under in order to stay within the linear regime of the CCDs (Bixel et al. 2019). However, during the data reduction (see Section 3.1), we found that flat-field corrections introduced additional noise in the data, similar to what was found with previous ACCESS observations (Rackham et al. 2017; Bixel et al. 2019; Espinoza et al. 2019; Weaver et al. 2020). Therefore, we decided not to apply the flat-field correction to any of our data sets. Finally, we took exposures with He, Ne, and Ar lamps using calibration masks identical to the science masks but with $0''5$ slit widths to increase the spectral resolution and prevent saturation of the CCDs. These lamps were used to wavelength-calibrate the spectra, as described in Section 3.1.

2.3. Photometric Monitoring

We analyzed photometric time series of the host star WASP-31 to constrain potential stellar contributions to the transmission spectrum of WASP-31b using observations from the Tennessee State University (TSU) Celestron 14-inch (C14) Automated Imaging Telescope (AIT) at Fairborn Observatory (Henry 1999), the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2014), and the All-Sky Automated Survey for Supernovae (ASAS-SN; Kochanek et al. 2017). The combined time baseline of these observations covers all our transit epochs, as shown in Figure 2. However, the photometric precision varies between surveys, so we decided to analyze each observing season for each survey separately.

2.3.1. AIT Analysis

TSU's AIT acquired 660 photometric observations through a Cousins R filter spanning six observing seasons (2011–12 through 2016–17). Standard deviations of a single observation from their seasonal means range between 3.2 and 4.2 mmag. This is near the limit of measurement precision with the AIT, as determined from the comparison stars in the field. The seasonal-mean differential magnitudes agree to a standard deviation of only 0.88 mmag, consistent with the absence of year-to-year variability and the low activity level of WASP-31. Periodogram analysis of each individual season resulted in the detection of significant periodicity only in the 2012–13 season. We detected a small peak-to-trough amplitude of 4 mmag with a period of 8.60 days, corresponding to the star's rotational period of 8.62 days as measured by Sing et al. (2015). This suggests that WASP-31 is a relatively inactive star.[19]

---

[18] Color distance, D, is defined in Rackham et al. (2017) as $D = \sqrt{[(B - V)_c - (B - V)_t]^2 + [(J - K)_c - (J - K)_t]^2}$, where t and c represent the target and comparison star magnitudes, respectively.

[19] See Sing et al. (2015) for further details of AIT observations, reduction, and analysis.

Table 2
Observing Log for WASP-31b Data Sets from Magellan/IMACS

Transit Date (UTC)	Obs. Start/ End (UTC)	Air Mass	Exposure Times (s)	Readout + Overhead (s)	Frames	Binning	Camera	min/max Resolution
2013 Feb 27	02:58/8:20	1.02–1.3	40	29	283	2 × 4	f/4	719/1126
2013 Apr 25	23:06/5:05	1.02–1.46	40	26	321	2 × 2	f/4	472/822
2014 Feb 23	02:32/08:44	1.02–1.43	27–36	26	363	2 × 2	f/4	295/1169
2019 Mar 15	03:04/07:20	1.02–1.3	60	31	168	2 × 2	f/2	739/1233

Note. The resolutions were calculated from the FWHM, which was estimated using the second moment method (Markevich & Gertner 1989). The average seeing conditions at LCO are 0″.6–0″.7 (Thomas-Osip et al. 2008), which is the resolution limiting factor.

2.3.2. TESS Analysis

WASP-31 was observed in TESS Sector 9 (2019 February 28–March 26) for 27 days with 2-minute cadence observations. To understand stellar variability, we used the out-of-transit Presearch Data Conditioning Simple Aperture Photometry (PDCSAP) light curve. After masking the transits of WASP-31b and binning to ∼3.8 hr, the standard deviation of this light curve was 350 ppm. The strongest periodicity we identified using Lomb–Scargle periodogram analysis (Lomb 1976; Scargle 1982) was that of a 2.88-day period with a peak-to-trough amplitude of 510 ppm. Like the AIT analysis, this supports the idea that WASP-31 is a low-activity—but not completely quiet—star. The period found with TESS is a third of that found with AIT. This type of aliased rotational period signal is common in the literature, likely owing to the evolving topology of stellar activity regions and harmonic peaks due to insufficient data (e.g., Charbonneau et al. 2009; Berta et al. 2011; Gillon et al. 2016; Luger et al. 2017; Mallonn et al. 2018). Specifically, the AIT light curves could have too sparse sampling or the TESS monitoring could not be long enough.

2.3.3. ASAS-SN Analysis

ASAS-SN monitors the entire visible sky (V-band filter) with 24 telescopes all over the world. There are 262 observations of WASP-31 from 2013 December to 2018 November with two ASAS-SN cameras (bg, be) at the Cerro Tololo Inter-American Observatory, Chile, and one camera (bc) at the Haleakala Observatory, Hawaii. The photometric precision of the ASAS-SN data is less than that of the AIT and TESS data, as seen in Figure 2. Thus, we cannot extract any evidence of stellar variability from the ASAS-SN light curves. Additionally, there is little ASAS-SN coverage of our transits; as such, we used the AIT and TESS data to constrain WASP-31's level of activity. The AIT and TESS phase-folded light curves are shown in Figure 3.

Our conclusion from the analysis of the AIT, TESS, and ASAS-SN data sets is that WASP-31b is a typical photometrically quiet F-type star. However, there is observational evidence that photometrically quiet F dwarfs are not completely inactive (e.g., López-Morales et al. 2016). Furthermore, photometric monitoring alone is not enough to fully characterize stellar activity levels, as axisymmetrically distributed active regions do not contribute to rotational variability (Rackham et al. 2019). Therefore, we still consider stellar activity in our retrieval analyses in Section 5.

3. Data Reduction and Light-curve Analyses

3.1. Reduction Pipeline

We reduced the raw IMACS data using the pipeline described in previous ACCESS papers (Jordán et al. 2013; Rackham et al. 2017; Espinoza et al. 2019; Bixel et al. 2019; Weaver et al. 2020). It performs standard bias and flat calibration, bad pixel and cosmic-ray correction, sky subtraction, spectrum extraction, and wavelength calibration. The pipeline produces sets of extracted, wavelength-calibrated spectra for each epoch, for the target and each of the comparison stars in the field.

The wavelength calibration was done by first fitting Lorentzian profiles to each spectral line on images taken with the calibration masks. Next, using these pixel positions and the known vacuum wavelengths, the wavelength solution for each spectrum was found by an iterative process in which a sixth-order polynomial was fitted to the wavelengths as a function of pixel position, the data point with the greatest deviation from the fit was removed, and the process was repeated until the rms error value of the fit was less than 2 km s^{-1} (∼0.05 Å; Espinoza 2017; Rackham et al. 2017).

The extracted, wavelength-calibrated spectra of WASP-31 for each of the four epochs are shown in Figure 1, which highlights the gaps between CCDs. In the epochs observed with the f/4 camera, the spectra have three gaps. For the UT190314 transit, the one gap of the f/2 mode was close to the blue edge of the spectrum. This inhibited proper calibration of wavelengths shorter than ∼0.48 μm, which we excluded.

The extracted spectra are then integrated across the entire wavelength range to produce white light curves or integrated over a narrow wavelength range to produce the binned light curves. These integrated white light curves are shown in the top panel of Figure 4.

3.2. Light-curve Detrending

We first modeled the light curves following the approach of Jordán et al. (2013), using principal component analysis (PCA). For transit data (time-series measurements) we assume that the light curve $L_k(t)$ for a given comparison star k is a linear combination of a set of signals $s_i(t)$, which represent the different instrumental and atmospheric effects on the light curves, i.e.,

$$L_k(t) = \sum_{i=1}^{K} A_{k,i} s_i(t), \qquad (1)$$

where $A_{k,i}$ is the weight of each signal $s_i(t)$ in each light curve.
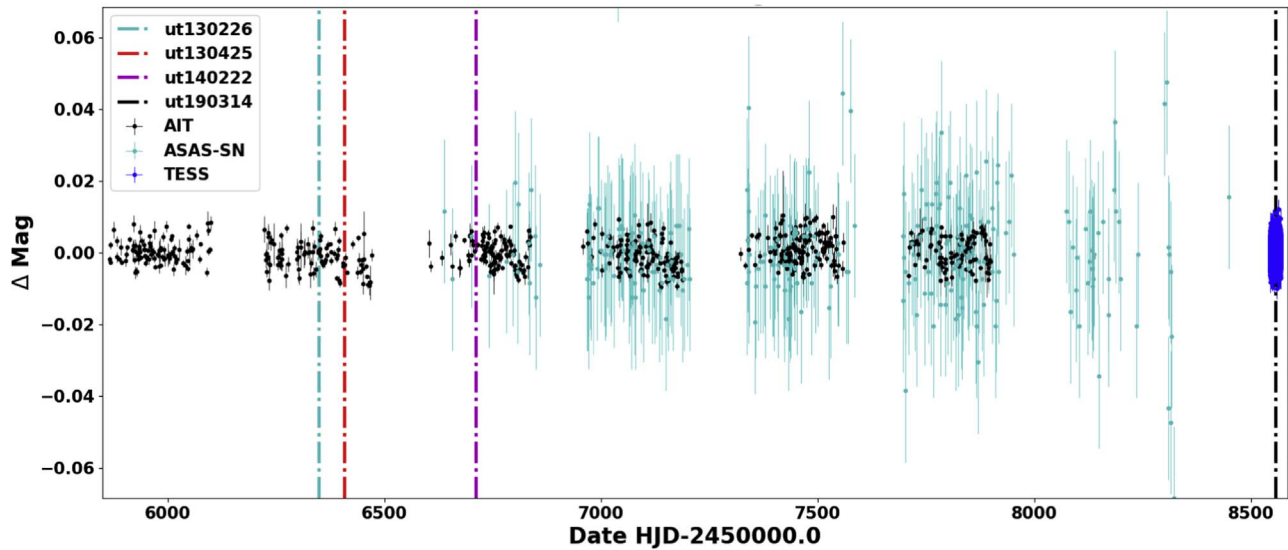
Figure 2. Photometric monitoring of WASP-31 with AIT (black), ASAS-SN (cyan), and TESS (blue). The times of the transits observed with Magellan are shown as dashed–dotted vertical lines. Given the differing observing bands of ASAS-SN (V), AIT (R), and TESS (650–1050 nm), we limit our comparison to relative magnitude changes. Thus, the values plotted are the differences from the mean magnitude of each survey. All observations suggests that the star is relatively inactive; however, this cannot confidently be constrained with photometry alone.

With each light curve $L_k(t)$ as a row in a matrix D, we perform singular value decomposition.[20] This returns the eigenvectors e_i and eigenvalues λ_i of the matrix. We obtain $s_i(t)$ from the product of the matrix of eigenvectors and the original data matrix D. Additionally, the importance of each signal for reconstructing D is given by the eigenvalues λ_i (largest λ_i, most important e_i, etc.). This allows for the sequential fitting of those signals to our target star in order to find the simplest model that preserves the most important information present in the comparison stars. For further details of PCA, we refer the reader to Jolliffe (2002), or Espinoza (2017) for its usage in high-S/N time-series photometry. We implemented PCA with the same procedure as Jordán et al. (2013) and Rackham et al. (2017) and found two, four, two, and three of the four PCA components to be the optimal number of components for each transit, chronologically.[21]

However, as can be seen in the middle panel of Figure 4, using this PCA method alone was not sufficient to properly correct all systematics. This is because there are likely additional systematics that uniquely affect WASP-31 owing to, e.g., color, detector, and instrumental differences with respect to the comparison stars. To account for this, we used Gaussian process (GP) regression to further model the target light curve in conjunction with PCA. A GP is a machine-learning technique for nonparametric regression modeling (Rasmussen & Williams 2005) and has proven effective at detrending radial velocity data (e.g., Aigrain et al. 2012) and transmission spectroscopy (e.g., Gibson et al. 2012). For our GP analysis, we define a joint probability distribution of the form $\mathcal{N}[0, \Sigma]$, where the covariance function (Σ) is defined as $K_{SE}(x_i, x_j) + \sigma_w^2 \delta_{i,j}$. Here σ_w^2, $\delta_{i,j}$, and $K_{SE}(x_i, x_j)$ are a jitter term, the Kroenecker delta function, and our *kernel*, respectively.

A multidimensional squared-exponential kernel was used for our GP regression:

$$K_{SE}(x_i, x_j) = \sigma_{GP}^2 \exp\left(-\sum_{d=1}^{D} \alpha_d (x_{d,i} - x_{d,j})^2\right), \quad (2)$$

where σ_{GP}^2 is the amplitude of the GP and α_d are the inverse (squared) length scales of each of the components of the GP. The x_i vectors here have components $x_{d,i}$, where each i denotes a time stamp and where each d corresponds to a different external parameter. The motivation for using the SE kernel is that it is a relatively smooth function and self-regulates the importance of the external parameters (x) (Micchelli et al. 2006; Duvenaud 2014).

Because the joint probability of the data can be defined by a multivariate Gaussian using GPs, the log marginal likelihood is known and given by

$$\ln \mathcal{L} = -\frac{1}{2} r^T \Sigma^{-1} r - \frac{1}{2} \ln|\Sigma| - \frac{N}{2} \ln(2\pi). \quad (3)$$

Here N is the number of measurements of the data. We used `george`[22] (Foreman-Mackey et al. 2014) to evaluate the marginalized likelihood of our data set and `PyMultiNest`[23] (Buchner et al. 2014) (a python wrapper of `MultiNest`; Feroz et al. 2009) to sample the hyperparameter space. In our analyses, we found including five external parameters x useful to model the PCA-corrected data with GPs. These time-dependent parameters were (i) the sky flux, (ii) the variation of the FWHM of the spectra, (iii) the air mass, (iv) the position of the central pixel trace (perpendicular to the dispersion axis), and (v) the drift of the wavelength solution.

The first three parameters are proxies of how Earth's atmosphere is affecting the stellar spectra. Over all exposures the FWHM ranges from 4.3 to 17.1 pixels and the air mass

---

[20] $\text{Cov}_t[s_i(t)s_j(t)] = 0$ assuming that the signals are uncorrelated.

[21] The optimal number of PCA components was determined with the k-fold cross-validation procedure (Hastie et al. 2007).

[22] https://github.com/dfm/george

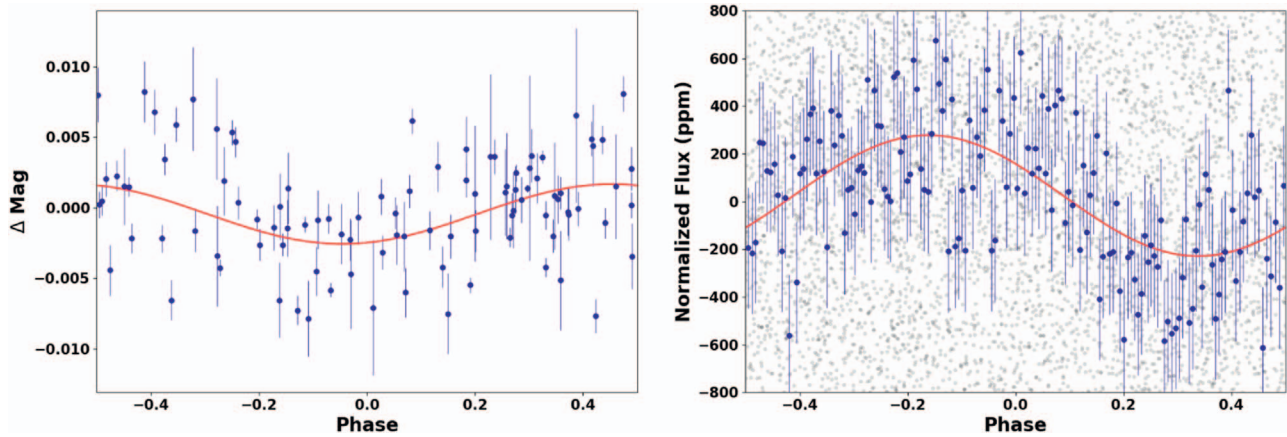[23] https://github.com/JohannesBuchner/PyMultiNest

Figure 3. Phase curves of the AIT and TESS observations, showing low-amplitude periodicity in periodogram analyses. AIT data (left) show an 8.6-day periodicity with an amplitude of 0.004 mag. For the TESS data (right), the transparent gray points are all ~15,500 out-of-transit 2-minute cadence observations from TESS Sector 9. The blue points are the same data binned to portray the periodicity better. The red line is the 2.88-day, 507 ppm amplitude best-fit Lomb–Scargle model.

varied from 1.7 to 1.0 over the course of a night for each transit. The last two parameters are proxies for how slight physical variations in the instruments/detector could affect the spectra. Ultimately, exactly how these parameters affect the spectra is not well understood, but GPs are optimal at finding correlations between parameters and observables (Rasmussen & Williams 2005).

The next task is to properly integrate the PCA, GP, and inherent transit model into one function defining our light curves. Our particular model used the magnitude of the target and comparison stars and was also used by Yan et al. (2020) to detrend ground-based spectroscopic transit data:

$$M_k(t) = c_k + \sum_{i=1}^{N_k} A_{k,i} s_i(t) - 2.51 \log_{10} T(t|\phi) + \epsilon, \quad (4)$$

where $M_k(t)$ is the (mean-subtracted) magnitude of the target star in the kth model, c_k is a magnitude offset, N_k is the number of PCA signals $s_i(t)$, $A_{i,k}$ is the weight for each signal in each model, $T(t|\phi)$ is the transit model with parameters ϕ,[24] and ϵ is a stochastic component here modeled as a GP.

Though PCA provides a weighting scheme for the importance level of the signals $s_i(t)$, the determination of which components improves our light-curve model is still nontrivial. In order to incorporate our ignorance of the functional form of nuisance signals into our model fit, we used the model averaging technique outlined by Gibson (2014). For our case, we averaged the posterior distributions of each $M_k(t)$, where the posteriors were explored using `PyMultiNest`. Our final model-averaged white light curves can be seen in the bottom row of Figure 4.

3.3. Light-curve Fitting

Using an improper limb-darkening (LD) law can introduce biases in the final light-curve fit (Espinoza & Jordán 2016), and as such we used the open-source package `ld-exosim`[25] to determine the most appropriate LD law for our transit fits. The square root law was found to be best for all white light fits,

given the data quality. We used the same law for the spectroscopic fits, described in Section 4.

Next, we applied a σ-clipping to each of the white light curves to remove data points that deviated more than 3σ from the model. This clipping was applied after obtaining an initial detrended light curve and corresponding best-fit model with all available data and lead to three, five, zero, and one data point being removed from epochs UT130226, UT130425, UT140222, and UT190314, respectively.

We then used the 3σ-clipped light curves to find the best transit parameters. The transit parameters used to fit the light curves were LD coefficients (q_1, q_2), planet orbital period (P), semimajor axis (in terms of star radius; a/R_s), the planet-to-star radius ratio (p), impact parameter (b), midtransit time (t_0), inclination (i), eccentricity (e), and the argument of periapsis (ω). We kept e and ω fixed to 0° and 90° (assuming a circular orbit), following the results of Anderson et al. (2011). We placed Gaussian priors on all other parameters using the results of Anderson et al. (2011) and then fit for them following the procedure described in Section 3.2. The best-fit white-light-curve orbital parameters for each epoch are summarized in Table 3.

4. Transmission Spectra

To produce the optical transmission spectrum of WASP-31b, we generated transit light curves in the same manner as described in Section 3, but using a set of narrower wavelength bins. We propagated the flagged 3σ outliers from the white-light analysis to each binned light curve. Our binning scheme was devised so it optimized spectrophotometric precision, while still letting us probe for atmospheric features, such as a scattering slope, and sodium and potassium lines. The average bin widths were 187 Å, where the lower-throughput, redder bins were made larger, and the regions around the Na I and K I features were 60 Å wide centered on each doublet (5892.9, 7681.2 Å). We found 60 Å to be enough to encompass both doublet lines of each feature, providing high signal from potential features while minimizing suppression of signal from the surrounding continuum. This is also comparable to the 78 Å bin widths that Sing et al. (2015) used to report a 4.2σ K I detection.

We employed the same analysis procedures as the white light curves (see Section 3.3) but with P, a/R_s, b, t_0, and i

---

[24] We used the Python package `batman` (Kreidberg 2015) to produce the analytic transit model.
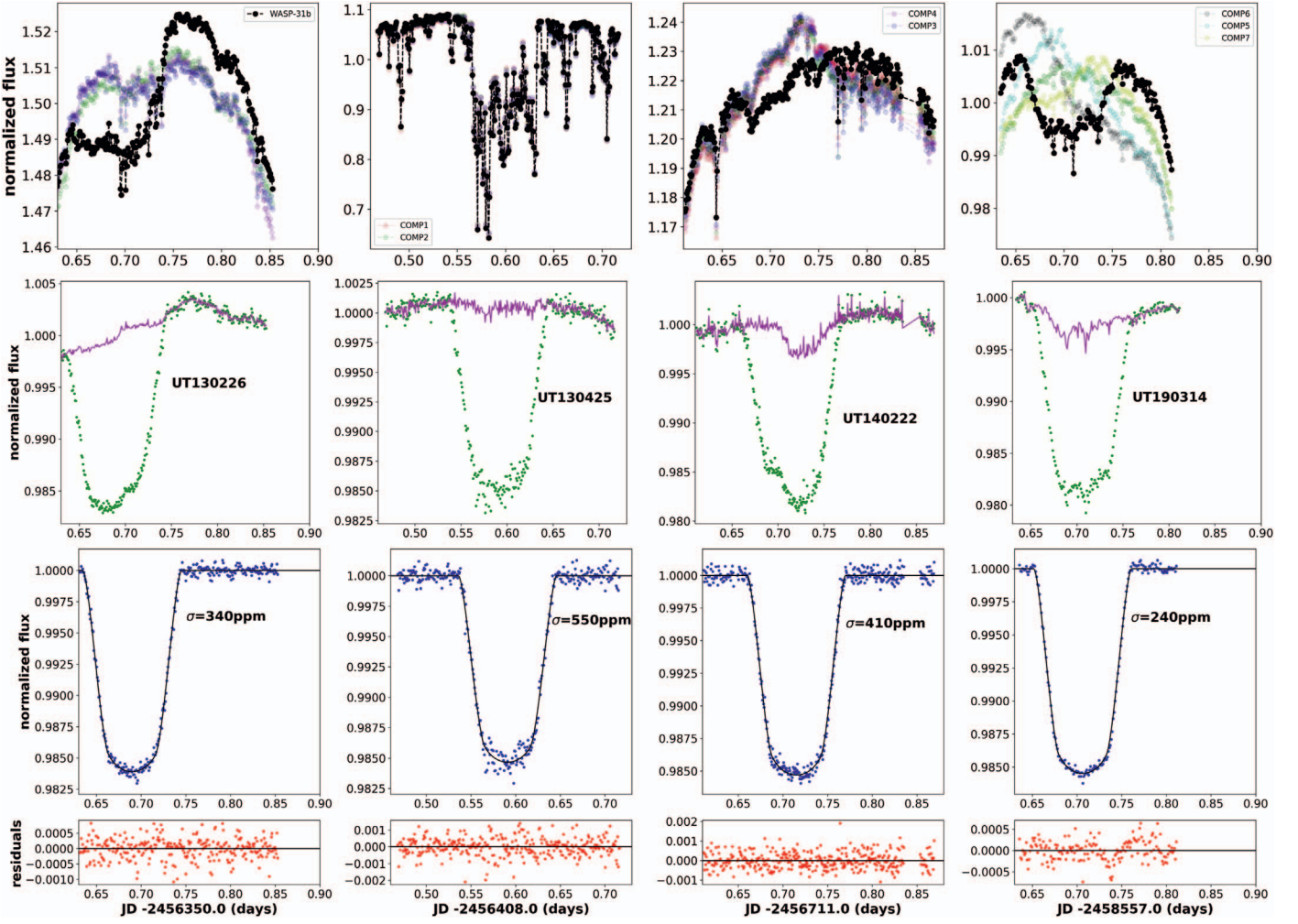[25] https://github.com/nespinoza/ld-exosim

Figure 4. Top: raw white light curves for our Magellan/IMACS observations directly from ACCESS's custom pipeline. Black is WASP-31, and transparent red, green, blue, magenta, cyan, gray, and yellow are comparison stars 1–7, respectively. See Table 1 for more information about the stars. Middle: PCA-corrected light curves are shown in green. The magenta lines are GP fits on the light curves to correct for systematics using the five external parameters. This figure depicts that PCA correction alone is not sufficient to remove all systematics, and as such model averaging of PCA and GPs are justified. Bottom: final white light curve with the model-averaged PCA and GP systematics removed (blue points), along with the corresponding residuals (red points). The best-fitting transit light curve given our model averaging procedure is depicted as a solid black line. The value of σ given in each panel is the standard deviation of the residuals in ppm. The middle panel is not an intermediate step to arrive at the bottom panel, given that the bottom panel is obtained directly when simultaneously doing PCA and GP model averaging of the data in the first panel. As can be seen in the right column, transit UT130226 has little pre-transit data, which might cause systematics in its transmission spectra. This is explored further in Appendix B.

Table 3
Fitted White-light-curve Values

Parameter	Definition	UT130226	UT130425	UT140222	UT190314
p	Planet radius/star radius	0.129 ± 0.002	0.125 ± 0.002	$0.124^{+0.003}_{-0.006}$	0.13 ± 0.002
$t_0 - 2.45e6$	Midtransit (JD)	$6350.6898^{+0.0004}_{-0.0005}$	6408.5909 ± 0.0002	$6711.7150^{+0.0003}_{-0.0002}$	8557.7067 ± 0.0002
P	Period (days)	3.405910 ± 0.000005	3.405910 ± 0.000005	3.405910 ± 0.000005	3.405910 ± 0.000005
a/R_s	Semimajor axis/star radius	$8.15^{+0.10}_{-0.11}$	8.04 ± 0.09	8.13 ± 0.09	$8.15^{+0.07}_{-0.06}$
b	Impact parameter	$0.772^{+0.007}_{-0.008}$	$0.784^{+0.007}_{-0.008}$	0.769 ± 0.006	$0.775^{+0.006}_{-0.008}$
i	Inclination (degrees)	84.6 ± 0.1	84.4 ± 0.1	84.6 ± 0.1	84.5 ± 0.1
q_1	LD coeff 1	0.7 ± 0.2	$0.4^{+0.3}_{-0.2}$	0.6 ± 0.3	$0.6^{+0.2}_{-0.3}$
q_2	LD coeff 2	0.4 ± 0.2	$0.5^{+0.2}_{-0.3}$	$0.7^{+0.2}_{-0.3}$	$0.8^{+0.1}_{-0.3}$

Note. These are calculated using all integrated flux and available comparisons, unique to each epoch. This causes variation in depths per night but improves the precision of the light-curve parameters needed for binned light-curve fits.

held fixed to values obtained from the corresponding white-light-curve fits, leaving only q_1, q_2, and p as free parameters. The detrended light curves, their best-fit model, and residuals for each bin of each epoch are shown in Appendix A. Appendix A also lists the values of p obtained for each bin.
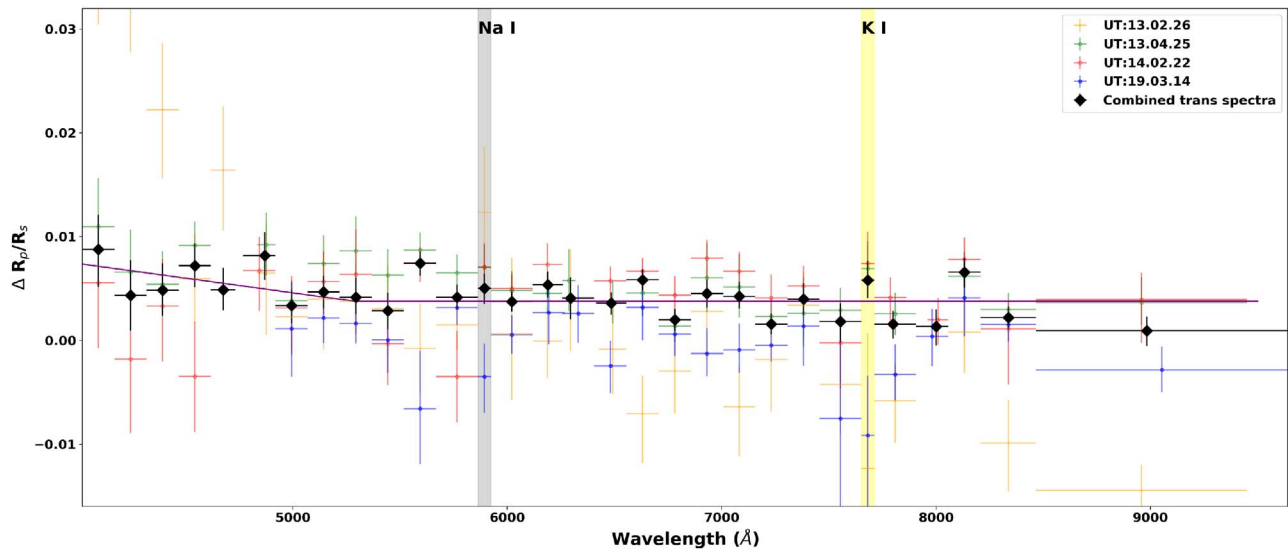
Figure 5. All four transmission spectra of WASP-31b taken by Magellan/IMACS, before an offset is applied (faded colors), and the weighted average of the last three spectra (black). The first transit, UT130226 (transparent orange), significantly deviates from the other three spectra and has its highest $\Delta R_p/R_s$ of 0.0385, but it cannot be seen in this figure because emphasis is placed on the other three consistent spectra. Transit UT130226 is likely an outlier because of insufficient baseline during observation or stellar activity (see Appendix B). The purple line is a two-part fit with a scattering slope blueward of 5500 Å and a featureless spectrum redward of 5500 Å, similar to fits done by Sing et al. (2015) and Gibson et al. (2017). The highlighted gray and yellow regions are centered around the Na I and K I lines, respectively.

4.1. Combined Transmission Spectrum

The average precision of each bin for each spectrum is $0.007 \, R_p/R_s$. This is larger than the $\sim 0.006 \, R_p/R_s$ expected signal from Na and K features for this planet (Sing et al. 2015). Therefore, we combined the transmission spectra from epochs UT130425, UT140222, and UT190314 to achieve higher precision. The UT130226 transit was not used in our combined data set because its transmission spectrum was inconsistent with the others, perhaps because of stellar activity or insufficient baseline before ingress. We discuss this further in Appendix B.

Our spectra were combined by first applying an offset to each spectrum equal to the difference between their white-light-curve depths and the overall mean white-light-curve depth. We then combined the measurements for each spectroscopic bin by applying a weighted average using each point's R_p/R_s error bars as weights. The combined transmission spectrum, shown as black diamonds in Figure 5, has an average precision per bin of $0.0033 \, R_p/R_s$. We emphasize that when combining transmission spectra of different epochs, as we did, information about how the star and the terminator of the planet's atmosphere are behaving at a given epoch is lost, leaving only persistent (average) behaviors in the planet–star system.

As can be seen in Table 3, the best-fit planet radius varies per epoch. These differences in our white-light-curve transit depths can be explained by the different locations of wavelength gaps in the detector for both the comparison stars and WASP-31, as shown in Figure 1, which causes the integrated flux to be slightly different per epoch. Additionally, some epochs use different comparison stars as shown in Table 1, which cause differences in white-light-curve flux ratios when dividing the flux of the targets by the flux of the comparison stars. We tested this hypothesis by using one matching comparison and a fixed wavelength range with no gaps for all data sets and found consistent transit depths for all three epochs. However, this approach significantly reduces the precision of the white light curves and the spectrophotometric bins, so we resorted back to

applying offsets between epochs in order to use the full wavelength coverages and all comparisons. This approach is justified because the most important information from the transmission spectra is the relative changes in depth between wavelength bins. In addition, to address this issue, we fit for offsets between our combined transmission spectrum and other data sets in our atmospheric retrieval analysis, as described in Section 5.

Figure 5 shows the individual transmission spectra for each transit epoch, together with our combined transmission spectrum. Overlaid on the combined spectrum is a two-part fit with a scattering slope blueward of 0.55 μm and a featureless spectrum redward of 0.55 μm, similar to fits done by Sing et al. (2015) and Gibson et al. (2017). The fit is favored over a completely featureless spectrum ($\chi^2 = 50.9$) with a χ^2 of 46.2.

4.2. Previous Optical Transmission Spectra

All available low-resolution optical transmission spectra of WASP-31b are shown in Figure 6. When qualitatively comparing the Magellan/IMACS data with the VLT/FORS2 (Gibson et al. 2017) and HST/STIS observations (Sing et al. 2015), we see that all three spectra can be modeled with a scattering slope in the blue and a featureless spectrum beyond. The simple models in Figure 6 were created the same way as in Figure 5, just as a comparison with the Sing et al. (2015) model. However, in our retrieval analysis (Section 5.2), we do not use a two-part model, but rather obtain scattering slopes with one continuous atmospheric model over the whole spectrum. This is also consistent with the findings of Barstow et al. (2016), who found that a "split" spectrum with the HST data is not significantly favored compared to a continuous model. Like the FORS2 observations, our IMACS observations show no signs of potassium or sodium absorption. This is contrary to the detection of potassium with HST/STIS. Though the overall shapes of the spectra agree with each other (aside from potassium absorption), we elected not to combine previous studies with our own in the subsequent analysis. This
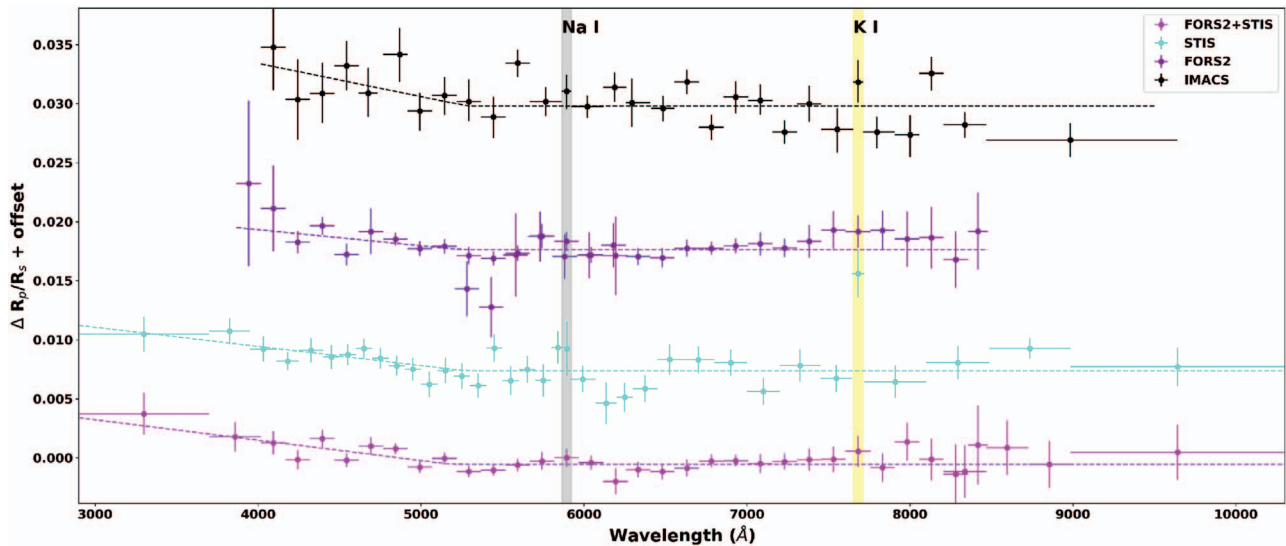
Figure 6. Transmission spectra of WASP-31b taken by Magellan/IMACS (excluding epoch UT130226; black), VLT/FORS2 (purple; Gibson et al. 2017), HST/ STIS (cyan; Sing et al. 2015), and combined FORS2 and reanalyzed STIS data (magenta; Gibson et al. 2017), with vertical offsets for clarity. Each spectrum has a two-part fit of a scattering slope blueward of 5500 Å and a featureless spectrum redward of 5500 Å, excluding the possible K signal centered at 7682.5 Å for the HST/ STIS data, overlaid as dashed lines. All three spectra have signs of a blue scattering slope and featureless optical spectrum. The highlighted gray and yellow regions are centered around the Na I and K I lines, respectively.

was because we are uncertain of the biases introduced to the data by the various detrending techniques for each analysis, and combining the spectra may lead to misinterpretation of signals lost or produced by compounded biases. This rationale is supported by the analysis of Gibson et al. (2017), which found significantly different results for the same data set depending on the detrending method. However, we did combine the IMACS optical transmission spectrum with near-IR HST/ WFC3 spectra and mid-IR (Spitzer/IRAC) data (Sing et al. 2015), because there is only one observation/analysis for each and including them will help better characterize WASP-31b's atmospheric structure. In total, the combined transmission spectrum covers a wavelength range from 0.4 to 5.06 μm, with average bin sizes of 187 and 200 Å for the optical (Magellan/ IMACS) and near-IR (HST/WFC3) data.

5. Atmospheric Retrieval Analyses

We use atmospheric retrievals to provide a quantitative measurement of the existence of features in the transmission spectra. Given that results can depend on the priors and the specific retrievals used (e.g., Kirk et al. 2019), we employed two independent retrievals to improve data interpretation: PLATON[26] (Zhang et al. 2019), and Exoretrievals (Espinoza et al. 2019).

PLATON and Exoretrievals complement each other in several ways: Exoretrievals considers a small number of user-defined molecules, but PLATON considers a large number (33) of predetermined molecules. Exoretrievals performs free retrievals, allowing for abundances of individual molecules to be retrieved with nonequilibrium chemistry, whereas PLATON assumes equilibrium chemistry throughout the atmosphere. Thus, using both retrievals gives a more thorough understanding of the planet. Additionally, finding agreeing results in overlapping parameters strengthens support of

accurate interpretation of the data, given that they use very different fundamental assumptions.

In detail, PLATON uses the planet's limb temperature (T_p), atmospheric metallicity ($Z/Z_\odot$), C/O ratio, cloud top pressure (P_0), a scattering slope (α), and an offset term between instruments to characterize observed transmission spectra using equilibrium chemistry. PLATON's sources of line lists for all 33 molecules considered are shown in Table 2 of Lupu et al. (2014), which mostly, but not exclusively, comes from HITRAN. PLATON includes collision-induced absorption, where all coefficients come from HITRAN. PLATON also simultaneously fits for stellar activity using the occulted temperature of the star ($T_{\rm star}$), the temperature of unocculted spots ($T_{\rm spot}$), and their area ($f_{\rm spot}$). PLATON uses the nested sampling algorithm dynesty[27] (Speagle 2020) to explore the parameter space.

Exoretrievals uses a semianalytical formalism to model the planet with an isothermal, isobaric atmosphere with nonequilibrium chemistry (see Heng & Kitzmann 2017). The cross sections we considered with Exoretrievals were CO, CO2, CH4, TiO, NH3, FeH, HCN, H2O, Na, and K, where the line list of the molecules came from HITEMP (Rothman et al. 2010) and ExoMol (Yurchenko et al. 2013; Tennyson et al. 2016), and TiO was synthesized using the latest (2012 version) line lists calculated by B. Plez, obtained from VALD (Ryabchikova et al. 2015)). The Na and K cross sections were determined from analytical Lorentzian-profiled doublet shapes used in MacDonald & Madhusudhan (2017). Unlike PLATON, Exoretrievals does not include collision-induced absorption. However, it does also account for stellar photospheric heterogeneities by modeling the occulted and unocculted regions of the stellar photosphere with PHOENIX models (Husser et al. 2013), following the approach of Rackham et al. (2018, 2019). The retrieval has three parameters to characterize the stellar surface: the temperature of the region occulted by the planet ($T_{\rm occ}$); the temperature of

---

[26] https://github.com/ideasrule/platon

[27] https://github.com/joshspeagle/dynesty

Table 4
Priors for Exoretrievals and PLATON

Exoretrievals			PLATON		
Parameter	Prior Function	Prior Bounds	Parameter	Prior Function	Prior Bounds
Cloud top Pressure (P_0, bars)	log-uniform	−8 to 2	Cloud top Pressure (P_0, pascals)	log-uniform	−3 to 7
Planetary atmospheric temperature (T_p)	uniform	800 to 1900K	Planetary atmospheric temperature (T_p)	uniform	800 to 1900K
Stellar temperature (T_{occ})	uniform	6000.0 to 6300.0K	Stellar temperature (T_{star})	Gaussian	mean = 6300 K, std = 150K
Stellar heterogeneities temperature (T_{het})	uniform	2300.0 to 7000.0K	Stellar heterogeneities temperature (T_{spot})	uniform	5900.0 to 6700.0K
Spot covering fraction (f_{het})	uniform	0 to 0.8	Spot covering fraction (f_{spot})	uniform	0 to 0.8
Offset (depth)	Gaussian	mean = 0, std = 1000ppm	Offset (depth)	uniform	−8400 to 8400ppm
Cloud/haze amplitude (a)	uniform	−30 to 30	Scattering factor	log-uniform	−10 to 10
Cloud/haze power law (γ)[a]	uniform	−10 to 4	Scattering slope (α)[b]	uniform	−4 to 10
Cloud absorbing cross section (σ_{cloud})	log-uniform	−80 to 80	Metallicity (Z/Z_{odot})	log-uniform	−1 to 3
Trace molecules' mixing ratios	log-uniform	−30 to 0	C/O	uniform	0.05 to 2
Reference radius factor (f)	uniform	0.8 to 1.2	Reference radius (R_p)	uniform	1.35 to 1.6R_j

Notes. These priors were set to allow for a wide parameter space to be surveyed, but contained within physical regimes. Not all parameters were included in each model fit (see Tables 5). We used 10,000 live points for each Exoretrievals run and 5000 for PLATON. For further description of the parameters of Exoretrievals, please refer to the Appendix of Espinoza et al. (2019).
[a] This is the exponent of the scattering slope power law, where −4 is a Rayleigh scattering slope.
[b] This is the wavelength dependence of scattering, with 4 being Rayleigh.

unoccculted heterogeneities (T_{het}), which can be hotter or colder than the mean photosphere; and the heterogeneity covering fraction (f_{het}). The planet's atmosphere is characterized by a factor (f) multiplied by the planetary radius to find a reference radius at which the atmosphere is optically thick (i.e., $R_0 = fR_p$), the corresponding pressure where the atmosphere is optically thick (P_0), the limb atmospheric temperature (T_p), mixing ratios of specified species (i.e., Na, K, and H_2O), offset terms for different instruments, and two parameters to characterize the cloud/haze extinction (κ). The two cloud/haze parameters are a *Rayleigh-enhancement factor* (a) and a *cloud/haze power law* (γ), and they quantify extinction with

$$\kappa(\lambda) = a\sigma_0(\lambda/\lambda_0)^\gamma, \quad (5)$$

where $\sigma_0 = 5.31 \times 10^{31}$ m^2 and $\lambda_0 = 350$ nm (MacDonald & Madhusudhan 2017; Sedaghati et al. 2017). The code uses nested sampling with PyMultiNest to measure the evidence for each of the retrieved models. The priors we used for both retrievals are shown in Table 4. Because Exoretrievals allows for individual molecules to be tested, we used the results of Exoretrievals to determine the presence of water, potassium, or sodium.

5.1. Combined Magellan Data

We run PLATON and Exoretrievals on the combined IMACS, HST/WFC3, and Spitzer data. For each retrieval run, we fit for an offset between the optical data and the IR data but no offset between the Spitzer and the HST/WFC3. The Spitzer and HST/WFC3 data were taken as a single IR data set, where Sing et al. (2015) analyzed them both using the same ephemeris and system parameters. Using the retrievals, we tested for different scenarios: (1) a featureless exoplanet atmosphere, where all potential signals are covered by high-altitude clouds,

(2) a spectrum with any signal detected in the spectrum coming from active regions on the surface of the star, (3) an atmosphere with scatterers, and (4) an atmosphere with scatterers and stellar activity contamination. Additionally, we tested Exoretrievals' models with the same four scenarios, but including only H_2O, Na, or K, individually and simultaneously in order to isolate their atmospheric effects.

Table 5 summarizes the Bayesian evidences of each scenario with respect to a featureless spectrum ($\Delta \ln Z$). Following Trotta (2008) and Benneke & Seager (2013), the frequentist significance of those $\Delta \ln Z$ values scales as follows: $|\Delta \ln Z|$ of 0 to 2.5 is inconclusive with $<2.7\sigma$ support for the higher evidence model, $|\Delta \ln Z|$ of 2.5 to 5 corresponds to a moderately significant detection of 2.7σ to 3.6σ, and $|\Delta \ln Z| \geqslant 5$ corresponds to strong support for one model over the other.

Exoretrievals could also fit for CO, CO_2, CH_4, TiO, NH_3, FeH, and HCN. Given the temperature (Fortney et al. 2008; Evans et al. 2018), wavelength range, and likely high-altitude clouds (see discussion in Section 6), we do not expect to detect any of those seven species. As a precaution, we did run Exoretrievals with the three cases discussed above (featureless, activity, scatters) and including each of the seven species individually. As expected, there was insufficient evidence to support any of those models.

The best model fit from Exoretrievals was a flat spectrum. However, the $|\Delta \ln Z|$ between the featureless model and the majority of the others was below 2.5, implying that those models cannot be distinguished from a flat spectrum. For PLATON the best fit was a model with a near-Rayleigh scattering slope ($5.3^{+2.9}_{-3.1}$), log metallicity of $1.3^{+1.1}_{-1.4}$, and C/O $= 0.52^{+0.87}_{-0.32}$, but again the support for this model is indistinguishable from the others ($\Delta \ln Z = 1.7$). The undefined features (i.e., lack of Na or K) and nominal detection of a scattering slope found here are consistent with the results of

Table 5
$\Delta \ln Z$ for Various Exoretrievals (left) and PLATON (right) Models Relative to a Featureless Spectrum with the Subset of Data That Included the Combined Magellan/IMACS (Excluding Transit UT130226), HST/WFC3, and Spitzer IR Data

	Exoretrievals					PLATON	
Model	Featureless	H_2O	Na	K	H_2O + K + Na	Model	
Featureless	0.0	−0.54	−1.09	−0.99	−1.19	Featureless	0.0
Scatterers	⋯	−1.77	−2.19	−2.42	−2.48	Scattering	1.7
Activity	−0.29	−0.41	−0.79	−0.97	−1.1	Stellar activity	0.63
Scatterers and activity	⋯	−2.66	−3.09	−2.72	−3.35	Both	1.04

Note. For Exoretrievals no model is significantly preferred, but a flat spectrum has the highest evidence. For PLATON a spectrum with an optical scattering slope is slightly favored but indistinguishable from the other models.

Gibson et al. (2017). Gibson et al. found no significant detection of K or Na and had only a tentative detection of a scattering slope with their FORS2 data. It was only when including the bluer HST/STIS data between 0.3 and 1.0 μm that they confidently detect a scattering slope. Unfortunately, our Magellan data do not reach that far into the blue.

Though the highest model with Exoretrievals was that of a featureless atmosphere, in order to compare results from both retrievals, we examine the Exoretrievals scattering model with the highest evidence ($\Delta \ln Z = -1.77$). This model is still indistinguishable from the featureless model. We find that all overlapping parameters agree with each other within their uncertainties. The planet's terminator temperature of $T_p = 1223^{+429}_{-292}$ K with Exoretrievals agrees with the calculated equilibrium temperature (1575 K; Anderson et al. 2011), and the T_p of 1219^{+262}_{-245} K retrieved with PLATON is nearly 1σ away from the calculated equilibrium temperature. A low retrieved terminator temperature is not uncommon, and an inherent bias is produced when using 1D transmission spectrum models for a 2D problem (MacDonald et al. 2020; Pluriel et al. 2020). However, using Equation (1) of López-Morales & Seager (2007) to calculate the equilibrium temperature of the planet, we find that the retrieved temperatures are still consistent if the planet has a bond albedo above 0.3.[28] This might be a slightly high albedo for a hot Jupiter but is completely reasonable relative to the solar system's Jovians (Rowe et al. 2008; Mallonn et al. 2019).

The log cloud top pressure of $-2.9^{+3.3}_{-3.2}$ bars with Exoretrievals and $-3.6^{+2.7}_{-2.1}$ bars with PLATON, while not well constrained, implies that there are high-altitude clouds. High-altitude clouds provide a good explanation for the data, given that features such as Na and K are not present, the possible water feature in the IR data is strongly muted, and the retrieved parameters have large uncertainties. But again, it cannot be confidently claimed that this is the case because of the lack of convergence of the cloud top pressure. The γ of $-3.8^{+3.1}_{-3.8}$ obtained with Exoretrievals and scattering slope of $5.3^{+2.9}_{-3.1}$ with PLATON are within 1σ of a Rayleigh scattering slope, which was detected at some level by previous studies (Sing et al. 2015; Gibson et al. 2017).

The highest $\Delta \ln Z$ transmission spectrum model retrieved is that of a scattering slope with PLATON and is overlaid on top of our combined Magellan/IMACS data in Figure 7, with corner plot results shown in Figure 12.

Previous detailed atmospheric studies on WASP-31b's atmosphere with data introduced in Sing et al. (2015) are

consistent with our retrieval analysis using IMACS data introduced in this manuscript. Specifically, Wakeford et al. (2017) fit models to the HST/STIS WASP-31b data and found that clouds made of enstatite ($MgSiO_3$) and iron act as a gray opacity source from log pressures of −4.1 to −1.7 in WASP-31. This agrees with the cloud deck pressure and muted features we retrieved. In fact, given the high uncertainties of our retrieved scattering slope, a near gray (little color-dependent scattering) cloud deck is also supported. Barstow et al. (2016) fit 3600 atmospheric models to the HST/STIS WASP-31 data and found a best fit for a gray cloud deck at 100 mbar, but there were a few high-evidence models with a Rayleigh slope and cloud top pressure as low as 0.01 mbar. Pinhas & Madhusudhan (2017) also found multiple indistinguishable fits for the HST/STIS WASP-31b data with their models but had a best-fit slope of -5.52 ± 1.27, which agrees with our best fits. These findings in conjunction with ours emphasize the difficulty in converging on the exact picture of WASP-31b's atmosphere, due to cloud formation.

5.2. Previous Data

In order to test the consistency of our IMACS observations and previous observations of WASP-31b, we ran the same retrieval analysis against the three separate optical data in the literature, which were the original HST/STIS (Sing et al. 2015), VLT/FORS2, and the combined FORS2 and reanalyzed STIS data (Gibson et al. 2017). We combined each of these subsets of data with the WFC3 and Spitzer IR data.

In summary, the VLT/FORS2 retrieval analysis produced nearly identical results to our IMACS analysis. That is, inconclusive support for a flat spectrum with Exoretrievals and a slightly favored, indistinguishable from a flat line ($\Delta \ln Z = 0.26$), near-Rayleigh optical scattering slope with Platon. Given the similar wavelength coverage of IMACS (0.4–0.96 μm) and FORS2 (0.39–0.85 μm) and their precision (0.0033 R_p/R_s), we expect to see this agreement.

We detected potassium ($\Delta \ln Z = 3.8$) and a scattering slope with Exoretrievals when performing retrieval analysis with the HST/STIS data. Interestingly, the best model with Platon was one with strong support ($\Delta \ln Z = 20.9$) toward only stellar activity, which is likely a portrayal of a degeneracy between scattering slope and stellar activity (Rackham et al. 2017).

The evidences obtained with Exoretrievals, utilizing the combined FORS2 and reanalyzed STIS data subset, favored the model with water and a scattering slope at $\Delta \ln Z = 2.5$. This data subset also moderately favored a scattering slope with Platon at $\Delta \ln Z = 4.2$.

With both the STIS and combined FORS2 and reanalyzed STIS data subsets, the retrieved scattering slopes were higher

---
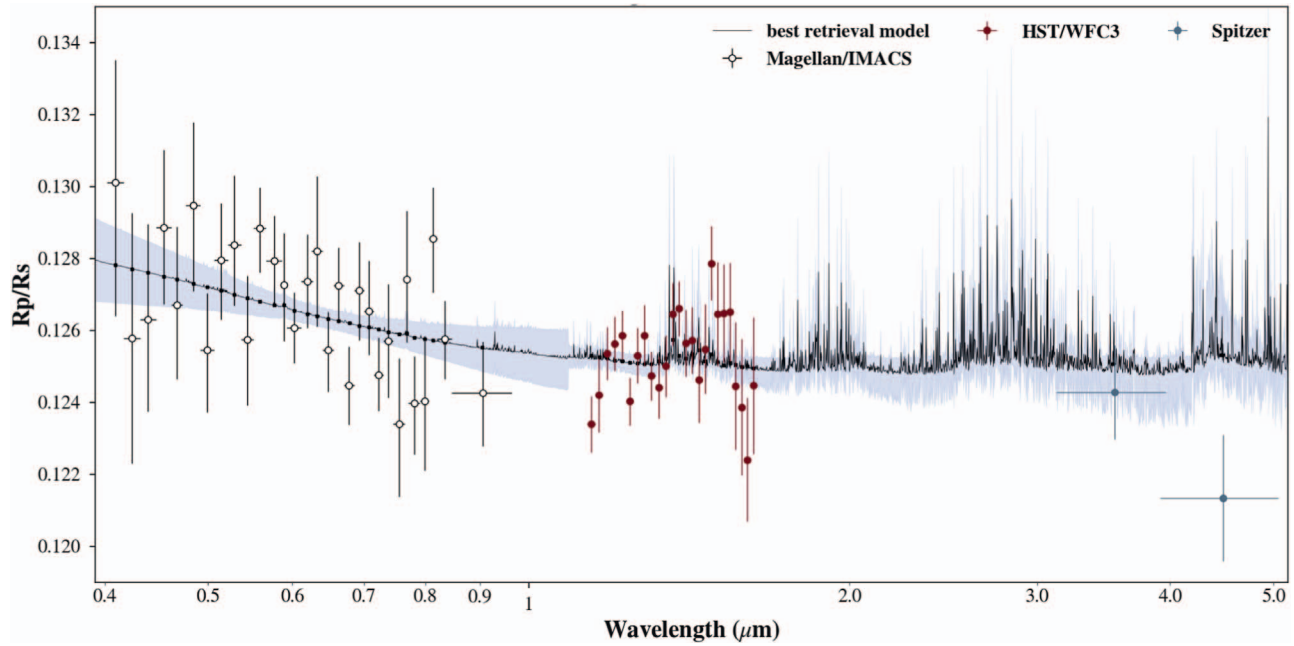[28] Assuming efficient heat distribution, $f = 1/4$, and using the retrieved system parameters.

Figure 7. Combined transmission spectrum of the three usable Magellan/IMACS optical transits, HST/WFC3, and Spitzer IR transits. With these data the best-fit model (highest $\Delta \ln Z$) was one utilizing `Platon` and obtained a scattering slope consistent with Rayleigh ($5.3^{+2.9}_{-3.1}$), log metallicity ($Z/Z_\odot$) of $1.3^{+1.1}_{-1.4}$, C/O ratio of $0.52^{+0.87}_{-0.32}$, planet limb temperature of 1220^{+260}_{-250} K, and log cloud top pressure of $-3.6^{+2.7}_{-2.1}$ bars. The solid black line is the retrieval's best fit (highest evidence), the shaded region is the 1σ confidence interval, and the colored dots with error bars are the data for various instruments.

($>2\sigma$) than Rayleigh. This is likely due to the fact that these data subsets include bluer (0.3–0.4 μm) wavelengths, which constrain the high scattering slope. We tested this by running the combined FORS2 and reanalyzed STIS data excluding the first two data points (centered at 3350 and 3859 Å) against `Platon`,[29] and we obtained a shallower slope that was consistent with Rayleigh scattering at 1σ.

6. Retrieval Interpretation

6.1. Optical Scattering Slope

As discussed in Sections 5.1 and 5.2, the Bayesian evidences of the FORS2 or IMACS data sets were unable to distinguish a flat spectrum from that with a scattering slope. The models that fit for a scattering slope with these data did retrieve slopes that were within Rayleigh, but those retrieved slopes had large uncertainties. In contrast, retrieval runs utilizing the STIS or combined FORS2 and reanalyzed STIS data subsets moderately detected a scattering slope. These slopes were above Rayleigh ($|\gamma| = \alpha = 8.3$) and better constrained with approximate error bars of ± 1.5. However, the STIS data set is also the only one where K and stellar activity are significantly retrieved, portraying our overall uncertainty of its retrieved parameters. Thus, it is unclear whether the slope anchored by the bluer STIS data points implies that WASP-31b does have a large scattering slope or that the STIS data set is an outlier. As such, we can only report that there are tentative signs of an optical Rayleigh scattering slope redward of 0.4 μm.

---

[29] `Platon` is only capable of calculating the spectrum from 0.3 to 30.0 μm. Therefore, we had to shift the central wavelength of the bluest data point by 50 Å. To do so, we did a linear interpolation between the three bluest bins, using the `linear` method of the `scipy.interpolate` package (Virtanen et al. 2020). We found a 45 ppm change in depth.

6.2. Sodium and Potassium Features

Of the four available WASP-31b optical data sets, only the STIS data had a Bayesian evidence that favored potassium absorption. As Gibson et al. (2017) imply, this is likely produced as a side effect of the detrending technique used. Another cause may be that the star was active during the HST observations, which Pont et al. (2013), Rackham et al. (2017), and Espinoza et al. (2019) have shown could mimic transmission spectral signals. As introduced in Section 5.2, the Bayesian evidence using `Platon` strongly favored a model characterizing the STIS transmission spectrum with stellar activity. This is supported by the AIT photometric monitoring that only showed signs of activity during the 2012–2013 season, which coincides with the STIS observations taken in 2012 June. However, Keck/HIRES observations of WASP-31's Ca II H and K emission lines suggest that the star was extremely quiet (log $R'_{HK} = -5.225$) when observed (2002–2009; Sing et al. 2016). Thus, it is possible, while not certain, that stellar activity could contribute to the cause of this unique transmission spectrum. No matter the cause of the potassium signal in the original STIS data, given that there are two independent low-resolution transmission spectra and a high-resolution transmission spectrum that show no signs of K, we conclude that WASP-31b shows no optical atomic absorption.

It could also be argued that because the K feature (7681 Å) is near prominent tellurics ($\sim$7580–7680 Å), ground-based observations would not be able to detect potassium. However, there are multiple cases where K is detected from the ground (Sing et al. 2011; Colón et al. 2012; Chen et al. 2018) with low resolution $R \sim 500$, and high resolution can deconvolve Earth's spectroscopic signal from a target's (Gibson et al. 2019). Thus, the Magellan/IMACS observations further paint the picture of WASP-31b lacking alkali absorption features. This is indicative of high-altitude clouds present in the planet's atmosphere, given that prominent Na and K absorption is expected in hot Jupiters like WASP-31b (Seager & Sasselov 2000). Authors such as

Morley et al. (2012) and Gao et al. (2020) have found that for gas giants at the temperature range of WASP-31b silicate clouds can heavily alter observed spectral features.

6.3. Water Features

None of the fits in Sections 5.1 and 5.2, aside from the FORS2 and reanalyzed STIS data subset, showed significant detection of H_2O. However, in general the models that include H_2O are more favored (see Table 5), suggesting that there are signs of water features in the IR data. This is consistent with the findings of Sing et al. (2015) and Stevenson (2016), where Stevenson (2016) found a weak $H_2O - J$ index (0.86 ± 0.48). Additionally, the retrieved C/O ratio with our Magellan/IMACS data of $0.52^{+0.87}_{-0.32}$, which is within uncertainties of the retrieved C/O ratios using the other data subsets, suggests an oxygen-rich atmosphere. The C/O ratio is relatively unconstrained, because the only observations that could constrain CO features are two wide wavelength data points from Spitzer. Nonetheless, this ratio and the relatively low temperature retrieved ($T_p = 1220^{+260}_{-250}$ K) imply that water would be a large absorber in WASP-31b's atmosphere (Madhusudhan 2012). Thus, the lack of significant detection of water is likely due to high-altitude clouds muting the water features.

6.4. Final Interpretation

The possible extremely muted water features, lack of optical absorption features from K and Na, and retrieved low pressures with all data subsets (see Sections 5.1 and 5.2) point toward WASP-31b possessing high-altitude clouds. Furthermore, comparing our spectra and retrieval results with atmospheric modeling by Wakeford & Sing (2015) and Pinhas & Madhusudhan (2017), the data suggest the following: (1) Their likely are a large number of particles of modal size between $\sim 3 \times 10^{-2}$ and 0.25 μm lofted high in WASP-31b's atmosphere. (2) The composition of these particles is predominately enstatite and other silicates. (3) These high-opacity structures are clouds, which is an accumulation of particles that condense in equilibrium or near-equilibrium chemistry, rather than hazes, which form from photochemistry or other nonequilibrium chemical processes. Given that this planet likely has clouds obscuring the optical and near-IR features, spectral observations in the IR (6–30 μm) could clarify the cloud composition and particle size distribution. As suggested by Wakeford & Sing (2015) and Pinhas & Madhusudhan (2017), JWST observations would be optimal for such clarity.

7. Summary and Conclusion

We observed four transit epochs of WASP-31b between 2013 and 2019 with the Magellan/IMACS spectrograph. From these observations we derived optical transmission spectra of the planet between 0.4 and 0.97 μm. We excluded the first transit (UT130226), which had inconsistent retrieved parameters, and combined the other three transits to produce a transmission spectrum with an average precision of $0.0033 R_p/R_s$ per 187 Å bin width. We combined our transmission spectrum with literature HST/WFC3 and Spitzer observations and modeled the full data sets using two independent retrieval codes. We find signs of a Rayleigh scattering slope and water features, but neither is significantly detected. We also find no Na or K signals likely owing to a high-altitude cloud deck. Our PLATON scattering slope model had the highest $\Delta \ln Z$ and retrieved a Rayleigh scattering

slope of $5.3^{+2.9}_{-3.1}$, log metallicity of $1.3^{+1.1}_{-1.4}$, C/O of $0.52^{+0.87}_{-0.32}$, and cloud top pressure of $-3.6^{+2.7}_{-2.1}$ bars.

In addition, we ran the retrievals against data from the literature (Sing et al. 2015; Gibson et al. 2017), which all showed weak signs of water absorption likely muted by high-altitude clouds. The retrieved scattering slopes for the data sets including the bluer STIS observations were steeper than Rayleigh by over 2σ. This could mean that WASP-31b's scattering slope is beyond Rayleigh, but we acknowledge that there could be other causes of this retrieved slope. Of all available data on WASP-31b, only the original STIS transit showed strong K absorption.

From this work we find that the atmosphere of WASP-31b is likely to be significantly covered by high-altitude clouds, explaining high uncertainties of parameters and low significance of feature detection.

It is imperative to understand which planets form high-altitude clouds or hazes, and how. Understanding this will be key in the JWST era. We need a large library of thoroughly analyzed optical-to-IR transmission spectra for this. In particular, the optical component allows for absolute abundances of molecules retrieved in the IR by constraining the level at which the molecular features are muted by clouds. Such a library would allow for better comparative studies on observed spectral features and other system parameters (e.g., stellar irradiation levels, spectral type, mass, radius, planet density) to find correlations and will drastically enhance the field of exoplanetology. Uniform data sets for a wide range of planets, such as the ones ACCESS is building, will be crucial for such analyses.

Appendix A
Light Curves

The detrended, binned light curves of transit UT130226 (Figure 8), UT130425 (Figure 9), UT140222 (Figure 10), and UT190314 (Figure 11) are shown below. Additionally the residuals of the detrended data from the best-fit transit model are shown in their adjacent figures. Table 6 shows the transit depths corresponding to Figures 8–11, and Table 7 shows the weighted mean depths per each bin.
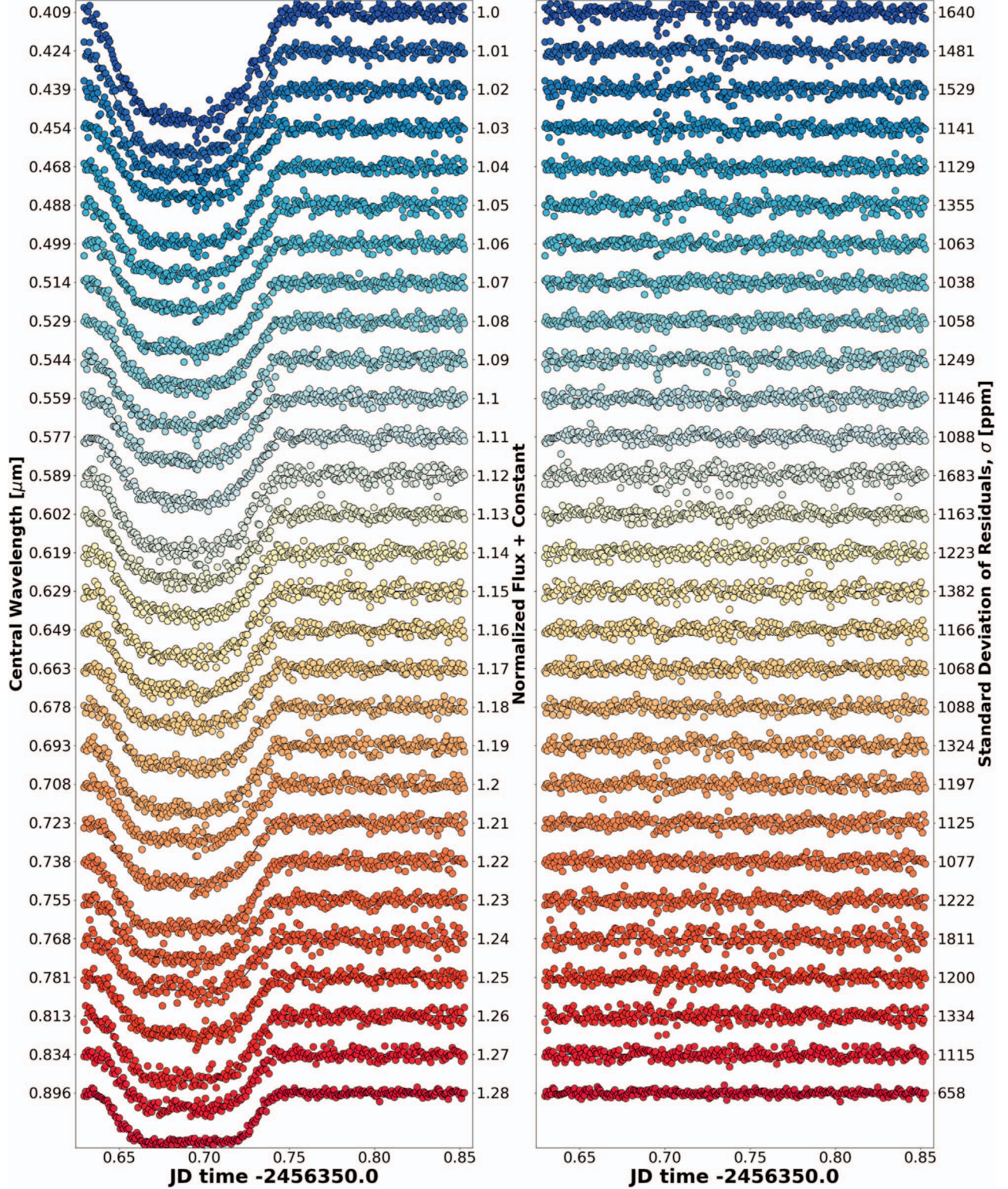

Figure 8. Detrended, binned light curves of transit UT130226 (left) with residuals of the detrended data from the best-fit transit model on the right. The first axis on the left panel displays the mean wavelength of that bin in microns, and the second axis displays the light curves' normalized flux, which were incrementally offset by 0.01 for visual purposes. The axis on the right panel displays the standard deviation (σ) of the residuals in ppm. Each bin was detrended with model averaging of PCA and GP machine-learning techniques. These light curves encompass the spectrum from 4018 to 9450 Å, aside from the spaces caused by chip gaps.
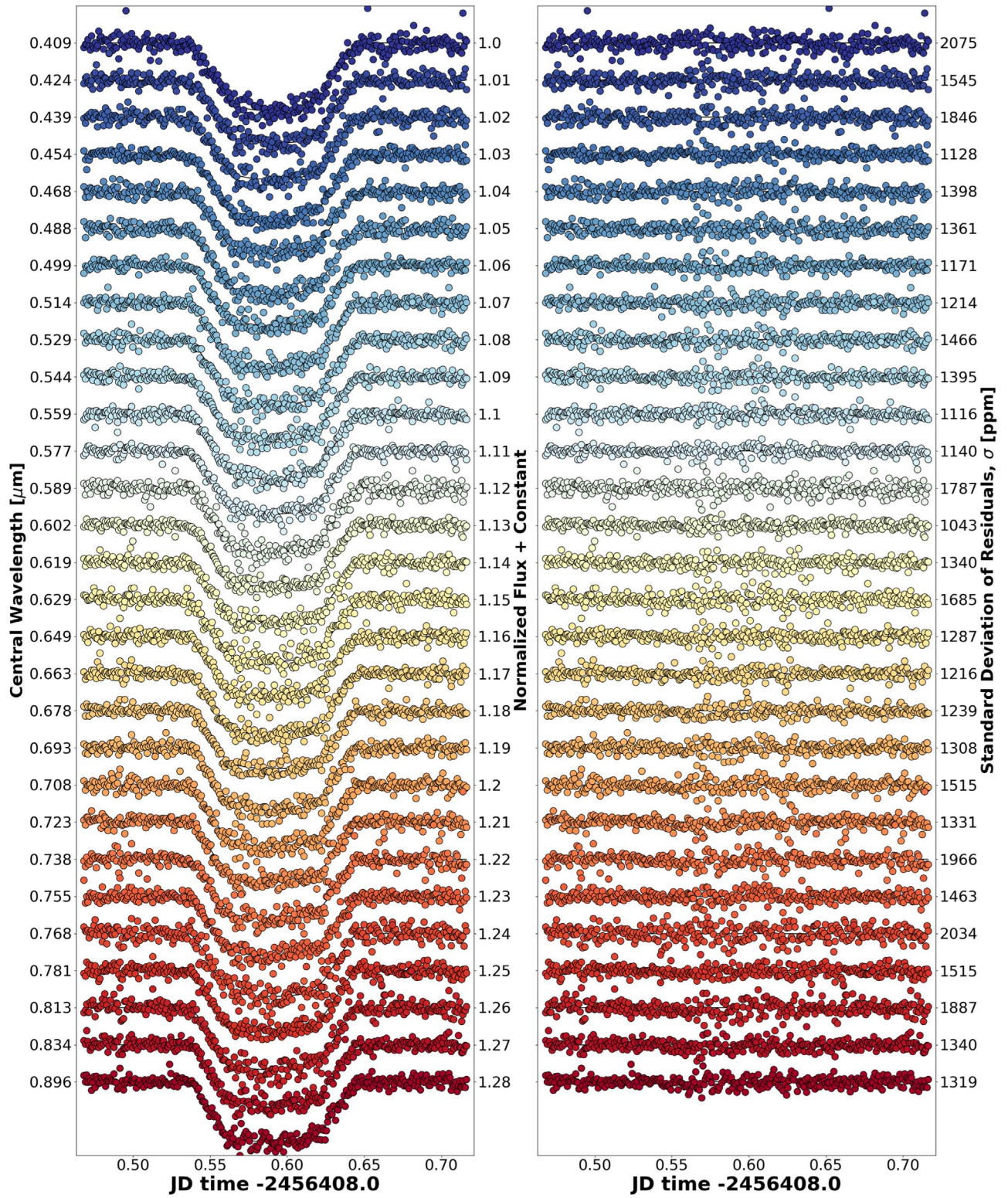
Figure 9. Same as Figure 8, but for transit UT130425.
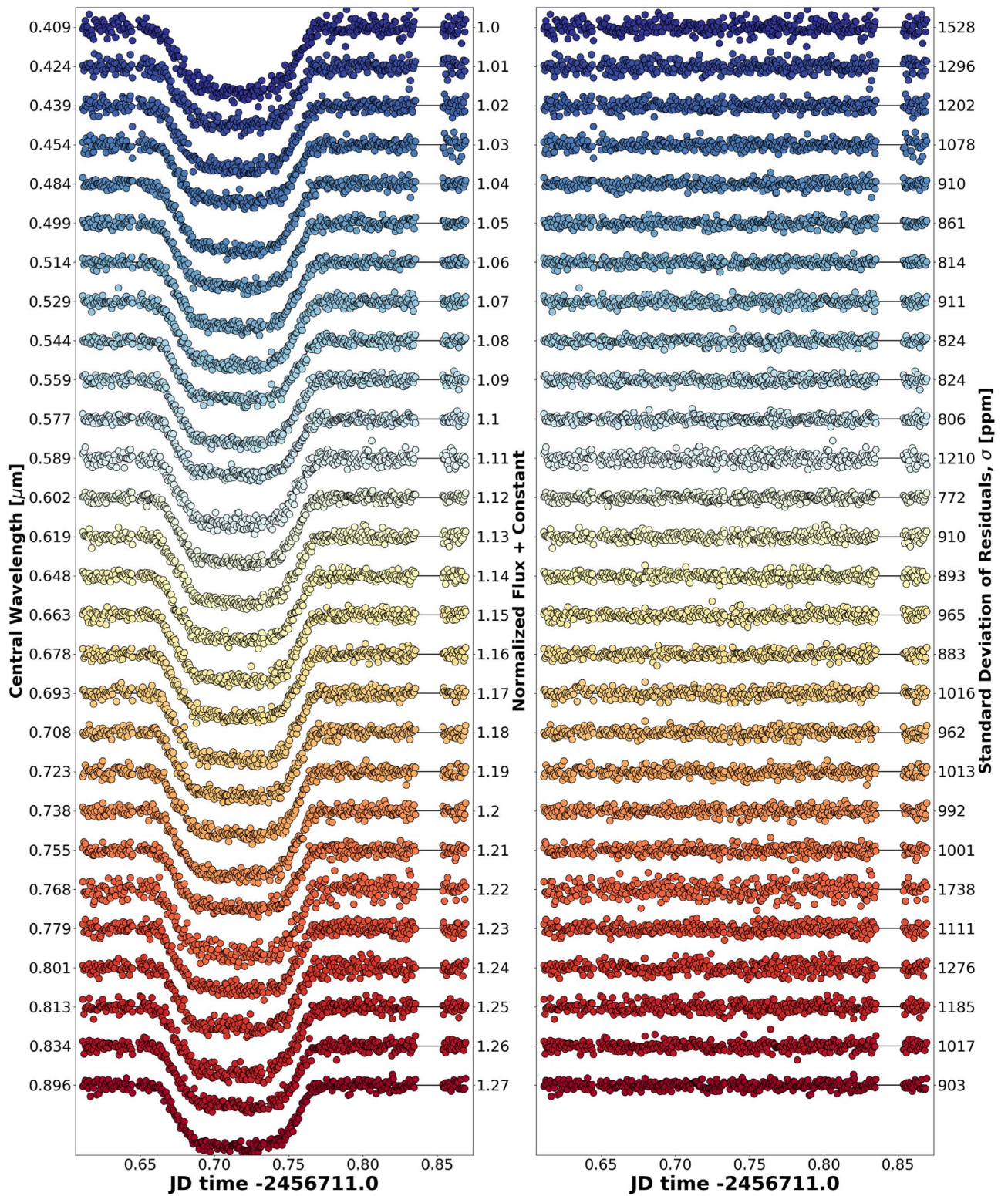
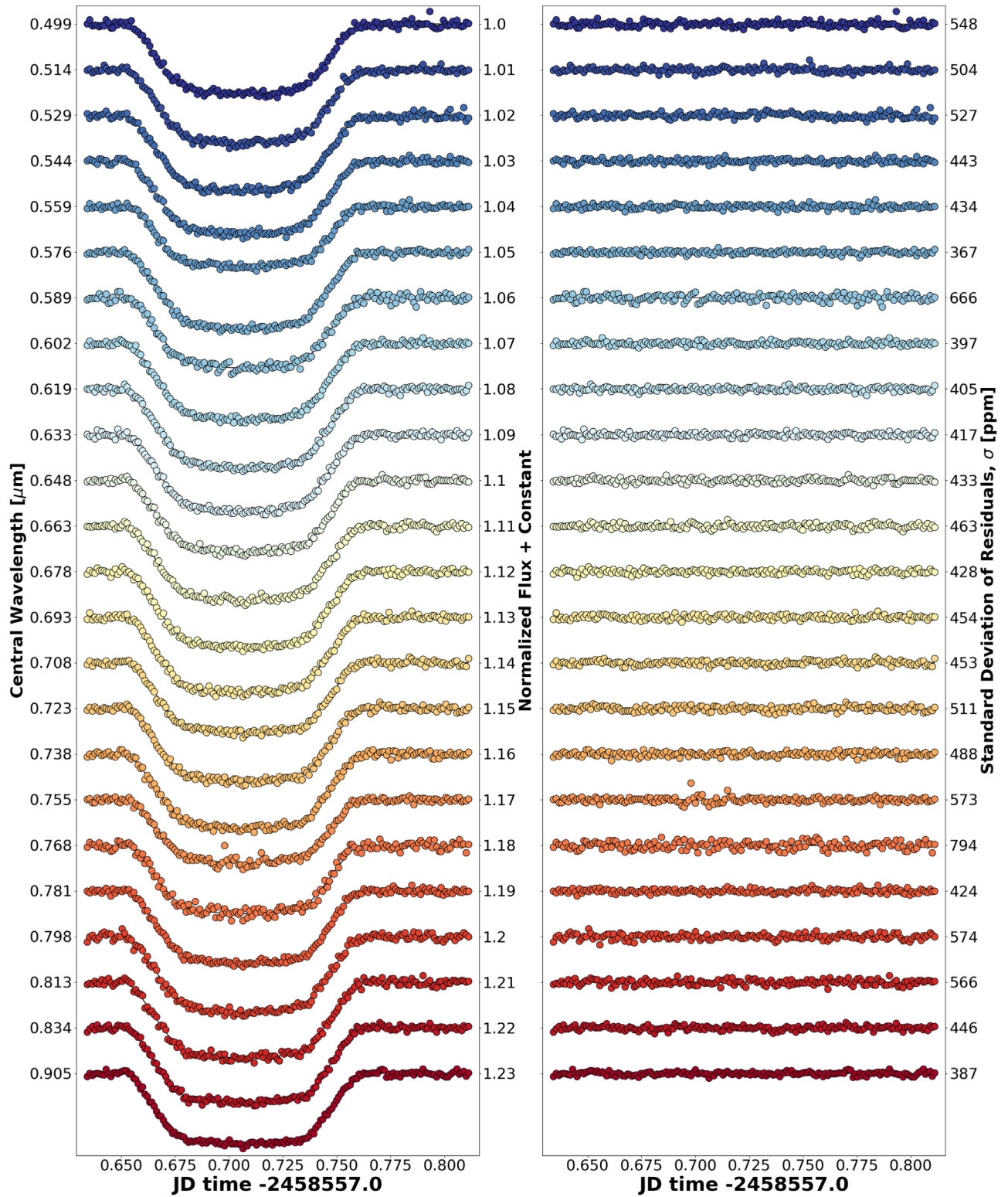**Figure 10.** Same as Figure 8, but for transit UT140222.

Figure 11. Same as Figure 8, but for transit UT190314. Because of the IMACS chip gaps, this particular data set fully encompasses the spectrum from 4918 to 9670 Å.

Table 6
Magellan/IMACS Optical Transmission Spectrum (Planet Radius/Star Radius, p) for Each of Our Four Transit Epochs: UT130226, UT130425, UT140222, and UT190314, Respectively

Wavelength Range (Å)	UT130226 R_p/R_s	UT130425 R_p/R_s	UT140222 R_p/R_s	UT190314 R_p/R_s
4018.0–4168.0	$0.1673^{+0.0071}_{-0.0076}$	0.1361 ± 0.0045	$0.1292^{+0.0047}_{-0.0065}$	⋯
4168.0–4318.0	$0.1632^{+0.0063}_{-0.0066}$	$0.1315^{+0.0043}_{-0.0039}$	$0.1216^{+0.0061}_{-0.0071}$	⋯
4318.0–4468.0	$0.1511^{+0.0067}_{-0.0062}$	$0.1305^{+0.0031}_{-0.0029}$	$0.1268^{+0.0049}_{-0.0052}$	⋯
4468.0–4618.0	$0.1347^{+0.0042}_{-0.0043}$	$0.1341^{+0.0024}_{-0.0023}$	$0.1203^{+0.0051}_{-0.0055}$	⋯
4618.0–4733.0	$0.1454^{+0.0064}_{-0.006}$	$0.1299^{+0.0022}_{-0.002}$	⋯	⋯
4768.0–4918.0	⋯	⋯	$0.1302^{+0.0033}_{-0.0037}$	⋯
4835.0–4918.0	$0.1351^{+0.005}_{-0.0062}$	$0.1343^{+0.0029}_{-0.003}$	⋯	⋯
4918.0–5068.0	$0.1312^{+0.0032}_{-0.0035}$	0.1287 ± 0.0019	$0.1267^{+0.0031}_{-0.0046}$	$0.1302^{+0.0044}_{-0.0049}$
5068.0–5218.0	$0.1329^{+0.0046}_{-0.0047}$	$0.1323^{+0.0028}_{-0.0029}$	$0.1293^{+0.0028}_{-0.003}$	0.1312 ± 0.0023
5218.0–5368.0	$0.1329^{+0.0036}_{-0.0039}$	$0.1338^{+0.0033}_{-0.0036}$	$0.1301^{+0.0043}_{-0.0056}$	$0.1308^{+0.0027}_{-0.002}$
5368.0–5518.0	$0.1318^{+0.0043}_{-0.0047}$	$0.1314^{+0.0026}_{-0.0027}$	$0.1233^{+0.0037}_{-0.0039}$	$0.1293^{+0.0032}_{-0.0033}$
5518.0–5668.0	0.1284 ± 0.0041	$0.1337^{+0.0016}_{-0.0017}$	$0.131^{+0.0014}_{-0.002}$	0.1222 ± 0.0051
5668.0–5862.9	$0.1303^{+0.0025}_{-0.0024}$	$0.1315^{+0.0018}_{-0.0019}$	$0.1202^{+0.0047}_{-0.0045}$	0.1323 ± 0.0018
5862.9–5922.9	$0.141^{+0.0065}_{-0.0082}$	0.132 ± 0.0021	$0.1306^{+0.0024}_{-0.003}$	$0.1254^{+0.0032}_{-0.0034}$
5922.9–6118.0	$0.1295^{+0.0073}_{-0.0062}$	$0.1295^{+0.0019}_{-0.0017}$	$0.1286^{+0.0014}_{-0.0016}$	$0.1297^{+0.0019}_{-0.0018}$
6118.0–6256.0	$0.1286^{+0.0033}_{-0.0037}$	$0.1295^{+0.0021}_{-0.0019}$	$0.1309^{+0.002}_{-0.0019}$	$0.1317^{+0.003}_{-0.0031}$
6256.0-6333.0	$0.133^{+0.0046}_{-0.0054}$	$0.1307^{+0.0031}_{-0.003}$	⋯	⋯
6256.0–6406.0	—-	⋯	⋯	$0.1316^{+0.0028}_{-0.0029}$
6406.0–6556.0	⋯	⋯	$0.1293^{+0.0014}_{-0.0017}$	$0.1265^{+0.0024}_{-0.0026}$
6426.0–6556.0	$0.1282^{+0.0045}_{-0.0042}$	$0.1285^{+0.002}_{-0.0019}$	⋯	⋯
6556.0–6706.0	$0.1218^{+0.0039}_{-0.0049}$	$0.1295^{+0.0023}_{-0.0022}$	0.1302 ± 0.0013	$0.1322^{+0.0028}_{-0.003}$
6706.0–6856.0	$0.1261^{+0.0035}_{-0.0041}$	0.1264 ± 0.0015	$0.1279^{+0.0018}_{-0.0023}$	$0.1299^{+0.0022}_{-0.0023}$
6856.0–7006.0	$0.1316^{+0.0043}_{-0.005}$	$0.1311^{+0.0034}_{-0.0033}$	$0.1315^{+0.0017}_{-0.002}$	$0.128^{+0.0024}_{-0.0022}$
7006.0–7156.0	$0.1224^{+0.0039}_{-0.0048}$	$0.1301^{+0.0031}_{-0.0029}$	$0.1301^{+0.0018}_{-0.0015}$	$0.1283^{+0.0026}_{-0.0023}$
7156.0–7306.0	$0.127^{+0.0046}_{-0.005}$	$0.1272^{+0.0016}_{-0.0015}$	0.1278 ± 0.0022	$0.1287^{+0.0016}_{-0.0015}$
7306.0–7456.0	$0.1322^{+0.0032}_{-0.003}$	$0.1276^{+0.0035}_{-0.0033}$	$0.1288^{+0.0021}_{-0.002}$	$0.1307^{+0.0034}_{-0.0037}$
7456.0–7651.25	$0.1244^{+0.0042}_{-0.0045}$	0.1278 ± 0.0022	$0.1234^{+0.0035}_{-0.0042}$	$0.1221^{+0.0079}_{-0.0121}$
7651.25–7711.25	$0.1164^{+0.0134}_{-0.008}$	0.1319 ± 0.0025	$0.1312^{+0.003}_{-0.0027}$	$0.1195^{+0.006}_{-0.0064}$
7711.25-7860.0	⋯	⋯	$0.1277^{+0.002}_{-0.0029}$	⋯
7711.25–7906.0	$0.1229^{+0.0037}_{-0.0041}$	0.1276 ± 0.0021	⋯	$0.1259^{+0.0028}_{-0.0026}$
7906.0–8056.0	⋯	⋯	⋯	$0.1296^{+0.0028}_{-0.0027}$
7959.0–8056.0	⋯	⋯	$0.1256^{+0.0021}_{-0.0024}$	⋯
8056.0–8206.0	$0.1297^{+0.0041}_{-0.004}$	$0.1311^{+0.0023}_{-0.0022}$	$0.1315^{+0.002}_{-0.0022}$	$0.1331^{+0.0031}_{-0.0039}$
8206.0–8466.0	$0.119^{+0.0041}_{-0.0047}$	$0.128^{+0.0016}_{-0.0015}$	$0.1247^{+0.0026}_{-0.0052}$	0.1308 ± 0.0016
8466.0-9450.0	$0.1142^{+0.0024}_{-0.0026}$	$0.1285^{+0.0024}_{-0.0021}$	$0.1269^{+0.0033}_{-0.0052}$	⋯
8466.0–9640.0	⋯	⋯	⋯	$0.125^{+0.0025}_{-0.0023}$

Note. The data were produced implementing the reduction and detrending processes discussed in Section 3. These depths do not include the offsets for combining each spectra. Gaps in the spectra (see Figure 1) prevented a few bins from exactly overlapping in wavelength space. Those bins were still weighted averaged together (x and y direction), with the resulting bin composed of the full wavelength width of all bins. In this table, each bin for the combined transmission spectrum is separated by a vertical line.

(This table is available in its entirety in machine-readable form.)

Table 7
Combined Transmission Spectrum from Transit Epochs UT130425,
UT140222, and UT190314

Wavelength Range (Å)	Mean (Å)	R_p/R_s
4018.0–4168.0	4093.00	$0.1334^{+0.1297}_{-0.1367}$
4168.0–4318.0	4243.00	$0.1288^{+0.1254}_{-0.1323}$
4318.0–4468.0	4393.00	$0.1295^{+0.1269}_{-0.1321}$
4468.0–4618.0	4543.00	$0.1319^{+0.1297}_{-0.134}$
4618.0–4733.0	4675.50	$0.1299^{+0.1279}_{-0.1321}$
4768.0–4918.0	4868.65	$0.1326^{+0.1303}_{-0.1348}$
4918.0–5068.0	4993.00	$0.1285^{+0.1269}_{-0.1301}$
5068.0–5218.0	5143.00	$0.131^{+0.1294}_{-0.1325}$
5218.0–5368.0	5293.00	$0.1316^{+0.1299}_{-0.1334}$
5368.0–5518.0	5443.00	$0.1289^{+0.1271}_{-0.1307}$
5518.0–5668.0	5593.00	$0.1319^{+0.1306}_{-0.1329}$
5668.0–5862.9	5765.30	$0.1311^{+0.1298}_{-0.1323}$
5862.9–5922.9	5892.90	$0.1303^{+0.1288}_{-0.1317}$
5922.9–6118.0	6020.45	$0.1292^{+0.1282}_{-0.1302}$
6118.0–6256.0	6188.78	$0.1305^{+0.1292}_{-0.1318}$
6256.0–6406.0	6294.62	$0.1312^{+0.1291}_{-0.1333}$
6406.0–6556.0	6485.00	$0.1285^{+0.1274}_{-0.1296}$
6556.0–6706.0	6631.00	$0.1303^{+0.1293}_{-0.1314}$
6706.0–6856.0	6781.00	$0.1276^{+0.1265}_{-0.1286}$
6856.0–7006.0	6931.00	$0.1303^{+0.1289}_{-0.1316}$
7006.0–7156.0	7081.00	$0.1296^{+0.1285}_{-0.131}$
7156.0–7306.0	7231.00	$0.128^{+0.127}_{-0.129}$
7306.0–7456.0	7381.00	$0.1289^{+0.1274}_{-0.1305}$
7456.0–7651.25	7553.62	$0.1266^{+0.1247}_{-0.1284}$
7651.25–7711.25	7681.25	$0.1306^{+0.1288}_{-0.1324}$
7711.25–7906.0	7798.01	$0.1272^{+0.1258}_{-0.1285}$
7906.0–8056.0	7999.69	$0.1272^{+0.1254}_{-0.1289}$
8056.0–8206.0	8131.00	$0.1316^{+0.1301}_{-0.1329}$
8206.0–8466.0	8336.00	$0.129^{+0.1279}_{-0.13}$
8466.0–9640.0	8982.70	$0.1269^{+0.1254}_{-0.1284}$

Note. The spectra were combined following the procedure outlined in Section 4.1. In the retrieval analysis, because the retrievals do not take asymmetric wavelength errors, the wavelength ranges of the few overlapping bins were recentered based on the weighted mean wavelength of the bin.

(This table is available in its entirety in machine-readable form.)

Appendix B
Transit UT130226

As can be seen in Figure 5, the transit depths of epoch UT130226's transmission spectrum varied widely over our wavelength coverage. Additionally, as can be seen in Figure 4, there was little baseline before ingress during this observation, which could cause improper detrending of systematics. We tested the impact of our insufficient baseline by comparing the transit duration of epoch UT130226 with those of other epochs when all transit parameters, aside from p, were fixed. We also compared the transit depth obtained for this epoch between this test case (with fixed parameters) and our nominal case (with other transit parameters free). For both cases, the transit durations and depths agreed within 1σ, suggesting that the baseline was sufficient. This is not conclusive support of sufficient baseline, but reasoning to explore whether the transmission spectrum is indeed caused by the system.

We used our retrieval framework to analyze just IMACS transit UT130226 and the IR data to further explore the cause of the outlier transmission spectrum.[30] It is possible that the transit can be explained by unique physical scenarios in the system during the single epoch, such as the host star being relatively active or the planet having an exceptionally large amount of scattering agents lofted high in the atmosphere. We tested the plausibility of these scenarios by running a variety of retrievals. Table 8 shows $\Delta \ln Z$ for these retrievals compared to a featureless spectrum. The evidences for both PLATON and Exoretrievals strongly support stellar activity and a large scattering slope ($\alpha = 9.61^{+0.28}_{-0.46}$ and $\gamma = -9.50^{+0.67}_{-0.35}$).

The retrieved large scattering slopes are inconsistent with the other Magellan/IMACS and VLT/FORS2 observations (discussed in Sections 5). Additionally, the best-fit terminator

---

[30] The signal of one Magellan/IMACS transit (0.007 R_p/R_s) is not enough to detect features like Na or K ($\sim$0.006R_p/R_s, Sing et al. 2015), but sufficient to interpret the large slope in the spectrum.

Table 8
Same as Table 5, but with the Subset of Data That Included Only the UT130226 Magellan/IMACS and IR Data

	Exoretrievals					PLATON	
Model	Featureless	H_2O	Na	K	$H_2O + K + Na$	Model	
Featureless	0.00	−0.70	−0.60	−1.21	−0.81	Featureless	0.0
Scatterers	⋯	14.53	14.13	14.00	13.66	Scattering	24.67
Activity	4.08	10.04	3.02	3.40	9.45	Stellar activity	12.01
Scatterers and activity	⋯	16.11	15.08	14.72	15.17	Both	22.93

Note. The retrievals with a strong scattering slope ($\gamma \sim -9.5$, $\alpha \sim 9.6$) and stellar activity ($T_{het} \sim 5700$ K and $f_{het} \sim 0.4$, $T_{spot} \sim 5900$ K and $f_{spot} \sim 0.3$) were heavily supported by the Exoretrievals and PLATON evidences, respectively.

temperatures of 1805^{+66}_{-126} K for Exoretrievals and 1846^{+35}_{-69} K for PLATON were high relative to the other IMACS and FORS2 data ($\sim$1200 K) and the calculated equilibrium temperature (1575 K; Anderson et al. 2011).

An exceptionally high level of stellar activity during the transit of WASP-31b also seems like an improbable scenario, given that our photometeric time-series analysis and previous Ca II H and K spectroscopic analysis (Sing et al. 2016) show signs of low-level activity. With Exoretrievals the

best-fit stellar activity model was that where $T_{het} = 5720 \pm 310$ K and $f_{het} = 0.4^{+0.28}_{-0.20}$, and for PLATON it was where $T_{spot} = 5950^{+240}_{-380}$ K and $f_{spot} = 0.31^{+0.28}_{-0.24}$. These high spot covering fractions imply a completely contrary level of activity from other activity indicators. Thus, we have no concrete explanation of this transmission spectrum's abnormal features and decided to exclude this transit from our combined data in order to prevent biasing our atmospheric interpretations.

Appendix C
Corner Plots

The corner plot of the highest evidence retrieval model with `PLATON` is shown in Figure 12. Though this particular model had the highest evidence, it is not significantly more favorable from the other models.
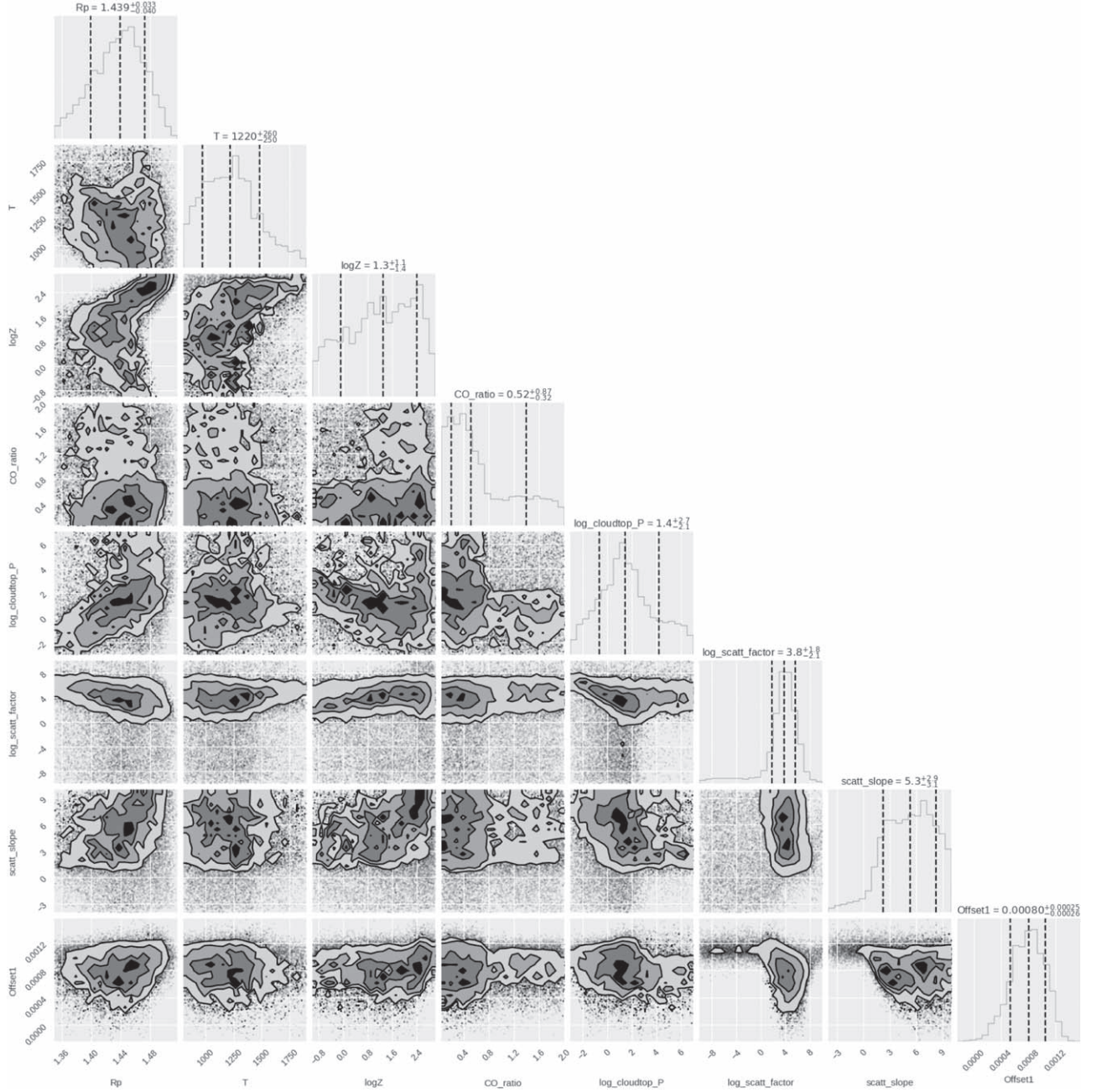

Figure 12. Corner plot of the best-fit retrieval on the combined Magellan/IMACS, HST/WFC3, and Spitzer data. This is using the `PLATON` retrieval, and its corresponding transmission spectrum is shown in Figure 7. Vertical dashed lines mark the 16% and 84% quantiles.

ORCID iDs

Chima D. McGruder ⊙ https://orcid.org/0000-0002-6167-3159
Mercedes López-Morales ⊙ https://orcid.org/0000-0003-3204-8183
Néstor Espinoza ⊙ https://orcid.org/0000-0001-9513-1449
Benjamin V. Rackham ⊙ https://orcid.org/0000-0002-3627-1676
Dániel Apai ⊙ https://orcid.org/0000-0003-3714-5855
Andrés Jordán ⊙ https://orcid.org/0000-0002-5389-3944
Munazza K. Alam ⊙ https://orcid.org/0000-0003-4157-832X
Alex Bixel ⊙ https://orcid.org/0000-0003-2831-1890
Jonathan J. Fortney ⊙ https://orcid.org/0000-0002-9843-4354
Gregory W. Henry ⊙ https://orcid.org/0000-0003-4155-8513
James Kirk ⊙ https://orcid.org/0000-0002-4207-6615
Nikole K. Lewis ⊙ https://orcid.org/0000-0002-8507-1304
Florian Rodler ⊙ https://orcid.org/0000-0003-0650-5723
Ian C. Weaver ⊙ https://orcid.org/0000-0001-6205-6315

References

Aigrain, S., Pont, F., & Zucker, S. 2012, MNRAS, 419, 3147
Anderson, D. R., Collier Cameron, A., Hellier, C., et al. 2011, A&A, 531, A60
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Barstow, J. K., Aigrain, S., Irwin, P. G. J., & Sing, D. K. 2016, ApJ, 834, 50
Benneke, B., & Seager, S. 2013, ApJ, 778, 153
Berta, Z. K., Charbonneau, D., Bean, J., et al. 2011, ApJ, 736, 12
Bixel, A., Rackham, B. V., Apai, D., et al. 2019, AJ, 157, 68
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Charbonneau, D., Berta, Z. K., Irwin, J., et al. 2009, Natur, 462, 891
Chen, G., Pallé, E., Welbanks, L., et al. 2018, A&A, 616, A145
Colón, K. D., Ford, E. B., Redfield, S., et al. 2012, MNRAS, 419, 2233
Diamond-Lowe, H., Berta-Thompson, Z., Charbonneau, D., & Kempton, E. M. R. 2018, AJ, 156, 42
Dressler, A., Bigelow, B., Hare, T., et al. 2011, PASP, 123, 288
Duvenaud, D. 2014, PhD thesis, Univ. Cambridge, https://doi.org/10.17863/CAM.14087
Espinoza, N. 2017, PhD thesis, Pontificia Universidad Católica de Chile, https://repositorio.uc.cl/handle/11534/21313
Espinoza, N., & Jordán, A. 2016, MNRAS, 457, 3573
Espinoza, N., Rackham, B. V., Jordán, A., et al. 2019, MNRAS, 482, 2065
Evans, T. M., Sing, D. K., Goyal, J. M., et al. 2018, AJ, 156, 283
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Foreman-Mackey, D. 2016, JOSS, 1, 24
Foreman-Mackey, D., Hoyer, S., Bernhard, J., & Angus, R. 2014, George: George (V0.2.0), v0.2.0, Zenodo, doi:10.5281/zenodo.11989
Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, ApJ, 678, 1419
Gao, P., Thorngren, D. P., Lee, G. K. H., et al. 2020, NatAs, 4, 951
Gibson, N. P. 2014, MNRAS, 445, 3401
Gibson, N. P., Aigrain, S., Roberts, S., et al. 2012, MNRAS, 419, 2683
Gibson, N. P., de Mooij, E. J. W., Evans, T. M., et al. 2019, MNRAS, 482, 606
Gibson, N. P., Nikolov, N., Sing, D. K., et al. 2017, MNRAS, 467, 4591
Gillon, M., Jehin, E., Lederer, S. M., et al. 2016, Natur, 533, 221
Hastie, T., Tibshirani, R., & Friedman, J. 2007, The Elements of Statistical Learning (2nd ed.; Berlin: Springer)
Heng, K., & Kitzmann, D. 2017, MNRAS, 470, 2972

Henry, G. W. 1999, PASP, 111, 845
Hunter, J. D. 2007, CSE, 9, 90
Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6
Jolliffe, I. 2002, Principal Component Analysis (New York: Springer)
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python, http://www.scipy.org/
Jordán, A., Espinoza, N., Rabus, M., et al. 2013, ApJ, 778, 184
Kirk, J., López-Morales, M., Wheatley, P. J., et al. 2019, AJ, 158, 144
Kochanek, C. S., Shappee, B. J., Stanek, K. Z., et al. 2017, PASP, 129, 104502
Kreidberg, L. 2015, PASP, 127, 1161
Lomb, N. R. 1976, Ap&SS, 39, 447
López-Morales, M., Haywood, R. D., Coughlin, J. L., et al. 2016, AJ, 152, 204
López-Morales, M., & Seager, S. 2007, ApJL, 667, L191
Luger, R., Sestovic, M., Kruse, E., et al. 2017, NatAs, 1, 0129
Lupu, R. E., Zahnle, K., Marley, M. S., et al. 2014, ApJ, 784, 27
MacDonald, R. J., Goyal, J. M., & Lewis, N. K. 2020, ApJL, 893, L43
MacDonald, R. J., & Madhusudhan, N. 2017, MNRAS, 469, 1979
Madhusudhan, N. 2012, ApJ, 758, 36
Mallonn, M., Herrero, E., Juvan, I. G., et al. 2018, A&A, 614, A35
Mallonn, M., Köhler, J., Alexoudi, X., et al. 2019, A&A, 624, A62
Markevich, N., & Gertner, I. 1989, NIMPA, 283, 72
McCullough, P. R., Crouzet, N., Deming, D., & Madhusudhan, N. 2014, ApJ, 791, 55
Micchelli, C., Xu, Y., & Zhang, H. 2006, JMLR, 7, 2651
Morley, C. V., Fortney, J. J., Marley, M. S., et al. 2012, ApJ, 756, 172
Oliphant, T. E. 2006, Guide to NumPy, https://numpy.org/
Pinhas, A., & Madhusudhan, N. 2017, MNRAS, 471, 4355
Pluriel, W., Zingales, T., Leconte, J., & Parmentier, V. 2020, A&A, 636, A66
Pont, F., Sing, D. K., Gibson, N. P., et al. 2013, MNRAS, 432, 2917
Rackham, B., Espinoza, N., Apai, D., et al. 2017, ApJ, 834, 151
Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, ApJ, 853, 122
Rackham, B. V., Apai, D., & Giampapa, M. S. 2019, AJ, 157, 96
Rasmussen, C., & Williams, C. 2005, Gaussian Processes for Machine Learning (Cambridge, MA: MIT Press)
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, Proc. SPIE, 9143, 914320
Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, JQSRT, 111, 2139
Rowe, J. F., Matthews, J. M., Seager, S., et al. 2008, ApJ, 689, 1345
Ryabchikova, T., Piskunov, N., Kurucz, R. L., et al. 2015, PhyS, 90, 054005
Scargle, J. D. 1982, ApJ, 263, 835
Seager, S., & Sasselov, D. D. 2000, ApJ, 537, 916
Sedaghati, E., Boffin, H. M. J., MacDonald, R. J., et al. 2017, Natur, 549, 238
Sing, D. K., Désert, J. M., Fortney, J. J., et al. 2011, A&A, 527, A73
Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, Natur, 529, 59
Sing, D. K., Wakeford, H. R., Showman, A. P., et al. 2015, MNRAS, 446, 2428
Southworth, J., Mancini, L., Madhusudhan, N., et al. 2017, AJ, 153, 191
Speagle, J. S. 2020, MNRAS, 492, 3132
Stevenson, K. B. 2016, ApJL, 817, L16
Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., et al. 2016, JMoSp, 327, 73
Thomas-Osip, J. E., Prieto, G., Johns, M., & Phillips, M. M. 2008, Proc. SPIE, 7012, 70121U
Trotta, R. 2008, ConPh, 49, 71
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Wakeford, H. R., & Sing, D. K. 2015, A&A, 573, A122
Wakeford, H. R., Stevenson, K. B., Lewis, N. K., et al. 2017, ApJL, 835, L12
Weaver, I. C., López-Morales, M., Espinoza, N., et al. 2020, AJ, 159, 13
Yan, F., Espinoza, N., Molaverdikhani, K., et al. 2020, A&A, 642, A98
Yurchenko, S. N., Tennyson, J., Barber, R. J., & Thiel, W. 2013, JMoSp, 291, 69
Zacharias, N., Finch, C. T., Girard, T. M., et al. 2013, AJ, 145, 44
Zhang, M., Chachan, Y., Kempton, E. M. R., & Knutson, H. A. 2019, PASP, 131, 034501