

Carbon monoxide emission lines reveal an inverted atmosphere in the ultra hot Jupiter WASP-33 b consistent with an eastward hot spot

Lennart van Sluijs,^{1,2★} Jayne L. Birkby^{1,2,3}, Joshua Lothringer,⁴ Elspeth K. H. Lee,⁵ Ian J. M. Crossfield,⁶ Vivien Parmentier,⁷ Matteo Brogi,^{8,9,10} Craig Kulesa,¹¹ Don McCarthy¹¹ and David Charbonneau³

¹Department of Astrophysics, University of Oxford, Oxford OX1 3RH, UK

²Anton Pannekoek Institute for Astronomy, University of Amsterdam, NL-1098 XH Amsterdam, the Netherlands

³Center for Astrophysics, Harvard and Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

⁴Department of Physics, Utah Valley University, Orem, UT 84058, USA

⁵Center for Space and Habitability, University of Bern, Gesellschaftsstrasse 6, CH-3012 Bern, Switzerland

⁶Department of Physics and Astronomy, University of Kansas, Lawrence, KS 66045, USA

⁷Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford OX1 3PU, UK

⁸Department of Physics, University of Warwick, Coventry CV4 7AL, UK

⁹INAF – Osservatorio Astrofisico di Torino, Via Osservatorio 20, I-10025 Pino Torinese, Italy

¹⁰Dipartimento di Fisica, Università degli Studi di Torino, via Pietro Giuria 1, I-10125 Torino, Italy

¹¹Steward Observatory, University of Arizona, Tucson, AZ 85721-0065, USA

Accepted 2023 April 5. Received 2023 March 10; in original form 2022 March 24

ABSTRACT

We report the first detection of CO emission at high spectral resolution in the day-side infrared thermal spectrum of an exoplanet. These emission lines, found in the atmosphere of the transiting ultra hot Jupiter (UHJ) WASP-33 b, provide unambiguous evidence of its thermal inversion. Using spectra from the MMT Exoplanet Atmosphere Survey (MEASURE, $R \sim 15\,000$), covering pre- and post-eclipse phases, we cross-correlate with 1D PHOENIX spectral templates to detect CO at $S/N = 7.9$ ($v_{\text{sys}} = 0.15_{-0.65}^{+0.64}$ km s⁻¹, $K_p = 229.5_{-1.0}^{+1.1}$ km s⁻¹). Moreover, using cross-correlation-to-log-likelihood mapping, we find that the scaling parameter which controls the spectral line contrast changes with phase. We thus use the general circulation model SPARC/MITgcm post-processed by the 3D gCMCRT radiative transfer code to interpret this variation, finding it consistent with an eastward-shifted hot spot. Pre-eclipse, when the hot spot faces Earth, the thermal profiles are shallower leading to smaller line contrast despite greater overall flux. Post-eclipse, the western part of the day-side faces Earth and has much steeper thermal profiles, leading to larger line contrast despite less overall flux. This demonstrates that within the log-likelihood framework, even relatively moderate resolution spectra can be used to understand the 3D nature of close-in exoplanets, and that resolution can be traded for photon-collecting power when the induced Doppler-shift is sufficiently large. We highlight CO as a good probe of UHJ thermal structure and dynamics that does not suffer from stellar activity, unlike species that are also present in the host star e.g. iron lines.

Key words: techniques: spectroscopic – planets and satellites: atmospheres – planets and satellites: fundamental parameters.

1 INTRODUCTION

Ultra hot Jupiters (UHJs) provide the unique opportunity to study gaseous planets in a strongly irradiated environment to learn about their composition and global circulation patterns (e.g. for a review see Showman, Tan & Parmentier 2020; Fortney, Dawson & Komacek 2021). Their high day-side temperatures above 2200 K provide a unique window to directly detect volatile species in their vapour phase (e.g. Visscher, Lodders & Fegley 2010; Hoeijmakers et al. 2018; Kitzmann et al. 2018; Lothringer, Barman & Koskinen 2018; Parmentier et al. 2018; Hoeijmakers et al. 2019; Ehrenreich et al. 2020; Cont et al. 2021; Merritt et al. 2021). Due to their tidally locked

rotation, they have an extreme day-to-night-side temperature contrast (e.g. Knutson et al. 2007). Consequently, global circulation models (GCMs) predict strong global jets and hot spots (Menou & Rauscher 2009; Showman et al. 2009; Parmentier et al. 2018; Tan & Komacek 2019). This combined with their relatively high star-to-planet contrast ratio, makes them ideal targets for atmospheric characterization with current observing facilities (e.g. Birkby 2018; Kreidberg 2018).

Hydrodynamic simulations and theoretical calculations predict that hot spots occur in ultra hot Jupiters with an eastward offset from the sub-stellar point due to global wind circulation patterns (e.g. Showman & Guillot 2002; Dobbs-Dixon & Lin 2008; Menou & Rauscher 2009; Dobbs-Dixon, Cumming & Lin 2010; Rauscher & Menou 2010; Showman & Polvani 2011; Perez-Becker & Showman 2013; Debras et al. 2020). Most observations support this eastward offset hot spot prediction (Harrington et al. 2006; Cowan, Agol &

* E-mail: lennart.vansluijs@physics.ox.ac.uk

Charbonneau 2007; Knutson et al. 2007; Charbonneau et al. 2008; Knutson et al. 2009; Swain et al. 2009; Crossfield et al. 2010; Wong et al. 2016), but for some UHJs westward offsets have been observed as well (Armstrong et al. 2016; Dang et al. 2018; Bell et al. 2019; Jackson et al. 2019; von Essen et al. 2020; Herman et al. 2022). Several mechanisms including cloud asymmetries, asynchronous rotation, and magnetohydrodynamical effects have been proposed to explain these westward offset hot spots (e.g. Hindle, Bushby & Rogers 2021).

Thermal inversions of the pressure–temperature profiles (P–T profiles) of UHJs were at first predicted due to the strong optical/UV absorbing molecules TiO and VO (e.g. Hubeny, Burrows & Sudarsky 2003). Later work found absorption by other atomic and molecular species, in addition to TiO and VO, can also result in thermal inversions (Arcangeli et al. 2018; Lothringer et al. 2018; Gandhi & Madhusudhan 2019). This is due to a combination of short-wavelength stellar irradiation around early-type stars and absorption of these wavelengths by continuous opacity sources, metal atoms, metal hydrides, metal oxides, SiO, or possibly even disequilibrium species like SH (Evans et al. 2018). Analysis of *Spitzer* and *Hubble Space Telescope (HST)* secondary eclipse observations support this scenario as they detect emission signatures suggestive of an inverted stratosphere for a handful of UHJs (e.g. Deming et al. 2012; Haynes et al. 2015; Evans et al. 2016, 2017; Sheppard et al. 2017; Arcangeli et al. 2018; Kreidberg et al. 2018; Mansfield et al. 2018; Garhart et al. 2020; Nugroho et al. 2020a; Yan et al. 2020; Baxter et al. 2021).

One well-studied UHJ for which a thermal inversion was predicted by atmospheric modelling is WASP-33 b. This planet orbits every 1.22 d around an A5 star (Collier Cameron et al. 2010). The first observational evidence of a thermal inversion of WASP-33 b's atmospheric profile was excess infrared emission observed by *Spitzer* from secondary eclipse observations (Deming et al. 2012; Zhang et al. 2018) and NIR low-resolution HST/Wide Field Camera 3 spectra, possibly due to TiO (Haynes et al. 2015). Furthermore, Zhang et al. (2018) find a phase angle offset by $-12.8 \pm 5.8^\circ$ in the *Spitzer* 3.6 μm -band, consistent with an eastward hot spot. On the contrary, von Essen et al. (2020) find a $28.7 \pm 7.1^\circ$ phase angle offset in the optical with *TESS*, suggesting a westward offset hot spot instead, although they mention host star variability may introduce a spurious westward offset. Recently, Herman et al. (2022) report day-to-night brightness contrast variations and a $22 \pm 12^\circ$ westward phase offset from their detection of Fe-I emission lines.

For these low resolution observations, stellar pulsations of the δ Scuti pulsating host star induce variability of the stellar continuum (Herrero et al. 2011), limiting observational constraints as they induce a quasi-sinusoidal trend in transit light-curve observations (e.g. Deming et al. 2012; de Mooij et al. 2013; von Essen et al. 2019, 2020; Garhart et al. 2020). Furthermore, stellar pulsations also affect high resolution spectral observations targeting any molecule present in both the stellar and exoplanetary atmosphere due to changes in the spectral line shape (Herrero et al. 2011). For WASP-33 b this has been observed for the detection of neutral iron lines by Nugroho et al. (2020b); Herman et al. (2022).

Pursuits to follow-up the tentative low-resolution detection of TiO in the atmosphere of WASP-33 b using high resolution spectra have lead to inconclusive results so far. Nugroho et al. (2017) detected the optical TiO molecular signature using the High Dispersion Spectrograph on the Subaru Telescope. This detection was later challenged by a re-analysis of the same data by Serindag et al. (2021) who unexpectedly found a slightly weaker signal using the updated TiO ExoMol TOTO line list. Furthermore, Herman et al. (2020) analysed emission and transmission spectra from the

Echelle SpectroPolarimetric Device for the Observation of Stars (ESPaDOnS) on the Canada–France–Hawaii Telescope and the High Resolution Echelle Spectrometer (HIRES) on the Keck telescope, but are unable corroborate their observations with a thermal inversion due to TiO. There is thus a lack of a current consensus regarding the observational evidence for TiO in the atmosphere of WASP-33 b.

Despite the elusive results for TiO, several new detections suggest WASP-33 b must have a thermally inverted atmosphere. Recently, emission signatures of Fe I (Nugroho et al. 2020b; Cont et al. 2021), neutral Si (Cont et al. 2022), and OH (Nugroho et al. 2021) have been detected. Nugroho et al. (2021) also marginally detect H₂O, consistent with the theoretical predictions that H₂O molecules dissociate into OH and H⁻ at the temperature regime of UHJs. Other atmospheric detections for WASP-33 b include Ca II (Yan et al. 2019), the hydrogen H α , H β , and H γ Balmer lines (Borsa et al. 2021; Cauley et al. 2021; Yan et al. 2021), evidence for AlO (von Essen et al. 2019), and lastly, Kesseli et al. (2020) place upper limits on the Volume Mixing Ratio (VMR) for FeH based on their null detection. These detections are thus in line with predictions by Lothringer et al. (2018); Gandhi & Madhusudhan (2019) of thermally inverted UHJ atmospheres, regardless of the presence of TiO or VO.

In this paper, we report CO emission lines in the atmosphere of WASP-33 b using observations with the MMT in Arizona, USA, equipped with the ARizona Infrared imager and Echelle Spectrograph (ARIES). It is the first reported detection of CO emission lines detected with high resolution cross-correlation spectroscopy (HRCCS). During the refereeing process of this paper, Yan et al. (2022) independently reported a detection of CO emission lines in the atmosphere of WASP-33 b using GIANO-B, and CO emission lines are also subsequently reported in the atmospheres of WASP-121 b Holmberg & Madhusudhan (2022) and WASP-18 b Brogi et al. (2022). However, we note that our detection here, which uses the large Doppler-shift induced by the planet's orbital motion, is at lower spectral resolution compared to other HRCCS observations, and still reveals information about the planet's atmospheric dynamics. The observations and the subsequent data reduction are discussed in Section 2. Our atmospheric models are described in Section 3. These models are used to characterize the exoplanet's atmosphere using HRCCS as described in Section 4. The results are presented in Section 5 and discussed in Section 6. Our main conclusions and recommendations for future work are summarized in Section 7.

2 OBSERVATIONS AND DATA PROCESSING

2.1 Observations

We observed the ultra hot Jupiter WASP-33 b with the ARizona Infrared imager and Echelle Spectrograph (ARIES; McCarthy et al. 1998; Sarlot et al. 1999) in combination with the f/15 adaptive secondary mirror and adaptive optics system at the 6.5-m MMT Observatory on Mt Hopkins in Arizona, USA. The adaptive secondary mirror provides a low thermal background, while augmenting the total throughput of the instrument. ARIES can provide spectral observations using both long slit and echelle spectrograph modes. It can observe in the 1–5 μm range and at spectral resolutions of 2000–30 000. The observations of WASP-33 b presented in this work are the first of a larger survey called MEASURE (MMT Exoplanet Atmosphere Survey) which contains observations of a diverse set of eleven exoplanets with a wide range of temperatures, masses, and radii.

For these observations of WASP-33 b, we obtained echelle spectra in the 1.37–2.56 μm wavelength range during three half nights in 2016 October using the ARIES/MMT combination. In total 211

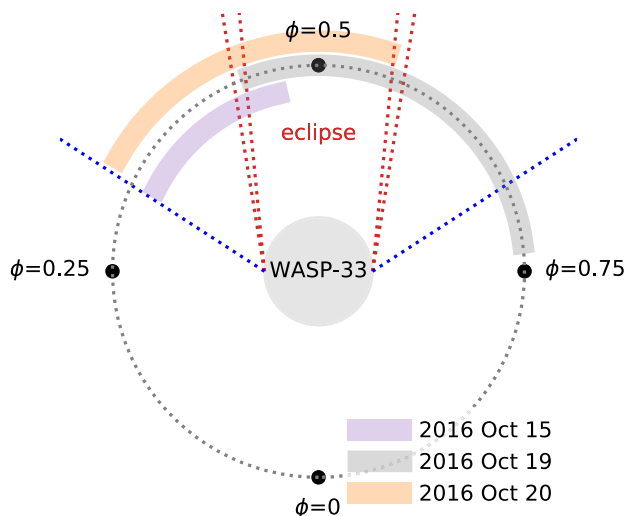


Figure 1. Orbital phase (ϕ) coverage for each observing night of the WASP-33 data set. As can be seen, all nights contain some spectra taken during full eclipse of WASP-33 b. The blue dotted line indicates up to which phases we have symmetric phase coverage for both the pre- and post-eclipse data.

Table 1. Observational parameters of the WASP-33 data set for each night.

Date (UTC)	Total exposure time (min)	No. spectra	Phase coverage	S/N*
2016 Oct 15	220	44	0.331–0.469	91
2016 Oct 19	465	93	0.461–0.731	26
2016 Oct 20	370	74	0.337–0.554	71

Note. *This quantity refers to the median S/N per wavelength bin as measured from the extracted and normalized spectral time series of the 23rd spectral order around the $2.3 \mu\text{m}$ CO emission region.

spectra were obtained with an exposure time of 300 s per frame at a theoretical instrumental resolution of $R = 30\,000$ using the $1 \text{ arcsec} \times 0.2 \text{ arcsec}$ slit and the $f/5.6$ camera mode. The simultaneously operated ARIES imager was used for in-slit guiding, with occasional manual offsets made to correct any sustained drift during an exposure not corrected by the AO guiding system.

An overview of our observed orbital phase coverage¹ is shown in Fig. 1 and our observational parameters and data quality are summarized in Table 1. The layout of the 26 echelle orders of ARIES is shown in the top left-hand panel of Fig. 2. The three observing nights cover both pre- and post-eclipse orbital phases, but the post-eclipse data covers significantly more of the orbit, almost to quadrature, where the planet shows half its day-side and half its night-side to the Earth.

At the beginning and end of each half night, we observed a set of calibration images. This included a set of dark frames obtained using a blank in the filter wheel and the ARIES entrance covered to prevent any light entering the instrument. We observed flat-fields with the grating in, using an incandescent light bulb arranged such that its light reflected off an aluminumized board into the dichroic of ARIES. We also observed a thorium–argon lamp in a similar manner at the start of the nights to focus the spectrograph, but used the observed telluric absorption lines for simultaneous wavelength calibration.

¹Orbital phases have been calculated using the ephemerides in Table A1, which contains all relevant WASP-33 system parameters used throughout this work.

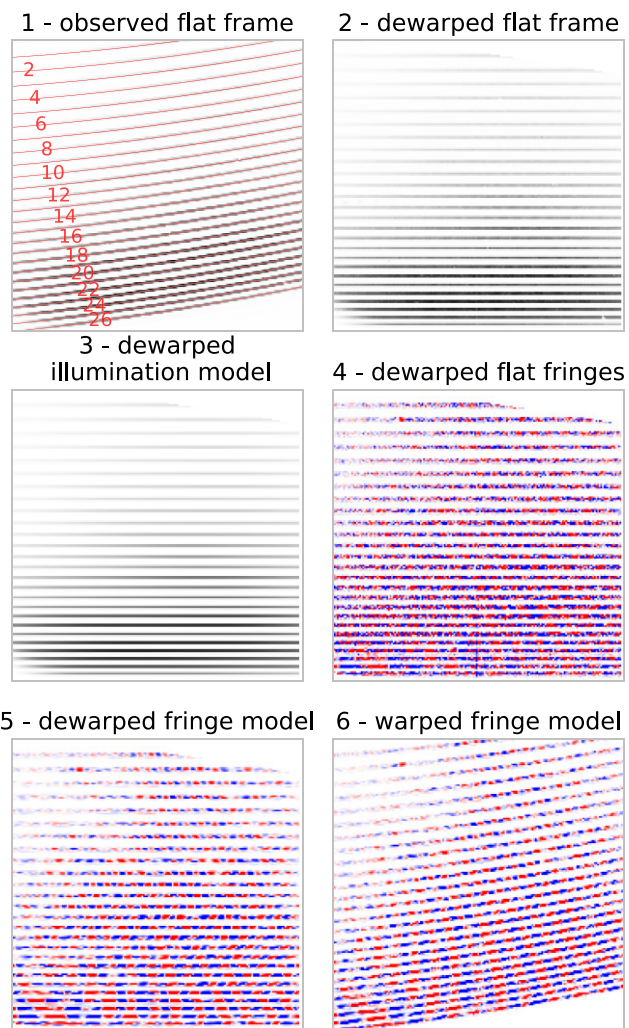


Figure 2. Flat fringe correction illustrated in six steps: (1) the observed flat frame with the fitted echelle traces overlaid in red (numbered from blue-to-red). (2) the flat frame after applying a 2D polynomial transform to straighten the echelle traces. (3) the 4th order polynomial row-by-row illumination model. (4) fringes in the dewarped flat frame. (5) row-by-row low-pass filter fringe model. (6) fringe model after applying inverse 2D polynomial transform.

2.2 Data reduction: pre-processing

Since this is the first time the MMT/ARIES combination has been used for exoplanet atmosphere characterization, we developed a purpose-built end-to-end data reduction pipeline that we describe here first. This new pre-processing pipeline is publicly available on the author’s GitHub.² As ARIES is currently being upgraded into MAPS (Morzinski et al. 2020), the instrument is no longer available in this exact form, but our data reduction pipeline uses a generalized approach which may be useful to reduction of the instrument when it returns. The procedure is described below, and is applied in the same way to all three nights of data:

(i) Cross-talk correction: ARIES experiences cross-talk between the four quadrants of its detector which needs to be corrected. A bright pixel in one quadrant produces distinct, cross-shaped, negative

²<https://github.com/lennartvansluijs/ARIES>

shadows at the same position in all other three quadrants on the detector. This effect is apparent in all dark, flat, science frames. We use the C-based CORQUAD routine provided by the ARIES team.³ Three input parameters define the convolution kernel size for each detector quadrant. The parameter space is explored using a grid-based search to find the best cross-talk convolution kernels, which we define as those that minimize the standard deviation in a 10×10 box around a visually identified prominent shadow feature. We apply the CORQUAD routine using the best convolution kernels parameters to all dark, flat, and science frames, before any further data reduction.

(ii) Dark correction: we need to remove hot pixels from individual darks and combine them into a master dark. We identify hot pixels as $\geq 3\sigma$ -outliers. Like the shadow cross-talk features, they are cross-shaped and we replace them by the median of their neighbours outside of a cross-shaped footing in a 11×11 box around each hot pixel. We median combine all frames of identical exposure time into a master dark. We subtract the appropriate master dark from each science and flat frame.

(iii) Flat correction: all flats contain fringes arising from Fabry–Pérot interference between the optical elements of the instrument. Fringes in the flat field introduce modulation in the continuum that would be erroneously divided into the science flat during flat field correction, reducing the sensitivity of our later procedures. To remove the fringes from the flats, we developed a modified version of the flat fringe correction by Stone et al. (2014). The procedure is shown in Fig. 2. We use the *getthem* routine from the publicly available PYTHON package CERES⁴ (Brahm, Jordán & Espinoza 2017) to fit the echelle traces as 4th order polynomials using a 10 pixel wide aperture. We use the solution of all 26 spectral orders to calculate a 2D polynomial transform to dewarp the flat frame. We fit each row by a 7th order polynomial to create a dewarped illumination model, essentially representing the dewarped blaze function. This model is subtracted from the dewarped flat frame to reveal the dewarped flat fringes. Most fringes have frequencies $< 0.025 \text{ pixel}^{-1}$. We apply a 6th order Butterworth high-pass filter with a cutoff-frequency $f_c = 0.025 \text{ pixel}^{-1}$ to create a dewarped fringe model. We then use the inverse 2D polynomial transform to warp the fringes model. We remove the fringes from the observed flat frame by subtracting the warped fringes model. It is key this subtraction is done in the observed coordinate frame such that the observed flat was not interpolated by a transform operation, thus preserving any intrapixel systematics.

After defringing all flats, we perform a hot pixel correction similar to the one done for the dark frames and create a master flat. The ARIES detector has a central column of dead pixels. We correct this column by replacing it by the median of the two neighbouring columns in the master flat and science frames. To create a master flat hot pixel map, we first normalize the master flat by dividing the illumination model and identifying pixels with values outside of the $[0.5, 1.5]$ range, as hot pixels. We combine the master flat hot pixel map with the master dark hot pixel map and replace all master flat hot pixels by the fit from the illumination model. For each science frame we correct each flagged hot pixel by their closest neighbours in the dispersion direction of the spectrograph. Finally, the master flat is divided out of all of the science frames. The science frames still contain fringes at this stage, but these will be removed by applying a high-pass-filter to the extracted spectral time series in the post-processing stage of the pipeline.

³See <http://66.194.178.32/~rfinn/pisces.html> to correct for the cross-talk. A PYTHON-based version is available at: <https://github.com/jordan-stone/ARIES>.

⁴<https://github.com/rabrahm/ceres>

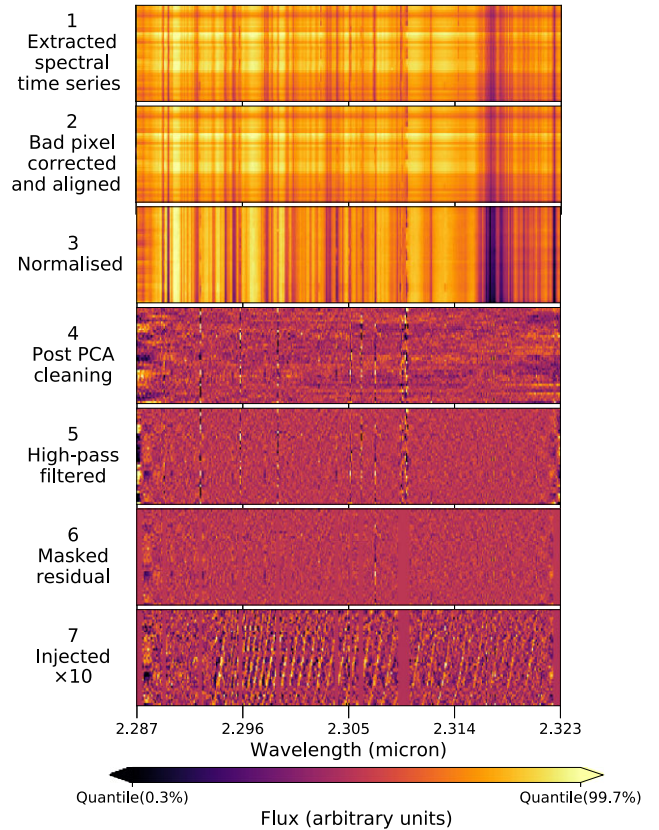


Figure 3. Example of post-processing procedure of a spectral time series for order 23 for the first observing night. First panel: The extracted spectral time series from the processed detector image is shown in the top panel. Second panel: after bad pixel/column correction and alignment with respect to the first spectrum. Third panel: normalized by the median of the 50 brightest pixels per spectrum. Fourth panel: after removing the first seven components using principal component analysis. Fifth panel: After applying a high-pass-filter with cutoff-frequency of 0.02 pixel^{-1} . Sixth panel: After masking the columns with the worst residual structure. Seventh panel: same as the previous panel, but with a $\times 10$ PHOENIX self-consistent model injected into the spectral time series between panels 2 and 3.

(iv) Spectral order extraction: after the flat correction we can extract the spectral orders from the detector images. We cannot directly fit the detector echelle traces for the science frames as these traces are often disrupted due to telluric absorption lines. Instead we fit the echelle traces using the CERES RETRACE routine. This treats the initial set of master flat traces as an aperture mask and cross-correlates it with each science frame to determine the drift in the cross-dispersion direction. We use the CERES MARSH-algorithm (Marsh 1989) routine with a 5 pixel aperture radius and clipping $> 5\sigma$ outliers. The final output of the pre-processing part of the data reduction pipeline is a spectral time series for each of the 26 spectral orders for each of the 3 observing nights of the system (see example for one order in the top panel of Fig. 3).

2.3 Data reduction: post-processing

From the spectral time series we would like to extract the planet spectrum, but it is faint and buried in the noise. Thus we first need to remove the telluric and stellar lines. An overview of all steps of the post-processing procedure for a single spectral order covering the most line dense part of the CO spectrum is shown in Fig. 3. We

perform the following steps to post-process the extracted spectral time series for each spectral order in a homogeneous manner:

(i) Bad pixel/column correction (see Fig. 3, panel 2): in each spectral time series, some bad pixel columns or detector artefacts may still be present in the spectral time series. To reveal these anomalies, we normalize each spectra by the median of each row, followed by subtracting the mean of each columns.

We use a Blob Detection Algorithm (an edge-detection algorithm), which uses a Laplacian of the Gaussian Filter to detect outliers, as described in more detail by Kong, Akakin & Sarma (2013). In our application we identify features of the size of a single pixel. These outliers are corrected by interpolating in the dispersion direction over the closest nearby good pixels. We ignore any convolution boundary effects here as they only affect the edge pixels and these will be clipped during the post-processing stage of the pipeline.

Bad columns in the spectral time series are identified by taking the median of each spectral channel and applying a median-filter using a three pixel width and subtract this from the original. As before we use the Blob Detection Algorithm to identify 5σ -outliers in the residual time series which are considered bad columns and are corrected by interpolation from their nearest neighbouring good columns. We iteratively perform the bad pixel and column correction procedure five times to ensure all outliers are corrected. In total only ~ 0.2 per cent of all pixels are identified as outliers.

(ii) Alignment of spectra (see Fig. 3, panel 2): effects such as instrumental flexure, variation of the airmass or misalignment of the star on the spectrograph slit, can all cause sub-pixel drift of the spectra across the detector. We need to align the spectra to correct for these effects, else they will leave residuals later in the telluric line removal procedure. We did not fit a telluric model to each spectrum directly, as this adds the complication that the telluric model needs to match the observing conditions for each frame well. Instead we use the data themselves as a reference: the first spectrum of every night is used as a reference spectrum, thus the alignment will be in the telluric rest frame of the first spectrum. The alignment procedure is done in three steps. First, the reference spectrum is cross-correlated with each frame. Secondly, we fit a Gaussian to the cross-correlation function (CCF) to determine the drift on a subpixel level for each frame. Finally, we then shift each spectra using linear interpolation to align them with the reference spectra.

(iii) Telluric wavelength calibration: after alignment, we perform wavelength calibration using the observed telluric lines. We generate a telluric model using ATRAN⁵ (Lord 1992). For each aligned spectral order we cross-match by-eye telluric lines in the model to the corresponding telluric lines in the median of our observed spectra. We robustly fit a polynomial to all these points for each order to obtain a wavelength solution. Robust here refers to the fact that we iteratively clip $\geq 5\sigma$ outliers and refitting a polynomial until no more outliers are found. By default we fit third order polynomials, but for some orders where there are few absorption lines towards the edges of the detector, we fitted second order polynomials instead, to avoid massive deviation of the wavelength solution towards the edges. All these fits were manually inspected to ensure a proper wavelength solution was obtained. For some orders the wavelength solutions were poorly constrained due to strong residual science fringes and/or weak telluric absorption features. For these reasons we excluded orders 7–9 and 11 from the rest of the analysis. The wavelength solution and root mean square (RMS) deviation for each spectral order is shown in Fig. A1 for the first observing night.

(iv) Throughput correction (see Fig. 3, panel 3): throughput variations can be caused by variations in airmass and misalignment of the target on the slit. We correct for these variations by dividing by the mean of the brightest 50 pixels per spectrum following the procedure described in Brogi & Line (2019).

(v) Telluric removal (see Fig. 3, panel 4–5): quasi-stationary trends mostly due to the telluric lines are still dominating the spectra. To remove these, principal component analysis (PCA) is used to decompose the spectral time series. In our implementation, we perform a singular value decomposition (similar to de Kok et al. 2013; Line et al. 2021). Visual inspection of other ARIES observations from other systems indicates that using seven principal components works well to balance telluric removal and retrieve an injected exoplanet signal at a high S/N (as described in Section 4.1), so we adopt this also for WASP-33 b. However, we emphasize that we keep the number PCA iterations fixed for all spectral orders and all observing nights to avoid optimization of order/night specific systematic effects and noise, which can produce artificially large S/N (e.g. Cabot et al. 2020; Spring et al. 2022).

Remaining low-frequency residuals persist due to several effects: (1) throughput variations due to small offsets of the target position on the spectrographs slit entrance; (2) residual airmass variations; (3) echelle traces drift on the detector due to the changing gravity vector during the observations, causing instrumental flexure; and (4) uncorrected science frame fringes. We apply a sixth order Butterworth high-pass filter with a cutoff frequency of 0.02 pixel^{-1} to remove most of these trends.

(vi) Masking (see Fig. 3, panel 6): remaining bad columns in the residual time series are masked. This is done by identifying $>3\sigma$ outliers in the residual matrix and flagging columns with more than five outliers. We combine flagged columns if there are more than two columns within a five pixel-wide sliding window. Additionally, the columns in a 50 pixel window from the edge of the detector are masked by default.

The final result of the post-processing pipeline is shown in panel 6 of Fig. 3.

2.4 Observed instrument performance

As this is the first time ARIES/MMT is used to characterize an exoplanet atmosphere, we compare the observed performance to the theoretical instrumental performance. We measured the instrumental throughput, resolving power, and precipitable water vapour (PWV) directly from the normalized spectral time series. The instrumental throughput is measured directly from the total S/N for each spectra. To measure the observed resolving power and PWV we largely follow procedure described in Chiavassa & Brogi (2019), where we cross-correlate a telluric model with the normalized spectral time series and evaluate the log-Likelihood defined by Zucker (2003). This method is sensitive to continuum variations and first we correct for them the following way: (1) a median filter removes any possible outliers, (2) each spectrum is binned and the maximum of each bin is calculated, (3) we fit a low third-order polynomial to these local maxima, and (4) we divide out this polynomial fit from each spectra. The TELFIT⁶ code is used to compute telluric spectra (Gullikson, Dodson-Robinson & Kraus 2014). We fix the airmass to the logged airmass for each frame which we obtain from the raw FITS-header information, instead of keeping it a free parameter as done by Chiavassa & Brogi (2019).

⁵<https://atran.arc.nasa.gov/cgi-bin/atran/atran.cgi>

⁶<https://github.com/kgullikson88/Telluric-Fitter>

The Markov Chain Monte Carlo (MCMC) PYMULTINEST⁷ (Buchner et al. 2014) package is used to constrain PWV and R.

The measured relative throughput, instrumental resolving power, and PWV are shown in Fig. 4. A key result is that the measured spectral resolving power is $R = 14\,000\text{--}16\,000$, notably lower than the nominal specification for ARIES of $R = 30\,000$. We hypothesize this large offset may be due to suboptimal focus of the spectrograph resulting in consistently lower observed spectral resolution. The other fluctuations of spectral resolving power in the course of each night could be due to a combination of target misalignment on the slit entrance, instrumental instability, and occasional opening of the AO loop. This analysis also highlights the importance of measuring the spectral resolving power from the observations when using HRCCS, as our cross-correlation templates will need to be convolved to the appropriate spectral resolution later on.

3 ATMOSPHERIC MODELS

As we use the HRCCS technique to characterize the atmosphere of WASP-33 b, we require forward models for the cross-correlation templates, i.e. synthetic spectra of WASP-33 b, which we discuss in this section. We use several atmospheric modelling codes in this work to address the structure, chemistry, and dynamics of the atmosphere of WASP-33 b. To identify absorption and emission features initially, we used PHOENIX atmospheric models. PHOENIX is a general purpose atmosphere code, well-tested on objects from cool planets to hot stars (e.g. Hauschildt, Baron & Allard 1997; Hauschildt, Allard & Baron 1999; Barman, Hauschildt & Allard 2001; Allard, Homeier & Freytag 2011; Barman et al. 2011; Lothringer et al. 2018; Lothringer & Barman 2019). We used both self-consistent atmosphere models in Section 3.1 in thermodynamic equilibrium, and a grid of models where the temperature structure and/or the molecular abundances have been manually varied in Section 3.2. We then further use the Monte Carlo Radiative Transfer code gCMCRT (Lee et al. 2019, 2021) to model 3D effects and orbital phase-dependent variations of the observed spectra in Section 3.3.

3.1 Self-consistent PHOENIX models

The self-consistent models used here are similar to the extremely irradiated hot Jupiter models presented in Lothringer et al. (2018). The self-consistent atmosphere models solve the atmosphere structure and composition iteratively, calculating the spectrum at each iteration and comparing it to radiative equilibrium. For our models the radiative–convective boundary is usually below the highest pressure we model and no convective adjustments were made in order to increase speed of convergence. Temperatures are then modified to approach radiative equilibrium after which the chemistry is calculated based on chemical equilibrium and the spectrum is recalculated. This process is repeated until the temperature corrections are small, indicating that the atmosphere is in radiative and chemical equilibrium. Each spectrum is calculated from 10 to 10^7 nm using planetary parameters by Haynes et al. (2015) (as summarized in table 1 of Lothringer et al. 2018) on a grid with varying atmospheric metallicity and a heat re-radiation efficiency f (where f is the parameter as defined in Madhusudhan & Seager 2009). For the metallicity, we explore a grid of models at $0.1\times$, $1\times$, and $10\times$ Solar metallicity. For the heat re-radiation factors f we explore the values: $f = 1/4$ for

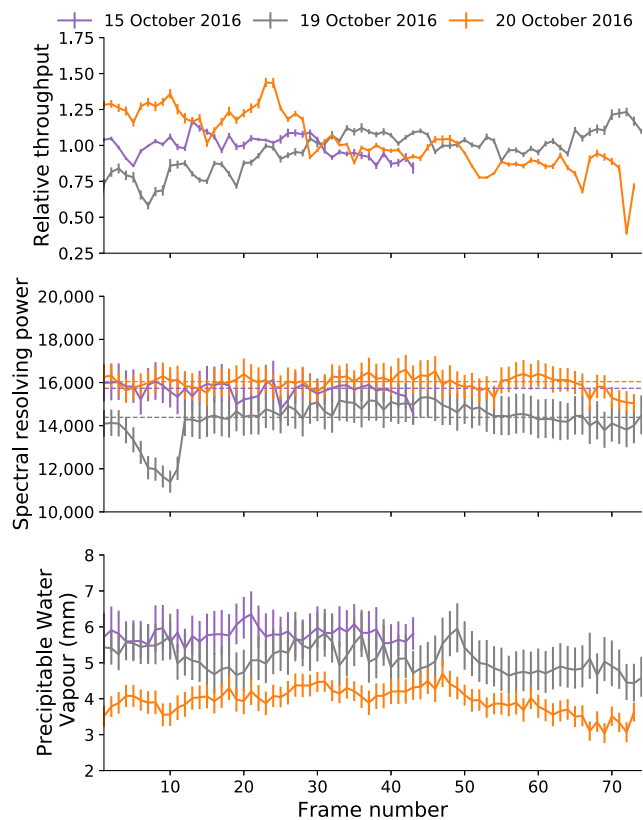


Figure 4. Relative throughput, spectral resolving power and PWV for each observing night, as measured from order 23 which is centred around the CO $2.3\ \mu\text{m}$ spectral feature. The dashed lines indicate the median resolving power used to convolve the spectral templates for each night. The overall consistently reduced resolution of ARIES is hypothesized as sub-optimal focusing of the spectrograph.

day-side and night-side heat redistribution, $f = 1/2$ for day-side-only heat redistribution, and $f = 2/3$ for instantaneous heat re-radiation.

These self-consistent models include (see Lothringer et al. 2018): bound-free opacity from H, H^- , He, C, N, O, Na, Mg, Al, Si, S, Ca, and Fe, free-free opacity from H, Mg, and Si, and scattering from e^- , H, He, H_2 . They also include opacities for TiO and VO, but as explained in Lothringer et al. (2018), TiO and VO are not the sole cause of temperature inversions in extremely irradiated hot Jupiters. Temperature increase at these pressures is likely due to the absorption of short-wavelength radiation by absorbers like atomic Fe and the lack of coolants such as H_2O . Such hot temperatures result in the thermal dissociation of most molecules in the atmosphere including H_2 and H_2O (see Fig. A2). The abundance of CO, which exhibits the strongest molecular bond, is also reduced, but to a lesser extent compared to other potential molecular opacity sources. The resulting spectra, shown in Fig. 5, predicts CO emission lines with signatures of H_2O and other molecules muted due to thermal dissociation combined with the isothermal lower atmosphere.

3.2 PHOENIX models with modified structure and/or abundances

In addition to the self-consistent modelling, we ran a grid of models with varying temperature structures and atmospheric metallicity to further explore the parameter space. Again, we explore metallicities of $0.1\times$, $1\times$, $10\times$, and $100\times$ Solar. For the P–T profiles, we use the parametrization from Madhusudhan & Seager (2009), but modified

⁷<https://johannesbuchner.github.io/PyMultiNest/index.html#citing-pymultinest>

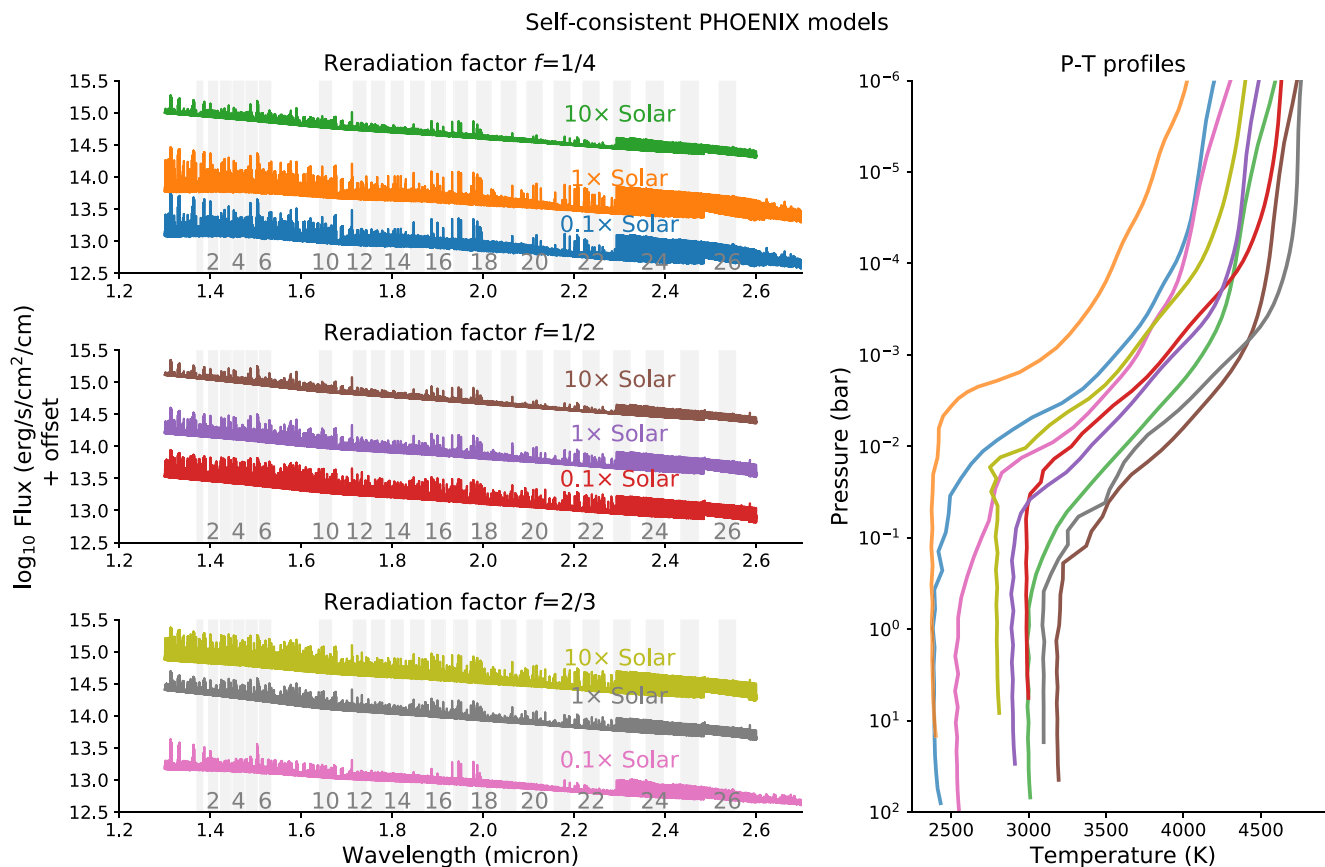


Figure 5. WASP-33 b 1D PHOENIX self-consistent synthetic emission spectra. Left-hand panels: PHOENIX self-consistent spectra as a function of the metallicity and re-radiation factor f : $f = 1/4$ for full heat-redistribution, $f = 1/2$ for day-side-only heat-redistribution, and $f = 2/3$ for instantaneous re-radiation. Different metallicity models have been offset by 0.75 along the y-axis with respect to the $0.1 \times$ Solar model. The ARIES spectral order wavelength ranges included are indicated by the grey bars. Right-hand panel: corresponding PHOENIX self-consistent P–T profiles.

to have an isothermal upper atmosphere. These simple structures used five parameters to describe the temperature throughout the atmosphere: P_1 , P_2 , T_1 , T_2 , and α_2 . P_1 describes the pressure at the tropopause, above which the temperature was set to T_1 . The temperature then varies until it reaches temperature T_2 at pressure P_2 , as set by the gradient of the inversion α_2 . We explored a wide range of P–T profiles with upper atmosphere temperatures from 2000 to about 10 100 K, shown in Fig. A3. In total we ran 576 forward models.

We also generated a small subset of these models to explore how removing a single opacity source affects the significance of the planet detection, i.e. we re-run the best matching model without H_2O , without CO, and without OH. We do this in order to identify potential absorbers or emitters in the atmosphere, while keeping the P–T profile fixed. Line lists for the investigated opacity sources were taken from: CO (Goorvitch 1994), OH (Barber et al. 2006), and H_2O (Rothman et al. 2009).

3.3 Global circulation model (GCM) and post-processing modelling

For the 3D GCM, we use output from a SPARC/MITGCM (Showman et al. 2009) of WASP-33 b with similar set-up to the simulations performed in the Parmentier et al. (2018) study for UHJs. The GCM output is then post-processed using the GCMCRT 3D RT code (Lee et al. 2021) to provide high-resolution emission spectra that account for variation in the spectra throughout the orbital phases of the

observation. We use line-list data from various sources, namely: OH (Hargreaves et al. 2019), H_2O (Polyansky et al. 2018), CH_4 (Hargreaves et al. 2020), CO (Li et al. 2015), CO_2 (Yurchenko et al. 2020), NH_3 (Coles, Yurchenko & Tennyson 2019), HCN (Barber et al. 2014). Cross-sections were calculated using the HELIOS-K opacity code (Grimm et al. 2021) and interpolated between 1.27 and 2.66 μm at a resolution of $R = 100\,000$.

The resulting synthetic high-resolution emission spectra include the Doppler shifting of spectral lines due to winds and planetary rotation towards the line of sight for each phase. We include all molecular species listed above as well as CIA H_2 -H-He pair continuum opacities from Karman et al. (2019) and H^- from John (1988).

The resulting phase-dependent GCM spectra and P–T profiles of the GCMCRT are shown in the bottom left-hand and right-hand panels of Fig. 6. It is computationally expensive and resource intensive to run the spectral calculation, as it uses a Monte–Carlo radiative transfer method on many GPUs. Therefore, we computed synthetic spectra at nine phases uniformly sampled over the phase coverage during our three observing nights. For those spectral templates required at intermediate phases, we linearly interpolate these spectral templates in phase. The GCM predicts an eastward hot spot at $+8^\circ$ east from the substellar point, which results in an asymmetry between P–T profiles with an equal longitudinal distance west or east as seen from the sub-stellar point. As the planet rotates in the same direction as it orbits its host star, we see more of the hot spot region pre-eclipse compared to post-eclipse. This results in

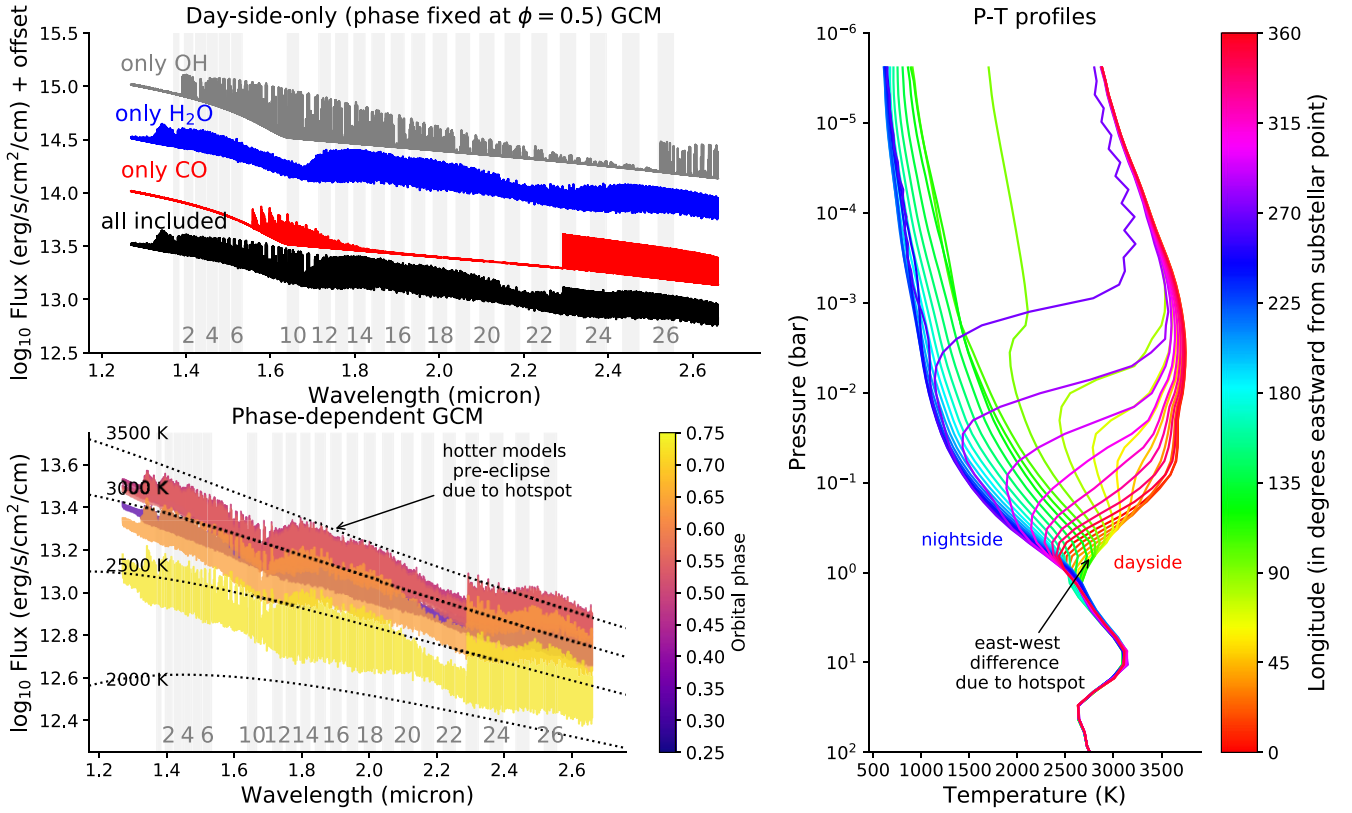


Figure 6. WASP-33 b GCM synthetic emission spectra. Top left-hand panel: Day-side-only (i.e. fixed at phase $\phi = 0.5$) GCM spectra with all opacity sources included or only specific molecules included. The different models have been offset by 0.5 along the y-axis with respect to the spectrum at the bottom. The ARIES spectral order window wavelength ranges included are indicated by the grey bars. Bottom left-hand panel: phase-dependent GCM spectra at different orbital phases. Planck black body curves are plotted at a range of temperatures. The GCM predicts hotter models pre-eclipse compared to post-eclipse due to an hot spot approximately $+8^\circ$ eastward seen from the substellar point. Right-hand panel: GCM P-T profiles as a function of longitude, where 0° is defined as the substellar point and increases towards the eastward direction as seen from the substellar point.

phase-dependent variations in both the continuum and relative line strengths (see bottom left-hand panel of Fig. 6). As noted by de Kok et al. (2014), while the overall flux on cooler night sides may be lower, the relative line-to-continuum strength, which HRCCS is sensitive to, is strong. This suggests then that the phase-dependent change in the relative line strengths from the GCM will also be detectable as WASP-33 b proceeds in its orbit.

To enable comparison and cross-check between the 3D GCM and the 1D PHOENIX model atmospheres, we also created a set of day-side-only GCM spectra, i.e. fixed at orbital phase $\phi = 0.5$ (see top left-hand panel of Fig. 6). These templates were created by simulating the emission assuming continuum opacity without the effects of Doppler shifting. We also again explore the impact of any single absorber or emitter on the planet signal by creating day-side-only templates with only a single opacity source included, i.e. only CO, only H₂O and only OH. CO is most prolific in spectral lines in the reddest parts of the ARIES spectra, while OH dominates its shorter wavelengths, and water persists throughout.

4 HIGH-RESOLUTION CROSS CORRELATION SPECTROSCOPY

The HRCCS technique has been developed in recent years to enable not simply detection of molecular species, i.e. the S/N-method, but into a framework that enables statistically rigorous assessment of atmospheric models that may match the data e.g. Cross-Correlation-

to-log-Likelihood (CC-to-log(L)) mapping (e.g. Brogi & Line 2019; Gibson et al. 2020). Both methods have different uses, and we use both in this work to characterize the composition, structure, and dynamics of the WASP-33 b atmosphere, as described below.

4.1 S/N method

The S/N-method cross-correlates each observed spectrum with a spectral template and the S/N is calculated from the maximum and standard deviation of the corresponding cross-correlation function. First the spectral template is scaled to the expected star-planet contrast ratio using

$$F_{\text{model, scaled}} = \frac{F_{\text{model}} R_p^2}{B_\star R_\star^2}, \quad (1)$$

with planet radius R_p , stellar radius R_{star} , and assuming a stellar blackbody flux B_\star at the effective stellar temperature (using values in Table A1). A stellar blackbody is a reasonable assumption for the host star WASP-33 given its hot effective temperature which results in few spectral stellar absorption lines in the wavelength region of interest. Previous works indicate that it is important to match the line shape correctly for HRCCS (e.g. Spring et al. 2022), hence before cross-correlating, we convolve each synthetic spectrum of WASP-33 b to the median observed resolving power for each night (see Fig. 4), using a Gaussian instrumental profile. Since we applied a sixth-order Butterworth high-pass filter at a cutoff-frequency of 0.02 pixel^{-1} to

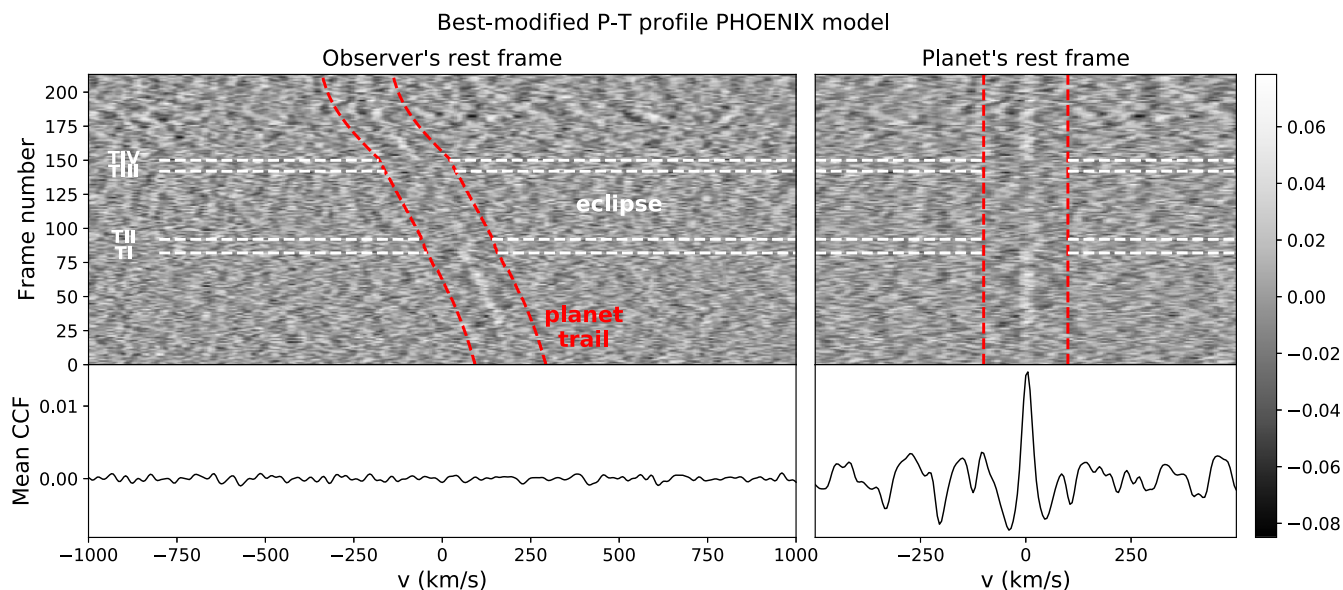


Figure 7. Cross-correlation functions for the best-modified P–T profile detected at a S/N of 7.9σ . Top left-hand panel: The CCF for each frame including all three observing nights. Frames are sorted by orbital phase but are not equally spaced, and the blue dotted line marks the post-eclipse frame with symmetric phase to the earliest pre-eclipse frame (see Fig. 1). The start/end of ingress/egress are indicated by the white dashed lines labelled TI–TIV for the four contact points. The two red lines are centred around the planet radial velocity trail. Top right-hand panel: Same as the top left-hand panel, but shifted to the planet rest frame. Bottom left-hand panel: the mean CCF along the y-axis. Frames where the planet is not fully visible (in between TI–TIV) were excluded when computing the CCF to prevent dilution of the planet signal. Bottom right-hand panel: same as the bottom right-hand panel, but shifted to the planet’s rest frame.

our data, we also apply it to the scaled model template, where we adjust the cutoff-frequency to account for the difference in sampling rate between the model and data wavelength grid. We calculate the correlation coefficient as defined by Brogi & Line (2019) for each observed residual for each spectral order for all observing nights with the scaled template. We explore a radial velocity range of $[-1000, 1000]$ km s^{-1} in steps of $\Delta v = 5 \text{ km s}^{-1}$, close to the theoretical instrumental resolution of ARIES. This results in a CCF time series for each spectral order, we will refer to this as the CCF-matrix (see Fig. 7). To robustly detect if the emission source is indeed the planet, it is common practice to shift each CCF-matrix using a grid of trial velocities into the proposed planet rest-frame velocity (see right-hand panel of Fig. 7):

$$v_{\text{trial}} = v_{\text{bary}} + v_{\text{sys}} + K_p \sin \phi_p, \quad (2)$$

with v_{bary} the barycentric velocity, v_{sys} the trial system velocity, trial K_p the semimajor amplitude of the planet’s orbital velocity, and ϕ_p the planet’s orbital phase, where we have assumed a circular Keplerian orbit. The PYTHON package BARYCORRPY⁸ (Wright & Eastman 2014; Kanodia & Wright 2018) is used to calculate v_{bary} at the observed times. We explore a 31×31 grid in (v_{sys}, K_p) -space around the expected values of $(v_{\text{sys}}, K_p) = (-0.3, 230.9) \text{ km s}^{-1}$ as previously found in the detection of OH by Nugroho et al. (2020b) within the range of $\pm 100 \text{ km s}^{-1}$ for v_{sys} and $\pm 150 \text{ km s}^{-1}$ for K_p on a 31×31 grid. For each trial velocity on the grid, we shift the CCF-matrix to v_{trial} and calculate the combined CCF by taking the mean along the time axis. The S/N is determined by calculating the maximum of the combined CCF and dividing it by the standard deviation of the CCF. Those exposures taken in between the first and last contact of the eclipse of WASP-33 b (see Fig. 1) are excluded from the

S/N calculation to avoid dilution of the combined CCF signal. Not all spectral orders are equally sensitive when cross-correlating with the spectral template. Lower sensitivity is expected for the bluer ARIES orders due to poorer data quality due to lower throughput and stronger spectral fringes and weaker spectral features in the models. To account for both, we use an order-specific weighting scheme in the S/N-method to combine the CCF-matrices, based on the data quality and model. To calculate the weights we follow Spring et al. (2022), where we first inject the model (before PCA cleaning) into the observations using the expected model strength and orbital parameters (using the values reported by Nugroho et al. 2020b). We create a CCF-matrix for each injected order, and then subtract its corresponding observed (no injection) CCF-matrix. For each order we calculate the S/N of this CCF difference. For each order n , weights w_n are assigned proportional to this S/N and normalized such that $\sum_{n=1}^N w_n = 1$, where N is the total number of spectral orders included. An example of the weighting scheme is shown in Fig. 8. As expected, using the self-consistent PHOENIX at $\times 1$ Solar metallicity with a re-radiation factor of $f = 1/2$ as an example, the reddest orders around the ~ 2.3 micron densely packed CO spectral feature region contribute the highest weights due to the numerous strong spectral lines and higher S/N of the observed spectra.

4.2 CC-to-log(L) mapping and model scaling parameter

While the S/N-method is widely used in literature and thus a convenient way to compare the significance of our detection with previous works, it is inherently normalized by definition and therefore insensitive to scaling of the model or observed spectra. This is why there was a need to implement a weighting scheme to combine multiple spectral orders in the S/N-method. The CC-to-log(L) Bayesian Framework described by Brogi & Line (2019) however does not suffer from this. CC-to-log(L) mapping maximizes

⁸<https://github.com/shbhuk/barycorrpy>

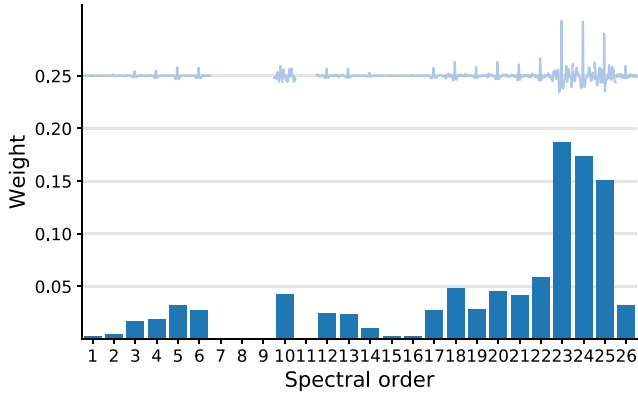


Figure 8. Weights per spectral order after cross-correlation with the best self-consistent PHOENIX spectral template used in our S/N-method adopted. Weights are calculated from the S/N of the difference between the injected and observed CCF, indicated by the miniature CCF above each order. Orders 7–9 and 11 were excluded from the analysis in the wavelength calibration step. Orders 23–25 around the CO 2.3 μm spectral feature are naturally assigned the highest weights.

Table 2. Priors used in CC-to-log(L) framework. Velocity offsets are given relative to the previous WASP-33 b velocity values of $(v_{\text{sys}}, K_p) = (-0.3, 230.9)$ km s^{-1} as reported by Nugroho et al. (2021).

Parameter	Symbol	Prior
System velocity offset	Δv_{sys}	Uniform($-50, 50$) km s^{-1}
Orbital velocity offset	ΔK_p	Uniform($-50, 50$) km s^{-1}
Scaling parameter	a	LogUniform($-2, 2$)

the log-likelihood as:

$$\log L = -\frac{N}{2} \log (s_g^2 - 2R + s_f^2) \quad (3)$$

with s_g and s_f the variance of model and data, respectively, R the cross-covariance, and N the number of spectral channels. When combining multiple nights and spectral orders, we can compute the total sum of their log-likelihood values. There is thus no need to weight each spectral order as this is accounted for in the CC-to-log(L) mapping (Brogi & Line 2019). Moreover, CC-to-log(L) mapping gives us a framework to extract uncertainties on the measured and derived quantities from the modelling.

Our implementation uses PYMULTINEST⁹ (Buchner et al. 2014), a generic PYTHON package connected to MULTINEST, a Bayesian inference tool¹⁰ (Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009; Feroz et al. 2019), for all our parameter estimations. Importantly, following Brogi & Line (2019), we also introduce one additional free parameter, the scaling factor a , which allows us to compensate for any potential unknown scaling of the model template, but ideally should be retrieved at $a = 1$. We multiply the template by a after applying the high-pass filter. The multidimensional parameter space is explored using the priors in Table 2.

To statistically compare models drawn from the same underlying parameter distribution amongst each other we apply Wilks' Theorem (Wilks 1938). In practical terms, this theorem states that the difference in $\log(L)$ -values, also known as the Bayes factor, between two models with n free parameters follows a χ^2 -distribution with

n degrees of freedom (for a concise statistical description see Pino et al. 2020). The confidence interval p -value can now be calculated from the corresponding χ^2 -distribution and converted to a familiar σ -value. This way, by its construction, we can only compare models relative to the best trial model.

5 RESULTS

We now present the results using both the S/N-method and the CC-to-log(L) mapping for our four different modelling suites described in Section 3. We detect the atmosphere of WASP-33 b and measure its properties to varying significance and confidence for the best matching model in each modelling suite. The key results are summarized in Table 3. The best-matching model from each suite results in the detection of the Doppler-shifting spectrum of the planet, as demonstrated by the radial velocity trail in the CCF-matrices shown in Figs 7 and A4. Notably, the trail disappears at the calculated start time of secondary eclipse ingress and reappears at the end of its egress, confirming that the detected signal is associated with the planet's spectrum as it is obscured by the star during secondary eclipse.

5.1 Thermal structure and composition

The model that results in the highest S/N detection is a modified PHOENIX model. The highest S/N model includes all opacity sources and has the following P–T profile parametrization: $T_1 = 3651$ K, $T_3 = 2000$ K, $P_1 = 10^{-3}$ bar, $P_3 = 10^{-6}$ bar, $\alpha_2 = 0.17$, and a metallicity of $10 \times$ Solar. As shown in the top row of Fig. 9 this model gives $S/N = 7.9\sigma$ when all nights are combined, and is higher during pre-eclipse orbital phases than post-eclipse.

In the calculation of the S/N-ratio (i.e. CCF peak/CCF standard deviation), we assume the distribution of the cross-correlation values is Gaussian and not impacted significantly by any residual correlated noise. To assess this, we plot the distribution of the cross-correlation values in-trail (three most central columns) and out-of-trail (outside of the ten most central columns). The result, based on the best matching model described above in this section, is shown in Fig. A5 for the highest S/N-ratio modified P–T profile PHOENIX spectral template. To quantify whether the samples are normally distributed, we performed a Shapiro Wilks Test on the in-trail and out-of-trail samples (that is after aligning to the planet's rest frame). Both samples pass the test with p -values > 0.05 ($p_{\text{in}} = 0.75$ and $p_{\text{out}} = 0.09$). Therefore, the distribution of the cross-correlation values follows a normal distribution, and we conclude the effect of residual correlated noise on the S/N-measurement is negligible.

The correspondingly similar results for the PHOENIX self-consistent models are shown in Fig. A6. Although the S/N of the detection varies with phase, this alone is not sufficient evidence to interpret it as due to a planetary feature e.g. an offset hot spot, day-to-night variations in brightness, abundances, or thermal structure. However, we discuss in Section 6.3 the additional evidence supplied by the CC-to-log(L) mapping scaling parameter that could support such interpretations.

The top left-hand panel of Fig. 10 shows the results from the log(L)-to-CC mapping for the best-matching modified PHOENIX model, giving the posterior distributions for the systemic velocity, orbital velocity, and the scaling parameter. The planet is detected at $v_{\text{sys}} = 0.15_{-0.65}^{+0.64}$ km s^{-1} and $K_p = 229.53_{-1.02}^{+1.11}$ km s^{-1} with this highest S/N spectral template, in agreement with previous literature e.g. $(v_{\text{sys}}, K_p) = (-0.3, 230.9)$ km s^{-1} (Nugroho et al. 2021). The scale factor posterior distribution, which would have $\log(a) = 0$ for

⁹<https://johannesbuchner.github.io/PyMultiNest/>

¹⁰<https://github.com/JohannesBuchner/MultiNest>

Table 3. Overview of the highest S/N-ratio of the sets of models explored in this work and their corresponding scaling parameter a from the CC-to-log(L) mapping. The final column shows $\Delta \log(a)$ i.e. the difference between the mean of the pre-eclipse (Nights 1 & 3) scaling parameter and that for the post-eclipse data subset for which we have symmetric phase coverage. The $\log(a)$ scaling parameter has a value of zero when the model is a good description of the data. The phase-dependent GCM gives the smallest discrepancy in $\log(a)$, indicating that it is accounting for the phase-dependence of the scaling parameter.

Model	Highest S/N	$\Delta \log(a) = \log(a_{\text{post-eclipse}}) - \log(a_{\text{pre-eclipse}})$ (symmetric phase coverage)
Modified P–T profile PHOENIX models (1D)	7.9	$0.21^{+0.03}_{-0.03}$
Self-consistent PHOENIX models (1D)	7.6	$0.24^{+0.03}_{-0.03}$
Day-side-only (fixed orbital phase $\phi = 0.5$) GCM (3D)	7.1	$0.27^{+0.03}_{-0.04}$
Phase-dependent GCM (3D)	7.1	$0.13^{+0.03}_{-0.03}$

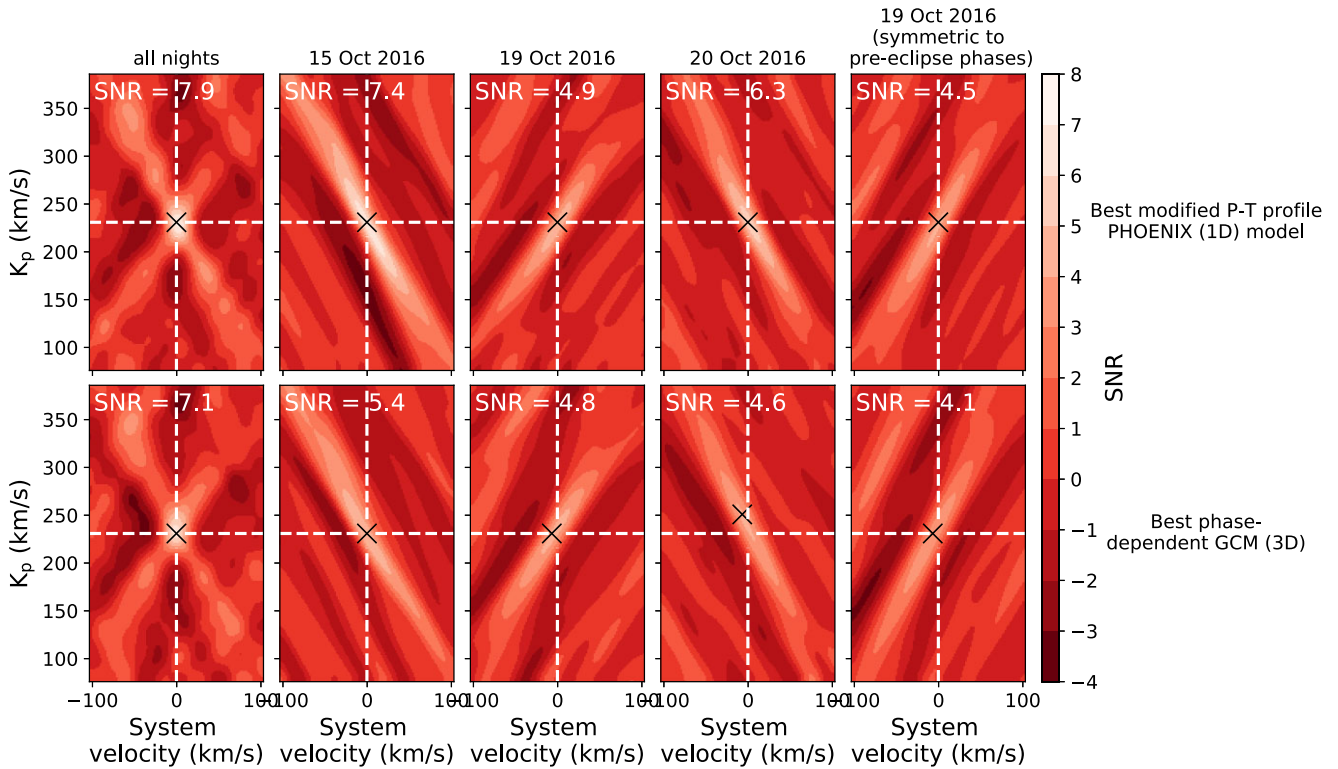


Figure 9. S/N as a function of the system velocity v_{sys} and orbital velocity K_p . The grey plus symbol marks the maximal S/N location. The dashed white grid indicates the expected location at $(v_{\text{sys}}, K_p) = (-0.3, 230.9) \text{ km s}^{-1}$ based on the OH detection by Nugroho et al. (2021). Our detection is in agreement with their (v_{sys}, K_p) -values for all nights combined. Top row: the highest S/N-ratio modified P–T profile PHOENIX model detected at 7.9σ . Bottom row: the phase-dependent GCM spectral template detected at 7.1σ .

a perfectly matched model, is systematically offset to smaller values, indicating that the model does not fully encapsulate all of the physics and chemistry to describe the atmosphere. The same thing is seen in the top right-hand panel of Fig. 10 for the best-matching self-consistent PHOENIX model. However, all nights are in agreement with their (v_{sys}, K_p) -values within 2σ . We discuss the variation of $\log(a)$ in later sections.

We compare the best-matching modified P–T profile PHOENIX model with the rest of the models in the suite, as well as that of the best-matching self-consistent PHOENIX model, to give confidence intervals with respect to the alternative P–T profiles, fixed at $10\times$ Solar metallicity. This was calculated using Wilks’ Theorem (with seven free parameters), and is shown in Fig. 11. The best-matching modified P–T profile is slightly cooler than the best-matching self-consistent PHOENIX model, and inverted P–T profiles with an upper and lower atmosphere temperature difference close to the

best-matching model are favoured. Both can be understood by the fact that we are sensitive to the relative line strength with respect to the continuum, which is set by the upper and lower atmosphere temperature contrast (or equivalently the strength of the inversion layer). On the contrary, we are less sensitive to the absolute temperature due to loss of the continuum information in the HRCCS processing, resulting in a range absolute lower and upper atmosphere temperatures inside the calculated 1σ -confidence interval.

To determine if any one particular species was contributing the majority of the detected signal, we fix the P–T profile to the best-matching modified P–T profile PHOENIX model (all opacities included), and then run the S/N analysis for each individual night without CO, H₂O, and OH, respectively. The results are shown in Fig. 12. Exclusion of CO significantly impacts the S/N on every night, and drops from 7.9σ to 4.3σ for all nights combined. On the

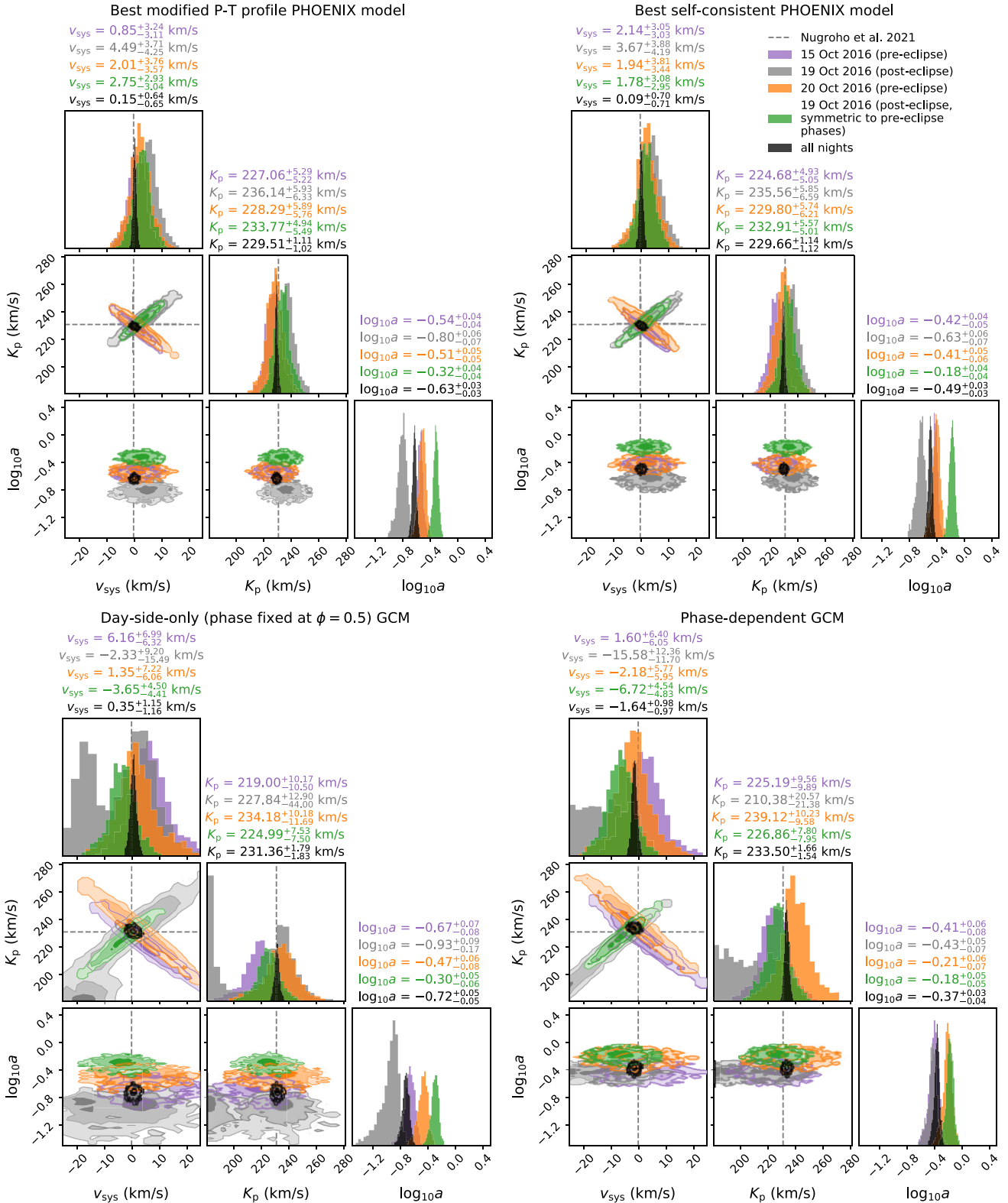


Figure 10. Marginalized distributions from the CC-to-log(L) framework based on Brogi & Line (2019). Results are shown for the best-matching modified P–T profile PHOENIX model (top left-hand panel), the best-matching self-consistent PHOENIX model (top right-hand panel), the day-side only GCM (phase fixed at $\phi = 0.5$; bottom left-hand panel) and the phase-dependent GCM (bottom right-hand panel). The contours indicate the 1σ , 2σ , and 3σ confidence intervals. All frames with phases during the eclipse have been excluded from this analysis to prevent dilution of the planetary signal. Results are shown for individual observing nights, all nights combined, and for the subset of post-eclipse phases that are symmetrically matched with the pre-eclipse phases (see Fig. 1). The dashed black grid indicates the expected location at $(v_{\text{sys}}, K_p) = (-0.3, 230.9)$ km s $^{-1}$ based on the OH detection by Nugroho et al. (2021).

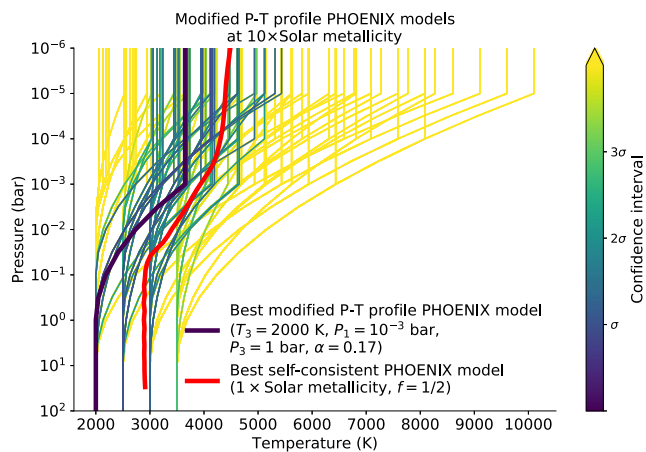


Figure 11. Confidence intervals for the modified P–T profile PHOENIX models at $10 \times$ Solar metallicity. These are for all observing nights combined and were calculated using Wilks’ Theorem. The P–T profile of the best modified structure model, the one with the highest log-likelihood, is plotted in dark green and has the following parametrization: $T_3 = 2000$ K, $P_1 = 10^{-3}$ bar, $P_3 = 1$ bar, $\alpha_2 = 0.17$. All confidence intervals here are with respect to this best model. The confidence intervals are converted to their corresponding σ -value (as described in the last paragraph of Section 4.2.) The best self-consistent PHOENIX model’s P–T profile, at $1 \times$ Solar metallicity and with a heat reradiation factor $f = 1/2$, is plotted in red for comparison.

other hand, exclusion of OH and H₂O only marginally reduces the measured S/N, indicating that we are not sensitive to these species at their abundance with this data. However, on the first night of pre-eclipse data, we still detect the planet signal to relatively high S/N, despite the exclusion of CO (the corresponding S/N matrices are shown in Fig. A7). This suggests that other species in the model are summing to give a detection, but we do not identify these individual species. The data on the third night, which covers the same orbital phase, is of lower quality and may explain why a similar feature is not seen during this night. We also run the same analysis using Wilks’ Theorem to compare the models with a single opacity source excluded to the model with all opacity sources included for the best modified P–T profile PHOENIX model (see left-hand panel of Fig. A8. We find for all nights combined models without CO are significantly worse ($\sigma = 11.1$). Interestingly, there also seems evidence models without H₂O ($\sigma = 5.4$) and marginal evidence OH ($\sigma = 3.4$) are worse which we will discuss further in Section 6.

While the modified PHOENIX models allow us to explore the structure and chemical species of the atmosphere of WASP-33 b, the self-consistent PHOENIX models allow us to assess its heat-redistribution efficient and metallicity. The highest log-likelihood in the PHOENIX self-consistent model suite corresponds to a day-side-only heat-redistribution efficiency $f = 0.5$ and a metallicity of $1 \times$ Solar. Confidence intervals comparing the self-consistent PHOENIX models computed using Wilks’ Theorem (with four free parameters) are shown in Fig. 13, where we use a visualization similar to that recently shown by Giacobbe et al. (2021). Models with high reradiation factors $f \geq 1/2$ and high metallicity $\geq 1 \times$ Solar are favoured and within 1σ of our best model. We note that the higher S/N ratio for the modified PHOENIX models compared to the self-consistent PHOENIX models is not necessarily surprising given the additional free parameters in the model and the larger number of models explored in the suite.

5.2 Atmospheric dynamics and longitudinal variation

The post-eclipse phase coverage on Night 2 is longer than the pre-eclipse coverage on Nights 1 and 3. To enable a direct comparison, we split the post-eclipse Night 2 spectra into a subset of data that contains only the symmetric phases corresponding to the pre-eclipse coverage. A comparison of results from Night 2 and this subset of data with the different model suites are shown in Fig. 10 by the grey and green contours, respectively, and in the right-hand column of the S/N maps in Fig. 9. The S/N for the subset data is reduced, but only by $\Delta S/N = 0.7$ at most, indicating that the data containing more night side hemisphere contributes less to the overall S/N of the detection. The green contours for the subset of phases are also tighter and better constrained in all cases in the $v_{\text{sys}} - K_p$ plots, again indicating a better match with the model when the night hemisphere contributions in the data set are reduced. As shown in Table 3, the scaling parameter from the CC-to-Log(L) mapping differs for the symmetric pre- and post-eclipse phases, consistent with a phase-dependence in the observed spectra of the planet due to asymmetry in the disc of the planet caused by e.g. an offset hot spot. To qualify this further, we use the GCM spectral templates described in Section 3.3.

To first check for general consistency between the GCM and the results from the PHOENIX 1D models, we perform our two analysis methods using the day-side-only (i.e. no phase dependence) 3D GCM with all opacities included. This is detected at $S/N = 7.1\sigma$, which is slightly lower than the PHOENIX models, and found at the expected (v_{sys}, K_p) -values for all nights combined (see bottom row of Fig. A6 and the corresponding cross-correlation trail is shown in the middle row of Fig. A4). In terms of their likelihoods, we find a Bayes factor of $+59.2$ between the best PHOENIX 1D model and GCM suggesting the best PHOENIX 1D model is preferred over the GCM. However, we cannot compare the GCM and PHOENIX 1D models quantitatively in terms of their likelihood as they are drawn from different underlying parameter spaces. The preference of the best 1D PHOENIX model over the GCM is not necessarily surprising. This is because first, we explore only a single GCM day-side-only spectral template versus many PHOENIX spectral templates, and secondly, the GCM lacks some of the important optical absorbers in its energy balance e.g. atomic iron. We can also assess in the day-side-only GCM model if any single species contributes the majority of the detection. The S/N results for single species models are shown in the right-hand panel of Fig. 12 and the corresponding S/N plots are shown in Fig. A9. From all opacity sources, only the CO GCM model results in a significant S/N of 6.6σ . The GCM H₂O spectral template has $S/N = 3.3\sigma$, which is insufficient for a robust detection in HRCCS as previous work demonstrated spurious signals at an S/N ratio $\leq 4\sigma$ often persist (Cabot et al. 2020; Spring et al. 2022), even though it appears at the expected planet velocity. OH is not detected. We also calculate the confidence intervals of each opacity source template to the case where all opacity sources are included. The results are shown in the right-hand panel of Fig. A8. We find a model with only CO is significantly better ($\sigma = 11.7$) compared to the day-side-only GCM with all opacity sources included, again in agreement with a robust detection of CO. Unsurprisingly, models with CO included are significantly better than models with just H₂O or OH. The CC-to-log(L) analysis shown in Fig. A11 further highlights the lack of robust detection of any species by CO with this model, in general agreement with the PHOENIX spectral templates. Although the low scale parameter for water in Fig. A11 could indicate water depletion from WASP-33 b’s atmosphere, we find it more compelling that this results from telluric residuals in the final spectra, which is supported by the low S/N ratio detection of the GCM H₂O spectral template

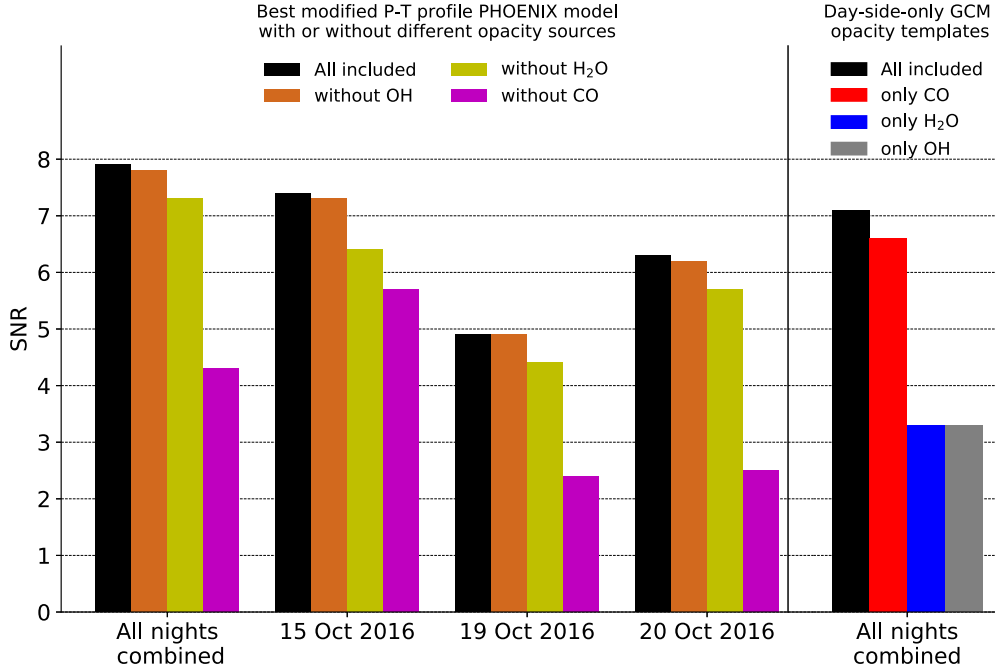


Figure 12. S/N results for the models with different opacity sources included or excluded. Left-hand panel: results from the best modified P–T profile PHOENIX models for the cases of all opacity sources included and with a single opacity source excluded, respectively models with OH, H₂O, and CO excluded. All these models fix the P–T profile to the best modified structure PHOENIX model where all opacity sources were included. Only the models without CO result in a significant drop of the S/N when compared to the all included case. Right-hand panel: Results from the dayside-only GCM for the cases of all opacity sources included and respectively with only CO, only H₂O and only OH included. From all opacity sources, only CO remains robustly detected.

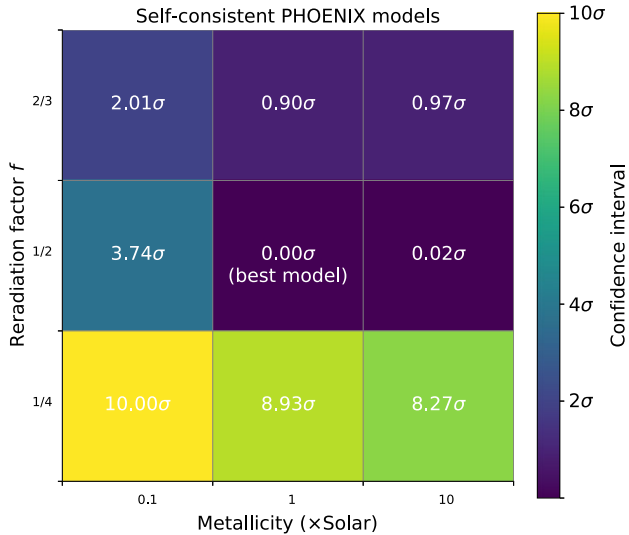


Figure 13. Confidence intervals for the PHOENIX self-consistent model grid for all observing nights combined using Wilks’ Theorem. The best model, the one with the highest log-likelihood, is indicated. By definition of Wilks’ Theorem, all confidence intervals here are with respect to this best model. The confidence intervals are converted to their corresponding σ -value (as described in the last paragraph of Section 4.2) which are annotated in the centre of each tile. Models with $f \geq 1/2$ and metallicity $\geq 1 \times$ Solar are all within 1σ .

and the large offset of this signal in (v_{sys}, K_p) -space away from the expected values for WASP-33 b.

Our main goal is to understand possible phase dependence in $\log(a)$. The bottom left-hand panel of Fig. 10 shows the results from the CC-to- $\log(L)$ framework for the GCM day-side-only model. The

results are qualitatively similar to the PHOENIX spectral templates, again with offsets in $\log(a)$ for the different phase ranges, but with larger error margins. We hypothesize these larger error contours may be due to the broadening of the lines by planet rotation in the GCM spectral template and its overall lower S/N detection. The broad agreement with the trend for offsets in $\log(a)$ is confirmed by the day-side-only GCM and thus we proceed to allow the GCM to have phase dependence and determine if this resolves the $\log(a)$ scaling discrepancy.

The phase-dependent GCM provides a different cross-correlation template for each observed spectrum at each phase, where the spectrum corresponds to a P–T profile that a combination of all the profiles from different longitudes visible across the planet disc at that time. The phase-dependent GCM is similarly detected at $S/N = 7.1\sigma$ at the expected (v_{sys}, K_p) -location for all night combined (the corresponding CCF-matrix and $K_p - v_{\text{sys}}$ maps are shown in the bottom rows of Figs A4 and 9, respectively). However, the results from the CC-to- $\log(L)$ mapping shown in the bottom right-hand panel Fig. 10 are notably different for the scaling parameter. The phase-dependent GCM appears to resolve the majority of the $\log(a)$ discrepancy between the pre- and post-eclipse symmetric phases, and brings all the data sets into broad agreement within 2σ . Although the phase-dependent GCM resolves most of the $\log(a)$ discrepancy seen for the 1D PHOENIX models, indicating that the inclusion of 3D dynamical effects such as hot spots and orbital phase brightness variation are a better match to the data, it does not fully describe the atmosphere of WASP-33 b. While the 1D model results in a higher S/N detection, it is possible that differences in full chemistry, or compensating with other parameters enable a better match, but only once scaled in the log-likelihood framework. The scaling parameter for the 1D models still indicates greater correction is needed for these models than the GCM. Thus there is scope for future work to improve

on 3D models for ultra hot Jupiters to give a more comprehensive description of the atmosphere that would result in a higher S/N than the 1D models. This missing physics in the GCM may account for the overall lower S/N ratio, compared to the S/N ratio found for the best 1D PHOENIX models.

6 DISCUSSION

6.1 Thermal structure and composition

The detection of CO emission lines with the modified PHOENIX P–T profiles provide unambiguous evidence for a thermal inversion in the atmosphere of WASP-33 b. A non-inverted atmosphere would have exhibited absorption lines instead. This is in agreement with previous observations of other atomic and molecular lines in emission in the atmosphere of WASP-33 b (e.g. Deming et al. 2012; Haynes et al. 2015; Zhang et al. 2018; Nugroho et al. 2020b; Cont et al. 2021, 2022; Herman et al. 2022) and expected from atmospheric modelling of UHJ atmospheres (e.g. Lothringer et al. 2018; Gandhi & Madhusudhan 2019). Although, we do not explore non-inverted or isothermal P–T profiles, our highest log-likelihood P–T profile (see Fig. 11) shows temperature contrasts with well-constrained upper- and lower limits when including a thermal inversion layer.

From all molecular templates investigated in this work, only CO has resulted in a robust detection. This is further supported by the CC-to-log(L) results shown in Fig. A11 where only CO has similar constraints on the velocities compared to when all opacity sources are included: $v_{\text{sys}} = 0.75^{+0.81}_{-0.79}$ km s⁻¹, $K_p = 228.74^{+1.26}_{-1.29}$ km s⁻¹. Furthermore, only the exclusion of CO had a significant impact on the S/N ratio of the planet signal. It does remain somewhat puzzling however that we still detect the best modified PHOENIX model without CO at an S/N = 5.8 in the first night of pre-eclipse data. We cannot attribute these to H₂O or OH alone in terms of the S/N ratio, and we do not expect significant contributions from iron lines in the ARIES wavelength range either. However, if we compare the same models in terms of the CC-to-log(L) framework, we find exclusion of water or OH does result in a significantly worse model (see Fig. A8), perhaps suggesting a marginal detection of these species or the combined set of all other opacity sources, besides CO, during the first night.

However, we do notice that the confidence intervals calculated from Wilks’ Theorem tend to suggest much higher σ -values than their S/N ratio counterparts. Possibly, this is due to the assumption that the total number of spectral channels N in equation (3) all provide independent and uncorrelated measurements according to the definition of the likelihood by Brogi & Line (2019). This assumption does not account for correlated noise between neighbouring pixels or for the fact that the full width at half-maximum (FWHM) of a resolution element may cover several pixels, particularly when observing at lower spectral resolutions. A full investigation into these effects on the CC-to-log(L) mapping and resulting σ -values is needed for this field of research.

We do not detect the model without CO during the third night, despite covering similar orbital phases. This may be explained by the better observing conditions during the first observing night relative to the third observing night (see S/N measurements in Table 1). Although there have been many detections of CO in absorption (Snellen et al. 2010; Brogi et al. 2012, 2014, 2016b; Rodler, Lopez-Morales & Ribas 2012; de Kok et al. 2013; Rodler et al. 2013; Lockwood et al. 2014), this is the first of CO emission lines using the HRCCS technique, using only the large Doppler-shift induced by the planet’s orbital motion, that is without the aid of high contrast imaging.

Previous works have indicated both water (Haynes et al. 2015) and OH (Nugroho et al. 2021) in the atmosphere of WASP-33 b, but we do not make a robust detection of these molecules in the ARIES data. To demonstrate that the very weak GCM H₂O signal in our data is caused by non-planetary residuals e.g. tellurics, we ran the CC-to-log(L) analysis using only frames obtained when the planet is not visible i.e. during the full eclipse of WASP-33 b. Spurious signals due to telluric or stellar residuals should persist whereas signals originating from WASP-33 b should disappear or move away from the planet’s expected (v_{sys}, K_p) location. For H₂O we found that a signal is still retrieved at the same offset (v_{sys}, K_p) position albeit with larger errors, while the CO signal was not obtained at the expected planet (v_{sys}, K_p) position anymore. This supports the interpretation that any H₂O signal in the ARIES data is caused by non-planetary residuals.

The non-detection of OH in with ARIES/MMT is not necessarily surprising either when compared to the OH detection by Nugroho et al. (2021) using Subaru/IRD data. The bluer wavelength region (0.97–1.75 μm) covered by Subaru/IRD with respect to ARIES/MMT contains stronger OH emission lines and Subaru/IRD’s has higher spectral resolution ($R = 70\,000$). Thus, the lack of an OH detection with ARIES can likely be explained by the poorer data quality of the spectral orders covering the OH line dense regions compared to CO line dense regions. The synthetic GCM spectrum in the left-hand top panel of Fig. 6 shows OH line dense regions in the wavelength range of the bluer spectral orders and order 26 at the far red-end. As explained in Section 4.1, lower weights are assigned to the bluer ARIES orders which have a lower throughput and contain residual fringing. The reddest order 26 falls at the edge of the detector and is consequently of poorer data quality. Lastly, telluric absorption is strong in some OH line dense spectral orders.

Due to the high dissociation temperature for CO, the abundances of CO are approximately constant in- and outside of the hotter and cooler regions of the planet, at the pressures probed by HRCCS (Lodders & Fegley 2002). This is in line with modelling results of WASP-33 b by Tsai et al. (2021) who also found CO to be the only molecule that can be stable against dissociation. This is in contrast to other molecular species such as OH, H₂O, or TiO which dissociate on the day-side (at least up to temperatures of ~ 4000 K at a pressure of 10^{-4} bar). For example, for H₂O this will result in an abundance difference of several orders of magnitude between the day- and night-side. Since we assume Solar metallicity, and the Solar C/O value is ~ 0.54 (Asplund et al. 2009; Caffau et al. 2011), the CO abundance will not really depend on the varying H₂O abundance, but rather the limiting factor to form more CO will be the amount of carbon atoms available. This is regardless of the extra freed-up oxygen atoms that become available as water dissociates at higher day-side temperatures. For species other than CO, their abundance has a strong temperature-dependence and introduces a degeneracy between the effects of the atmospheric abundance (or the dissociation rate) and thermal structure, complicating the interpretation of the line contrast. This stability of CO against dissociation means the CO line contrast is a reliable tracer of the thermal structure of the atmosphere. We therefore advocate for CO as a better probe of the thermal structure and inversions of UHJs compared to H₂O in HRCCS.

6.2 Robustness of CO as a temperature tracer in the presence of stellar activity

The CCFs inside the planet radial velocity trail in Fig. 7 are distinct and easily seen even by eye. We do not see contaminating signatures of the δ Scuti pulsations. On the contrary, the contamination by stellar

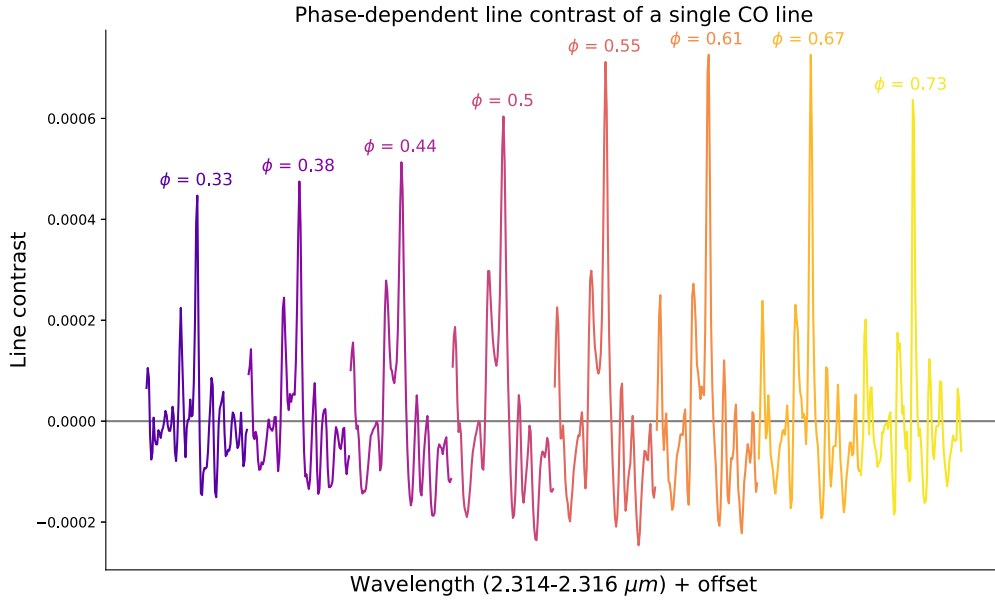


Figure 14. Line contrast (flux–mean(flux)) of a single CO line (2.314–2.316 μm) in the phase-dependent GCM as function of orbital phase (ϕ) coverage. The phase-dependent variations show shallower lines pre-eclipse compared to post-eclipse for symmetric phases around the secondary eclipse. This is due to a shallower average temperature gradient in the pre-eclipse phases compared to post-eclipse.

pulsations can be seen clearly in the cross-correlation analyses neutral iron emission lines (Nugroho et al. 2020a; Herman et al. 2022). This is not exclusive to WASP-33 b; any pulsating star with opacity sources present in both the star and planet will contaminate HRCCS. This contamination limits the available effective orbital phase coverage, as frames obtained at phases close to secondary-eclipse have to be either heavily processed to remove the stellar pulsations (Johnson et al. 2015; Temple et al. 2017; van Sluijs et al. 2019) or disregarded (Nugroho et al. 2020a; Herman et al. 2022; Spring et al. 2022). In contrast to atomic species often present in both the planet and star, CO is dissociated in the hot stellar atmosphere of early-type A/F stars. These type of stars are also more likely to be pulsating stars due to their location within the Hertzsprung–Russell diagram’s instability strip (e.g. Gautschi & Saio 1996). This highlights a further advantage of using CO in the NIR to probe the atmospheric temperature structure and inversions of UHJ around hot pulsating stars.

6.3 Atmospheric dynamics and longitudinal variation

Hydrodynamical simulations predict the thermal structure of hot Jupiters to be asymmetric around the sub-stellar point, with the hottest point being shifted eastward. This shift arises from the competition between heat transport by winds of planetary-scale waves and the radiative cooling of the parcel of gas (e.g. Showman & Guillot 2002; Perez-Becker & Showman 2013). As the exoplanet orbits its host star, its synchronous rotation causes more of its eastward side to be visible pre-eclipse, and more of its westward side post-eclipse. Consequently, if an eastward hot spot exists, the pre-eclipse phases are more dominated by the hot spot than the post-eclipse phases.

Previous evidence for a hot spot on WASP-33 b comes from NIR and optical photometric phase curve observations, where eastward hot spot results in negative phase curve offset. Zhang et al. (2018) report the first evidence of an eastward hot spot on WASP-33 b,

measuring a phase curve offset of $-12.8 \pm 5.8^\circ$ in the *Spitzer* 3.6 μm band. Conversely, von Essen et al. (2020) found a $+28.7 \pm 7.1^\circ$ westward phase curve offset in the optical *TESS* light curve, and a range of theories including magnetohydrodynamic effects, non-synchronous rotation, and clouds have been invoked to explain such westward offset hot spots (e.g. Dang et al. 2018; Hindle et al. 2021). Recent work by Herman et al. (2022) further report a $+22 \pm 8^\circ$ westward offset based on the phase offset of neutral iron thermal emission lines at high spectral resolution. Our CO emission line detection also supports a phase-dependence in the planet spectrum. However, the GCM that we use to model it includes an eastward (not westward) hot spot, which is oriented towards us during the pre-eclipse phases of WASP-33 b.

It is important to note that the phase-dependence detected in this work and in Herman et al. (2022) is found via a model scaling parameter (a in this work, and a similar $A_p(\phi)$ in equation 4 of the other study). Both studies detect that a larger scaling is required to best model the observations just after eclipse. Herman et al. (2022) interpret the larger $A_p(\phi)$ post-eclipse as an overall brightness variation, indicating more flux west from the substellar point and hence a westward hot spot. This is the typical interpretation for photometric phase curve measurements, which are sensitive to absolute flux as a function of phase. Instead, we highlight that because the HRCCS we use in this work divides out the continuum, that the scaling parameter is a measure only of line contrast, i.e. that the lines in the planet spectrum are deeper just after eclipse, as shown in Fig. 14. The phase-dependent line contrast can then be understood in the context of our GCM spectral template. First, the eastward hot spot results in a hotter continuum and thus greater overall brightness for pre-eclipse phases as shown in the bottom left-hand panel of Fig. 6. But, since our post-processing of the high resolution spectra removes the continuum information we argue we are insensitive to these variations. Fig. 6 then shows that the GCM predicts thermal profiles that are shallower in the eastward-shifted hot spot and steeper on the western part of the day-side. Overall, this

leads to a disc integrated temperature gradient that is shallower during pre-eclipse when it contains the eastward hot spot (where our line-of-sight is centred on longitudes $\sim 17^\circ$ – 61° east of the sub-stellar point), and steeper during post-eclipse when it contains more of the western day-side hemisphere (where our line-of-sight is centred on longitudes $\sim 277^\circ$ – 341° east of the sub-stellar point). Consequently, the average line contrast associated with these temperature gradients will be shallower during pre-eclipse and stronger during post-eclipse. For a fixed line contrast amongst all phases, such as the PHOENIX models or the day-side-only GCM, this results in a smaller scale parameter pre-eclipse compared to post-eclipse, as observed in our data set. But for the 3D phase-dependent model, this results in a more consistent scale parameter between the pre- and post-eclipse data, again as observed in our data set. We therefore conclude that despite our insensitivity to the absolute brightness variations of the hot spot, the phase-dependent variation of the scale parameter is still sensitive to the 3D nature of the atmosphere and consistent with an eastward offset hot spot scenario.

As a check, we also investigate if this result is robust against the alternative scenario of a westward offset hot spot. Such a scenario is not predicted by our GCM, but might arise from magnetohydrodynamical effects on the atmospheric dynamics (Hindle et al. 2021). Including magnetohydrodynamical physics into the GCM is beyond the scope of this work. Instead we simply mirror our phase-dependent GCM model such that eastward phases becomes westward and vice versa, effectively turning our eastward offset model into a westward offset model. The result is shown in Fig. A10. The discrepancy between pre- and post-eclipse phases has significantly increased ($\Delta \log(a) = 0.41_{-0.04}^{+0.03}$) due to the even shallower line contrast post-eclipse compared to pre-eclipse. We conclude a model with a westward hot spot offset is inconsistent with our data, but that the asymmetry therefore is consistent with an eastward offset.

The lower S/N ratio of the phase-dependent GCM compared to the PHOENIX models may seem unexpected at first, also in the light of results from Beltz et al. (2021) who find a significant increase of the S/N detections for the exoplanet HD 209458b when using a 3D GCM compared to a 1D model. In part this may be explained by the lower spectral resolution of this data compared to Beltz et al. (2021) as the 3D effects of the line shape and Doppler shift of the spectral lines will be more pronounced at higher spectral resolutions where they can be resolved. We also compare a single GCM ($N = 1$) to a suite of PHOENIX models ($N = 576$ modified models and $N = 9$ self-consistent models), which is an unfair comparison. It may be explained by current limitations of the GCM, mainly the exclusion of Fe and Fe⁺. Lothringer et al. (2018) find Fe and Fe⁺ are important opacity sources in the atmospheres of UHJ that contribute to their thermal inversions. None the less, we emphasize this phase-dependent change of P–T profile with longitude is robust to the choice inclusion or exclusion of iron opacities in our GCM. When an eastward hot spot shift is present, air west of the hot spot shift is experiencing a transition from cold to hot. This transition happens first at low pressures and propagates downwards. This naturally makes a very strong thermal gradient at this transition between the cold and hot profile and happens at the pressure these observations probe. Once air flows from the hot spot to the cold side the cooling happens more homogeneously over the P–T profile, leading to a less steep gradient. This behaviour does not depend on the details of the hot and cold profiles and should be the same regardless of whether iron opacities were included or not. Instead the phase-dependent variation of the thermal gradient and thus line contrast should be seen as a natural outcome of having a shifted hot spot.

We do not find significant evidence for a relative blueshift pre-eclipse compared to post-eclipse that would further indicate an eastward hot spot. Such net Doppler shifts are predicted of ~ 1 – 3 km s^{−1} as the hot spot rotates in and out of the observers view over while orbiting its host star (Zhang, Kempton & Rauscher 2017; Beltz et al. 2021). However, our velocity sampling is ~ 4 – 6 km s^{−1} for the individual observing nights. Thus we cannot resolve such Doppler shifts at our spectral resolution.

We note that when including all orbital phases out to quadrature in the post-eclipse night, the scaling parameter prefers shallower lines (see grey contours in Fig. 10), despite the strong line contrasts shown in Fig. 14. While these later phases add only a small contribution to the overall S/N, the change in line contrast may be in part explained by the post-processing of the data. The rate of change of the radial velocity for close-to-quadrature phases is small, and therefore our PCA cleaning will be more effective in removal of the exoplanetary signal, and this has also been reported by Brogi et al. (2022) and Pino et al. (2022). We thus emphasize it is key to only compare symmetric phases, ensuring both the pre- and post-eclipse frames included were homogeneously processed by our post-processing steps. We further find a systematic bias towards small scaling parameter $\log a < 0$ in all models. Rather than missing physics, we consider that the data cleaning steps of the HRCCS can affect the line shape and strength of planetary signal. PCA likely degrades more of the planetary signal at lower resolution, where the lines spread across multiple pixels. Applying PCA to our template in every iteration of our CC-to-log(L) PYMULTINEST implementation is computationally expensive. As an alternative, we investigated these effects by injecting the best modified PHOENIX model at $a = 1$ into the frames observed during the eclipse of WASP-33 b. Since the change in radial velocity would be relatively large during the eclipse, potentially making it more robust against PCA, we shifted the injected model in phase to cover similar radial velocities as observed during the first observing night. We found a scale parameter of $\log a = -0.23_{-0.03}^{+0.03}$ for the recovered signal. This result supports our hypothesis that our data cleaning steps cause at least part of the bias towards lower scale factors. Given that we apply the data processing in a homogeneous way to all data sets, and while we do not give large weight to the absolute values of $\log(a)$, we still give weight to the interpretation of the relative differences.

Future work adopting novel model-filtering techniques such as the one introduced by Gibson et al. (2022) may resolve the issue of comparing asymmetric phase coverage pre- and post-eclipse and alleviate the systematic bias towards lower scaling parameters.

6.4 Low resolution limit for HRCCS technique

For the HRCCS technique, to the first order, the S/N ratio of the planet S/N_p is given by

$$S/N_p = \frac{S_p \sqrt{N_{\text{lines}}}}{\sqrt{S_\star + \sigma_{\text{bg}}^2 + \sigma_{\text{read}}^2 + \sigma_{\text{dark}}^2}}, \quad (4)$$

where S_p is the planet's flux, S_\star the stellar flux, N_{lines} the number of resolved spectral lines, σ_{bg} the noise from the sky and telescope background, σ_{read} the detector read-out noise and σ_{dark} the detector dark current (Snellen et al. 2015). In the photon-limited regime, equation (4) reduces to

$$S/N_p = \left(\frac{S_p}{S_\star} \right) S/N_\star \sqrt{N_{\text{lines}}}, \quad (5)$$

where S/N_* is the stellar S/N ratio (Birkby 2018). Our observations are photon-limited thanks to the bright host star. However, residual noise, likely from residual telluric lines, persists in the final CCF time series (see Fig. 7), but the impact of this is substantially mitigated by the large Doppler shift of WASP-33 b.

For the close-in UHJs, we cannot be aided by high contrast imaging to spatially resolve the exoplanet and increase our planet-to-star contrast ratio. We thus rely solely on disentangling the exoplanet's lines from the Doppler shift induced by its orbital motion. We can thus increase our exoplanet's S/N in two ways: First, we can increase our photon collecting power such that the S/N_* increases or secondly increase the total number of spectral lines observed by either expanding our instantaneous wavelength coverage or our spectral resolving power.

Since S/N is proportional to $\sqrt{N_{\text{lines}}}$ it is crucial to have a sufficiently high spectral resolution to resolve more lines. Previously, the lowest spectral resolution at which HRCCS unaided by high contrast imaging was successfully demonstrated was $R = 25\,000$, as shown by the detection of water vapour in the atmosphere of τ Boo with the pre-upgraded NIRSPEC1.0/Keck combination as reported by Lockwood et al. (2014). Other evidence for a detection of the thermal spectrum of HD 88 133 b, also with NIRSPEC1.0/Keck was reported by Piskorz et al. (2016) using a multi-epoch approach, but later on disputed (Buzard et al. 2021). Our detection with ARIES/MMT shows the Doppler-only HRCCS can still be used even at spectral resolutions as low as $R = 14\,000\text{--}16\,000$.

Our detection of WASP-33 b at this lower spectral resolution provides proof-of-concept HRCCS is still possible at a lower resolving power when there is a sufficiently large orbital Doppler shift, despite less and diluted spectral lines being observed. Therefore, one can potentially trade off spectral resolving power for photon collecting power. This is particularly interesting for short period exoplanets where we do not need a high spectral resolving power to resolve their large orbital Doppler shifts. Examples include other UHJs (for example KELT-9b) or terrestrial lava worlds (for example CoRoT-7b) as both have orbital velocities of similar order of magnitude as WASP-33 b.

7 CONCLUSIONS

In this work, we presented the first high resolution spectroscopy observations from the NIR spectrograph ARIES mounted behind the MMT targeting WASP-33 b. We develop a public pre-processing pipeline for ARIES to extract a spectral time series from the raw detector images. We use the HRCCS technique with PCA to characterize WASP-33b's atmosphere using both the S/N method and CC-to-log(L) mapping. The primary conclusions of this work are the following:

(i) We present the first robust detection (7.9σ) of CO emission lines using HRCCS in an exoplanet atmosphere. The detection of CO emission lines in the atmosphere of WASP-33 b is unambiguous evidence of an inverted P–T profile.

(ii) The best self-consistent PHOENIX atmospheric models prefer day-side-only heat redistribution.

(iii) The modified P–T profile PHOENIX models show we are sensitive to the temperature gradient of the inversion layer, but less sensitive to the absolute temperatures of the lower atmosphere.

(iv) The CC-to-log(L) mapping reveals a phase-dependent scaling factor, which shows that the planet spectral line contrast is greater just after secondary eclipse than before. We demonstrate that the P–T profiles from a phase-dependent GCM that includes an eastward

offset hot spot can explain the majority of the phase-dependence in log(a). We emphasize that inclusion of a scaling parameter in the CC-to-log(L) mapping can reveal the longitudinal phase-dependent thermal structure of the atmosphere, even in spectra that lack the resolving power required to detect the additional induced Doppler shift caused by the offset hot spot.

(v) We find no evidence of stellar pulsations affecting our detection of CO, in contrast to previous detections of atomic emission lines in the optical. We advocate that CO is therefore an advantageous and good probe of the thermal atmospheric structure of UHJs at pressures seen with HRCCS.

(vi) At a measured resolving power of $R \sim 15\,000$, these observations show that HRCCS can be used even at spectral resolutions of $R \sim 15\,000$ when using only the large Doppler shift induced by the exoplanet's orbital motion to disentangle its spectrum from its host star and tellurics. That is without the use of any spatial information from High Contrast Imaging. Thus spectral resolving power can be traded for photon collecting power in cases where the induced orbital Doppler shift is sufficient to move the planet spectrum over multiple pixels such as expected for close-in terrestrial lava planets and ultra hot Jupiters.

ARIES is currently being upgraded into the MMT: AO exoPlanet characterization System (MAPS)¹¹ It will have an updated AO system, broader instantaneous wavelength coverage and modes at a higher spectral resolution. A modified version of the ARIES data reduction pipeline presented in this work may prove helpful in analysing future output from MAPS as well.

Finally, future low-resolution observations of WASP-33 b with the *JWST* can be combined with ground-based high-resolution observations to mitigate the effects of stellar pulsations and give tighter constraints on the P–T profile and chemical abundances of OH, H₂O, CO, and CO₂ (Beichman & Greene 2018). UHJ have the additional benefit that refractory elements can be measured as well (Lothringer et al. 2021). Combined this will enable calculation of abundance ratios which can in turn be used to constrain different planet formation pathways (e.g. Öberg, Murray-Clay & Bergin 2011; Madhusudhan 2012; Cridland, Pudritz & Alessi 2016; Cridland et al. 2017, 2019; Khorshid et al. 2021). Finally, combining phase-dependent low resolution data, for example the recently observed diurnal variations in the atmosphere of WASP-121b (Mikal-Evans et al. 2022), and high resolution data using the CC-to-log(L) (e.g. Brogi et al. 2016a; Brogi & Line 2019; Gibson et al. 2020) mapping will greatly improve our understanding of the longitudinal chemical and thermal phase-dependence.

ACKNOWLEDGEMENTS

We thank the anonymous reviewer for the careful reading of our manuscript and the many insightful comments and suggestions. JLB extends special thanks to all the MMT Observatory staff, particularly the telescope and AO operators, for their assistance and valuable advice in carrying out the MMT Exoplanet Atmosphere Survey. We thank K. Powell for all his assistance with the AO instrumentation operation during the process of collecting these data. This research is part of the exoZoo project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement No 805445. Observations reported here were obtained at

¹¹More information can be found here: <https://www.as.arizona.edu/~ktmorz/maps.html>

the MMT Observatory, a joint facility of the Smithsonian Institution and the University of Arizona. This work was performed in part under contract with the Jet Propulsion Laboratory (JPL) funded by NASA through the Sagan Fellowship Program executed by the NASA Exoplanet Science Institute. MB acknowledges support from the STFC research grant ST/T000406/1.

DATA AVAILABILITY

All the original data and pre-processing pipeline is publicly available in the author's Github repository: <https://github.com/lennartvansluij/s/ARIES>.

REFERENCES

- Allard F., Homeier D., Freytag B., 2011, in Johns-Krull C., Browning M. K., West A. A. eds, ASP Conf. Ser. Vol. 448, 16th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun. Astron. Soc. Pac., San Francisco, p. 91
- Arcangeli J. et al., 2018, *ApJ*, 855, L30
- Armstrong D. J., de Mooij E., Barstow J., Osborn H. P., Blake J., Saniee N. F., 2016, *Nat. Astron.*, 1, 0004
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
- Barber R., Tennyson J., Harris G., Tolchenov R., 2006, *MNRAS*, 368, 1087
- Barber R. J., Strange J. K., Hill C., Polyansky O. L., Mellau G. C., Yurchenko S. N., Tennyson J., 2014, *MNRAS*, 437, 1828
- Barman T. S., Hauschildt P. H., Allard F., 2001, *ApJ*, 556, 885
- Barman T. S., Macintosh B., Konopacky Q. M., Marois C., 2011, *ApJ*, 733, 65
- Baxter C. et al., 2021, *A&A*, 648, A127
- Beichman C. A., Greene T. P., 2018, *Observing Exoplanets with the JWST*. Springer International Publishing, New York, p. 85
- Bell T. J. et al., 2019, *MNRAS*, 489, 1995
- Beltz H., Rauscher E., Brogi M., Kempton E. M. R., 2021, *AJ*, 161, 1
- Birkby J. L., 2018, preprint ([arXiv:1806.04617](https://arxiv.org/abs/1806.04617))
- Borsa F. et al., 2021, *A&A*, 653, A104
- Brahm R., Jordán A., Espinoza N., 2017, *PASP*, 129, 034002
- Brogi M., 2022, *Nat. Astron.*, 6, 182
- Brogi M., Line M. R., 2019, *AJ*, 157, 114
- Brogi M., Snellen I. A. G., de Kok R. J., Albrecht S., Birkby J., de Mooij E. J. W., 2012, *Nature*, 486, 502
- Brogi M., de Kok R. J., Birkby J. L., Schwarz H., Snellen I. A. G., 2014, *A&A*, 565, A124
- Brogi M., de Kok R. J., Albrecht S., Snellen I. A. G., Birkby J. L., Schwarz H., 2016, *ApJ*, 817, 106
- Brogi M., Line M., Bean J., Désert J.-M., Schwarz H., 2017, *ApJL*, 389, L2
- Buchner J. et al., 2014, *A&A*, 564, A125
- Buzard C., Piskorz D., Lockwood A. C., Blake G., Barman T. S., Benneke B., Bender C. F., Carr J. S., 2021, *AJ*, 162, 269
- Cabot S. H. C., Madhusudhan N., Welbanks L., Piette A., Gandhi S., 2020, *MNRAS*, 494, 363
- Caffau E., Ludwig H. G., Steffen M., Freytag B., Bonifacio P., 2011, *Sol. Phys.*, 268, 255
- Cauley P. W., Wang J., Shkolnik E. L., Ilyin I., Strassmeier K. G., Redfield S., Jensen A. G., 2021, *AJ*, 161, 152
- Chakrabarty A., Sengupta S., 2019, *AJ*, 158, 39
- Charbonneau D., Knutson H. A., Barman T., Allen L. E., Mayor M., Megeath S. T., Queloz D., Udry S., 2008, *ApJ*, 686, 1341
- Chiavassa A., Brogi M., 2019, *A&A*, 631, A100
- Coles P. A., Yurchenko S. N., Tennyson J., 2019, *MNRAS*, 490, 4638
- Collier Cameron A. et al., 2010, *MNRAS*, 407, 507
- Cont D. et al., 2021, *A&A*, 651, A33
- Cont D. et al., 2022, *A&A*, 657, L2
- Cowan N. B., Agol E., Charbonneau D., 2007, *MNRAS*, 379, 641
- Cridland A. J., Pudritz R. E., Alessi M., 2016, *MNRAS*, 461, 3274
- Cridland A. J., Pudritz R. E., Birnstiel T., Cleeves L. I., Bergin E. A., 2017, *MNRAS*, 469, 3910
- Cridland A. J., van Dishoeck E. F., Alessi M., Pudritz R. E., 2019, *A&A*, 632, A63
- Crossfield I. J. M., Hansen B. M. S., Harrington J., Cho J. Y. K., Deming D., Menou K., Seager S., 2010, *ApJ*, 723, 1436
- Dang L. et al., 2018, *Nat. Astron.*, 2, 220
- Debras F., Mayne N., Baraffe I., Jaupart E., Mourier P., Laibe G., Goffrey T., Thuburn J., 2020, *A&A*, 633, A2
- de Kok R. J., Brogi M., Snellen I. A. G., Birkby J., Albrecht S., de Mooij E. J. W., 2013, *A&A*, 554, A82
- de Kok R. J., Birkby J., Brogi M., Schwarz H., Albrecht S., de Mooij E. J. W., Snellen I. A. G., 2014, *A&A*, 561, A150
- Deming D. et al., 2012, *ApJ*, 754, 106
- de Mooij E. J. W., Brogi M., de Kok R. J., Snellen I. A. G., Kenworthy M. A., Karjalainen R., 2013, *A&A*, 550, A54
- Dobbs-Dixon I., Lin D. N. C., 2008, *ApJ*, 673, 513
- Dobbs-Dixon I., Cumming A., Lin D. N. C., 2010, *ApJ*, 710, 1395
- Ehrenreich D. et al., 2020, *Nature*, 580, 597
- Evans T. M. et al., 2016, *ApJ*, 822, L4
- Evans T. M. et al., 2017, *Nature*, 548, 58
- Evans T. M. et al., 2018, *AJ*, 156, 283
- Feroz F., Hobson M. P., 2008, *MNRAS*, 384, 449
- Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601
- Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2019, *Open J. Astrophys.*, 2, 10
- Fortney J. J., Dawson R. I., Komacek T. D., 2021, *J. Geophys. Res.*, 126, e06629
- Gandhi S., Madhusudhan N., 2019, *MNRAS*, 485, 5817
- Garhart E. et al., 2020, *AJ*, 159, 137
- Gautschy A., Saio H., 1996, *ARA&A*, 34, 551
- Giacobbe P. et al., 2021, *Nature*, 592, 205
- Gibson N. P. et al., 2020, *MNRAS*, 493, 2215
- Gibson N. P., Nugroho S. K., Lothringer J., Maguire C., Sing D. K., 2022, *MNRAS*, 512, 4618
- Goorvitch D., 1994, *ApJS*, 95, 535
- Grimm S. L. et al., 2021, *ApJS*, 253, 30
- Gullikson K., Dodson-Robinson S., Kraus A., 2014, *AJ*, 148, 53
- Hargreaves R., Gordon I., Kochanov R., Rothman L., 2019, *J. Quant. Spectrosc. Radiat. Transfer*, 232, 35
- Hargreaves R. J., Gordon I. E., Rey M., Nikitin A. V., Tyuterev V. G., Kochanov R. V., Rothman L. S., 2020, *ApJS*, 247, 55
- Harrington J., Hansen B. M., Luszcz S. H., Seager S., Deming D., Menou K., Cho J. Y. K., Richardson L. J., 2006, *Science*, 314, 623
- Hauschildt P. H., Baron E., Allard F., 1997, *ApJ*, 483, 390
- Hauschildt P. H., Allard F., Baron E., 1999, *ApJ*, 512, 377
- Haynes K., Mandell A. M., Madhusudhan N., Deming D., Knutson H., 2015, *ApJ*, 806, 146
- Herman M. K., de Mooij E. J. W., Jayawardhana R., Brogi M., 2020, *AJ*, 160, 93
- Herman M. K., de Mooij E. J. W., Nugroho S. K., Gibson N. P., Jayawardhana R., 2022, *AJ*, 163, 248
- Herrero E., Morales J. C., Ribas I., Naves R., 2011, *A&A*, 526, L10
- Hindle A. W., Bushby P. J., Rogers T. M., 2021, *ApJ*, 922, 176
- Hoeijmakers H. J. et al., 2018, *Nature*, 560, 453
- Hoeijmakers H. J. et al., 2019, *A&A*, 627, A165
- Holmberg M., Madhusudhan N., 2022, *AJ*, 164, 79
- Hubeny I., Burrows A., Sudarsky D., 2003, *ApJ*, 594, 1011
- Jackson B., Adams E., Sandidge W., Kreyche S., Briggs J., 2019, *AJ*, 157, 239
- John T. L., 1988, *A&A*, 193, 189
- Johnson M. C., Cochran W. D., Collier Cameron A., Bayliss D., 2015, *ApJ*, 810, L23
- Kanodia S., Wright J., 2018, *Res. Notes Am. Astron. Soc.*, 2, 4
- Karman T. et al., 2019, *Icarus*, 328, 160
- Kesseli A. Y., Snellen I. A. G., Alonso-Floriano F. J., Mollière P., Serindag D. B., 2020, *AJ*, 160, 228

- Khorshid N., Min M., Désert J. M., Woitke P., Dominik C., 2022, *A&A*, 667, A147
- Kitzmann D. et al., 2018, *ApJ*, 863, 183
- Knutson H. A. et al., 2007, *Nature*, 447, 183
- Knutson H. A. et al., 2009, *ApJ*, 690, 822
- Kong H., Akakin H. C., Sarma S. E., 2013, *IEEE Trans. Cybernet.*, 43, 1719
- Kreidberg L., 2018, in Deeg H. J., Belmonte J. A., eds, *Handbook of Exoplanets*. Springer, Berlin, p. 2083
- Kreidberg L. et al., 2018, *AJ*, 156, 17
- Lee E., Taylor J., Grimm S. L., Baudino J.-L., Garland R., Irwin P. G. J., Wood K., 2019, *MNRAS*, 487, 2082
- Lee E. K. H. et al., 2022, *ApJ*, 929, 180
- Lehmann H., Guenther E., Sebastian D., Döllinger M., Hartmann M., Mkrtychian D. E., 2015, *A&A*, 578, L4
- Li G., Gordon I. E., Rothman L. S., Tan Y., Hu S.-M., Kassi S., Campargue A., Medvedev E. S., 2015, *ApJS*, 216, 15
- Line M. R. et al., 2021, *Nature*, 598, 580
- Lockwood A. C., Johnson J. A., Bender C. F., Carr J. S., Barman T., Richert A. J. W., Blake G. A., 2014, *ApJ*, 783, L29
- Lodders K., Fegley B., 2002, *Icarus*, 155, 393
- Lord S. D., 1992, in NASA Technical Memorandum 103957. Available at <https://ntrs.nasa.gov/citations/19930010877>
- Lothringer J. D., Barman T., 2019, *ApJ*, 876, 69
- Lothringer J. D., Barman T., Koskinen T., 2018, *ApJ*, 866, 27
- Lothringer J. D., Rustamkulov Z., Sing D. K., Gibson N. P., Wilson J., Schlaufman K. C., 2021, *ApJ*, 914, 12
- Madhusudhan N., 2012, *ApJ*, 758, 36
- Madhusudhan N., Seager S., 2009, *ApJ*, 707, 24
- Mansfield M. et al., 2018, *AJ*, 156, 10
- Marsh T. R., 1989, *PASP*, 101, 1032
- McCarthy D. W., Burge J. H., Angel J. R. P., Ge J., Sarlot R. J., Fitz-Patrick B. C., Hinz J. L., 1998, in Fowler A. M. ed., *Proc. SPIE Conf. Ser. Vol. 3354, Infrared Astronomical Instrumentation*. SPIE, Bellingham, p. 750
- Menou K., Rauscher E., 2009, *ApJ*, 700, 887
- Merritt S. R. et al., 2021, *MNRAS*, 506, 3853
- Mikal-Evans T. et al., 2022, *Nat. Astron.*, 6, 471
- Morzinski K. M. et al., 2020, in Schreiber L., Schmidt D., Vernet E., eds, *Proc. SPIE Conf. Ser. Vol. 11448, Adaptive Optics Systems VII*. SPIE, Bellingham, 114481L
- Nugroho S. K., Kawahara H., Masuda K., Hirano T., Kotani T., Tajitsu A., 2017, *AJ*, 154, 221
- Nugroho S. K., Gibson N. P., de Mooij E. J. W., Watson C. A., Kawahara H., Merritt S., 2020a, *MNRAS*, 496, 504
- Nugroho S. K., Gibson N. P., de Mooij E. J. W., Herman M. K., Watson C. A., Kawahara H., Merritt S. R., 2020b, *ApJ*, 898, L31
- Nugroho S. K. et al., 2021, *ApJ*, 910, L9
- Öberg K. I., Murray-Clay R., Bergin E. A., 2011, *ApJ*, 743, L16
- Parmentier V. et al., 2018, *A&A*, 617, A110
- Perez-Becker D., Showman A. P., 2013, *ApJ*, 776, 134
- Pino L. et al., 2020, *ApJ*, 894, L27
- Pino L. et al., 2022, *A&A*, 668, A176
- Piskorz D. et al., 2016, *ApJ*, 832, 131
- Polyansky O. L., Kyuberis A. A., Zobov N. F., Tennyson J., Yurchenko S. N., Lodi L., 2018, *MNRAS*, 480, 2597
- Rauscher E., Menou K., 2010, *ApJ*, 714, 1334
- Rodler F., Lopez-Morales M., Ribas I., 2012, *ApJ*, 753, L25
- Rodler F. et al., 2013, *MNRAS*, 432, 1980
- Rothman L. S. et al., 2009, *J. Quant. Spectrosc. Radiat. Transfer*, 110, 533
- Sarlot R. J., McCarthy D. W., Burge J. H., Ge J., 1999, in Fischer R. E., Smith W. J. eds, *Proc. SPIE Conf. Ser. Vol. 3779, Current Developments in Optical Design and Optical Engineering VIII*. SPIE, Bellingham, p. 274
- Serindag D. B., Nugroho S. K., Mollière P., de Mooij E. J. W., Gibson N. P., Snellen I. A. G., 2021, *A&A*, 645, A90
- Sheppard K. B., Mandell A. M., Tamburo P., Gandhi S., Pinhas A., Madhusudhan N., Deming D., 2017, *ApJ*, 850, L32
- Showman A. P., Guillot T., 2002, *A&A*, 385, 166
- Showman A. P., Polvani L. M., 2011, *ApJ*, 738, 71
- Showman A. P., Fortney J. J., Lian Y., Marley M. S., Freedman R. S., Knutson H. A., Charbonneau D., 2009, *ApJ*, 699, 564
- Showman A. P., Tan X., Parmentier V., 2020, *Space Sci. Rev.*, 216, 139
- Smith A. M. S., Anderson D. R., Skillen I., Collier Cameron A., Smalley B., 2011, *MNRAS*, 416, 2096
- Snellen I. A. G., de Kok R. J., de Mooij E. J. W., Albrecht S., 2010, *Nature*, 465, 1049
- Snellen I. et al., 2015, *A&A*, 576, A59
- Spring E. F. et al., 2022, *A&A*, 659, A121
- Stone J. M., Eisner J. A., Salyk C., Kulesa C., McCarthy D., 2014, *ApJ*, 792, 56
- Swain M. R., Vasisht G., Tinetti G., Bouwman J., Chen P., Yung Y., Deming D., Deroo P., 2009, *ApJ*, 690, L114
- Tan X., Komacek T. D., 2019, *ApJ*, 886, 26
- Temple L. Y. et al., 2017, *MNRAS*, 471, 2743
- Tsai S.-M., Malik M., Kitzmann D., Lyons J. R., Fateev A., Lee E., Heng K., 2021, *ApJ*, 923, 264
- van Sluijs L. et al., 2019, *A&A*, 626, A97
- Visscher C., Lodders K., Fegley Bruce, J., 2010, *ApJ*, 716, 1060
- von Essen C., Mallonn M., Welbanks L., Madhusudhan N., Pinhas A., Bouy H., Weis Hansen P., 2019, *A&A*, 622, A71
- von Essen C., Mallonn M., Borre C. C., Antoci V., Stassun K. G., Khalafinejad S., Tautvaišienė G., 2020, *A&A*, 639, A34
- Wilks S. S., 1938, *Ann. Math. Stat.*, 9, 60
- Wong I. et al., 2016, *ApJ*, 823, 122
- Wright J. T., Eastman J. D., 2014, *PASP*, 126, 838
- Yan F. et al., 2019, *A&A*, 632, A69
- Yan F. et al., 2020, *A&A*, 640, L5
- Yan F. et al., 2021, *A&A*, 645, A22
- Yan F. et al., 2022, *A&A*, 661, L6
- Yurchenko S. N., Mellor T. M., Freedman R. S., Tennyson J., 2020, *MNRAS*, 496, 5282
- Zhang J., Kempton E. M. R., Rauscher E., 2017, *ApJ*, 851, 84
- Zhang M. et al., 2018, *AJ*, 155, 83
- Zucker S., 2003, *MNRAS*, 342, 1291

APPENDIX A: ADDITIONAL TABLES AND FIGURES

For transparency, completeness, and reproducibility of this work, we include some additional figures and tables here.

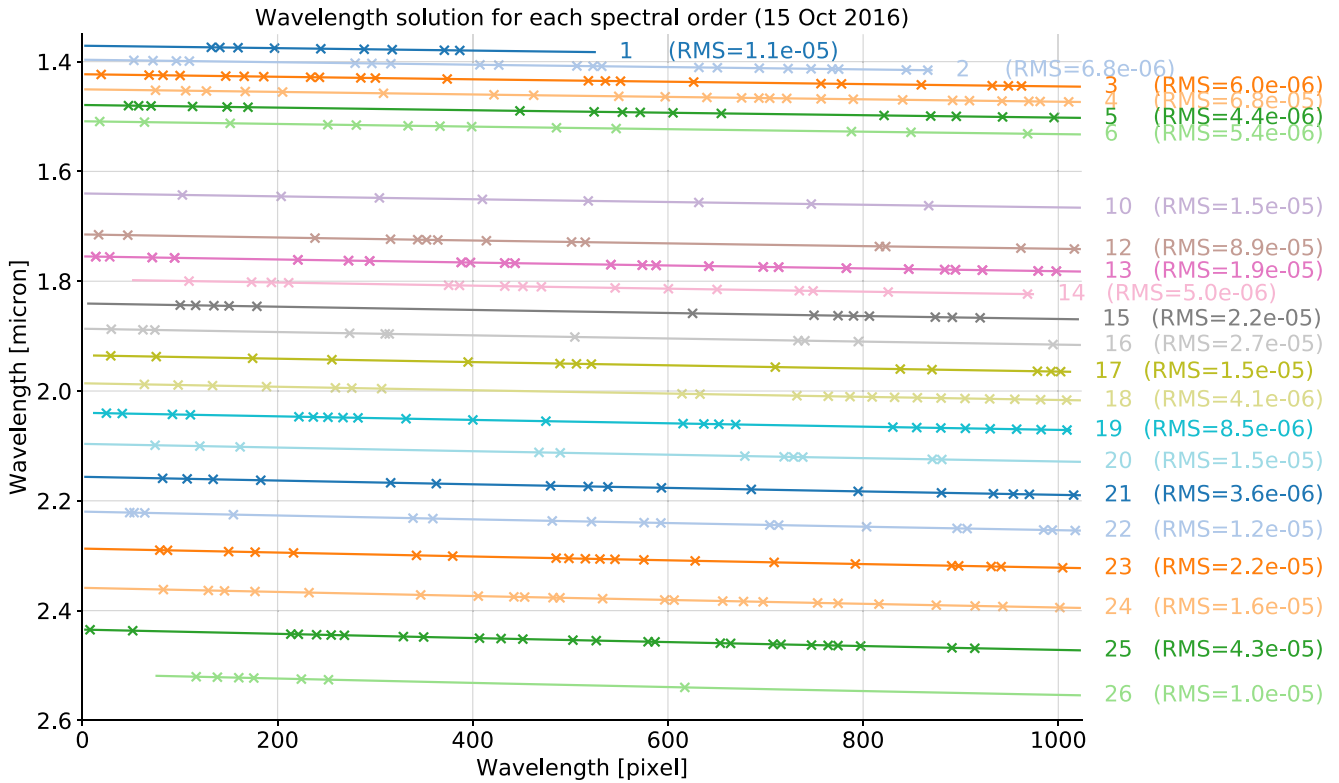


Figure A1. Wavelength solution for each spectral order for the first observing night. For each order the points indicate visually identified telluric absorption lines and the lines indicate the best-fitting polynomial. The root mean square (RMS) values of the difference between the points and the wavelength solution are noted to indicate quality of the fit for each spectral order. Some masked orders for which no good wavelength solution was found have been omitted.

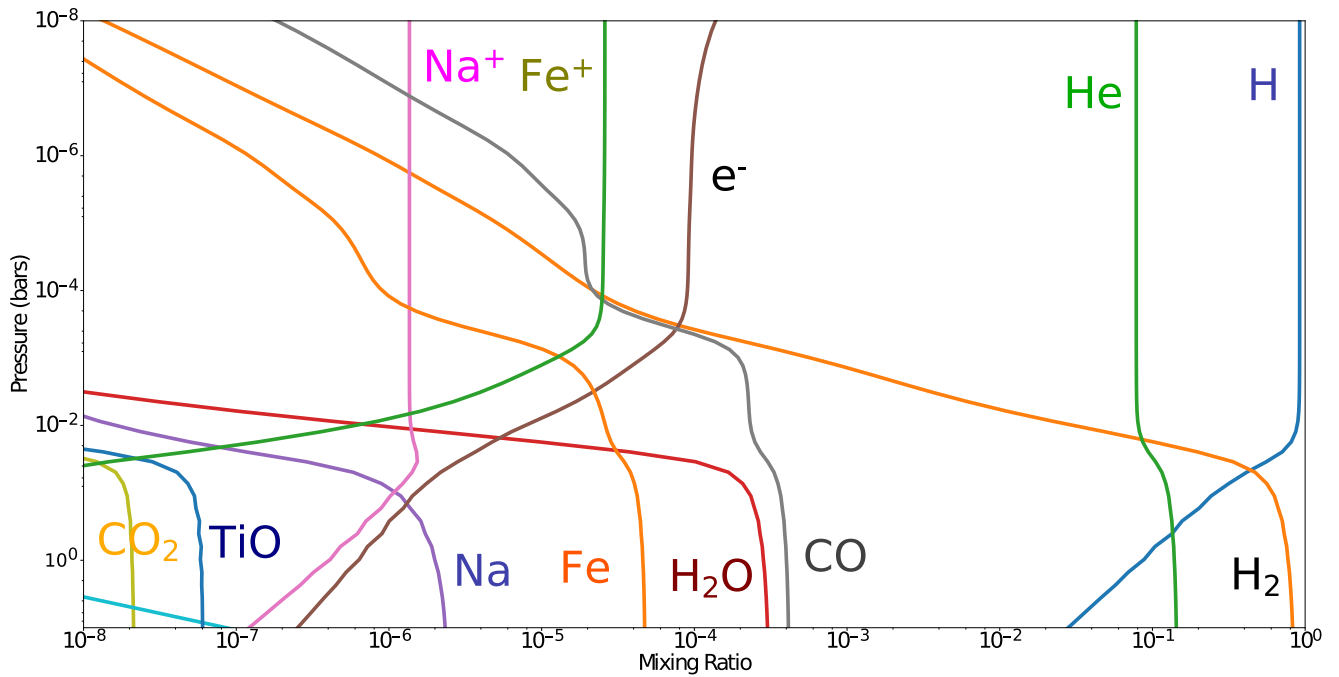
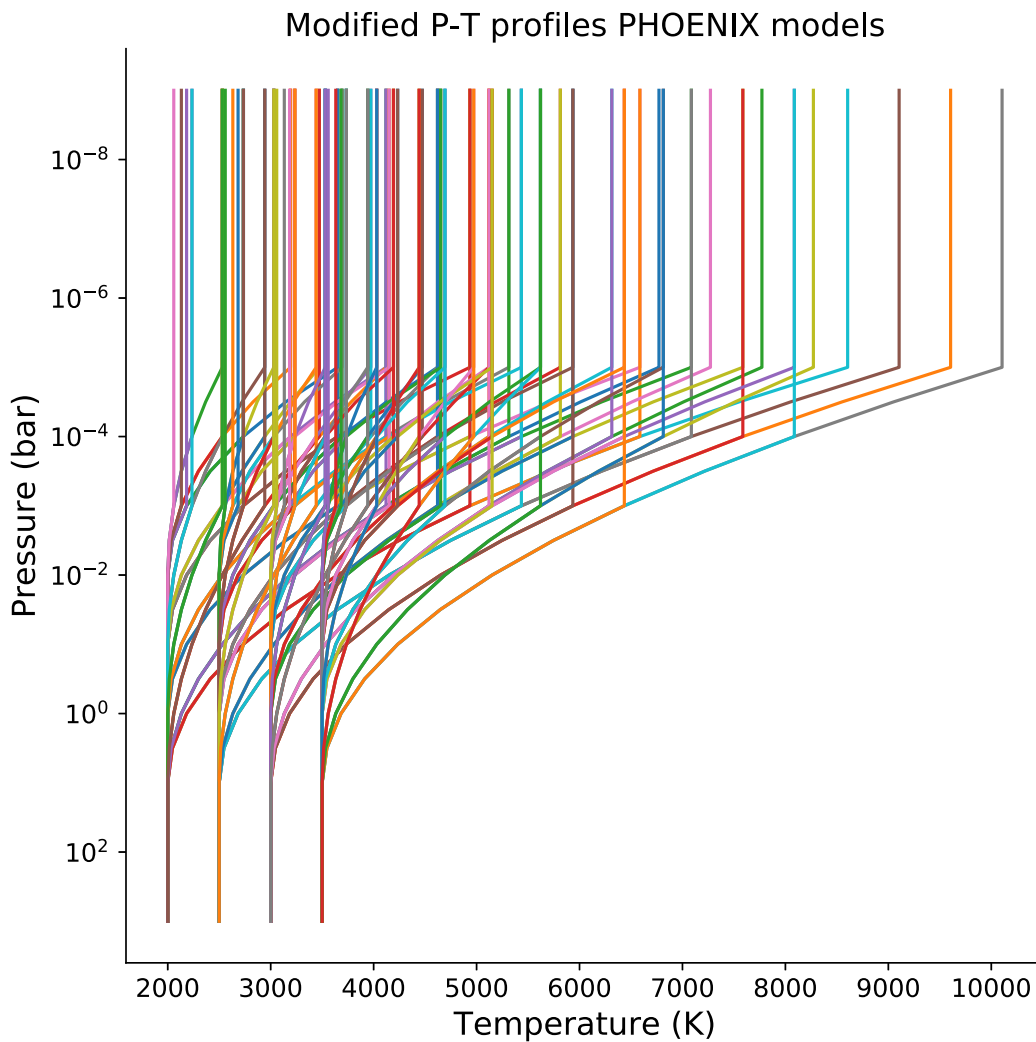


Figure A2. Mixing ratios in chemical equilibrium as functions of pressure for different species in the self-consistent PHOENIX model atmospheres of WASP-33 b. This is for the day-side-only heat redistribution model $f = 0.5$ at $\times 1$ Solar metallicity.

Table A1. An overview of relevant WASP-33 system parameters used in this work.

Quantity	Value	Unit	Reference
Stellar mass	1.495	M_{\odot}	Collier Cameron et al. (2010)
Stellar effective temperature	7400	K	Collier Cameron et al. (2010)
Stellar radius	1.444	R_{\odot}	Collier Cameron et al. (2010)
Primary transit time	2454590.17936	BJD	Smith et al. (2011)
Eccentricity	0.0	–	Smith et al. (2011)
Semimajor axis	0.02558	au	Smith et al. (2011)
Orbital period	1.21986967	days	Smith et al. (2011)
Orbital inclination	87.7	degree	Lehmann et al. (2015)
Planetary radius	1.603	R_{Jup}	Lehmann et al. (2015)
Impact parameter	0.21	–	Chakrabarty & Sengupta (2019)
Expected system velocity	–0.3	km s^{-1}	Nugroho et al. (2020b)
Expected orbital velocity	230.9	km s^{-1}	Nugroho et al. (2020b)

**Figure A3.** Overview of all modified PHOENIX P–T profiles explored in this work.

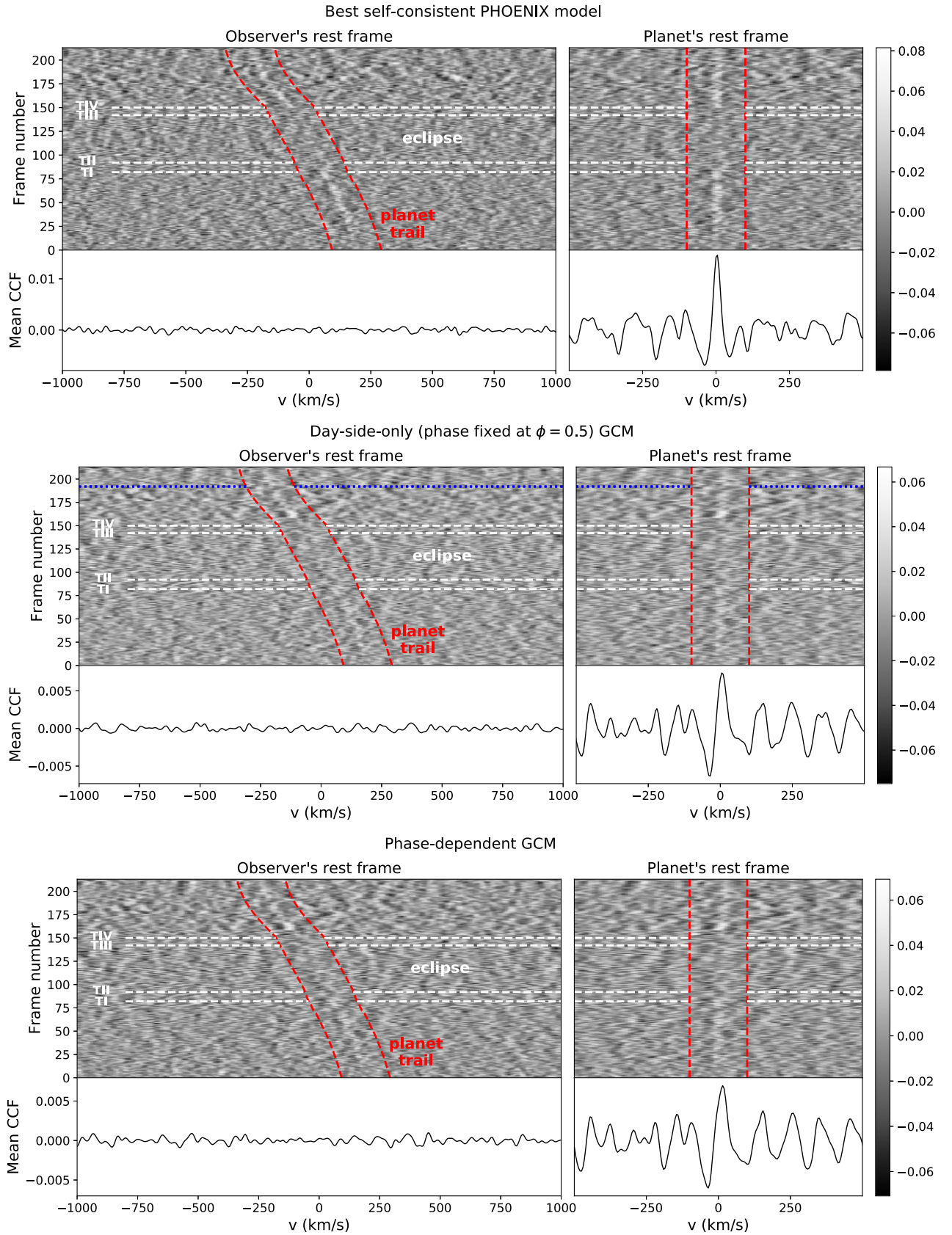


Figure A4. Similar to Fig. 7, but for the highest S/N-ratio self-consistent PHOENIX model, day-side-only GCM (middle panel) and phase-dependent GCM (bottom panel).

Best-modified P-T profile PHOENIX model

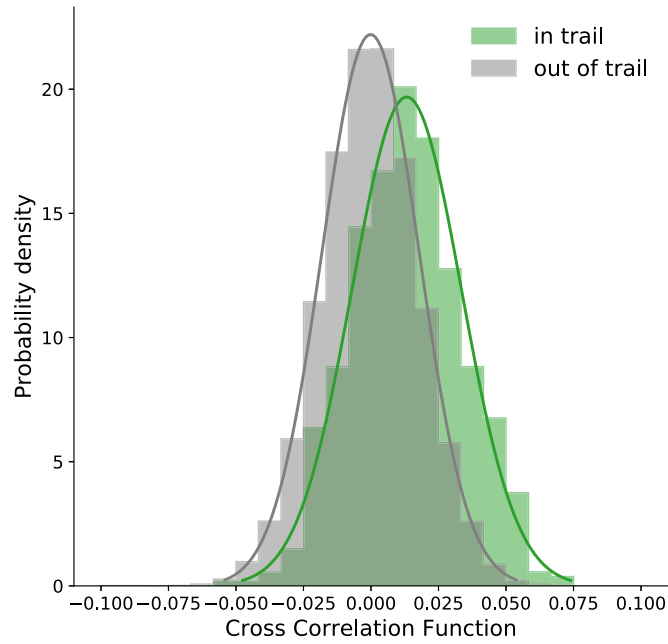


Figure A5. Histogram showing the probability density of the cross-correlation values in-trail (three most central columns) and out-of-trail (outside of the ten most central columns) of the exoplanet's velocity trail after aligning to the planet's rest frame. Values are shown for the PHOENIX best modified P-T profile model. The overlotted solid lines are Gaussian fits to the distributions, which are well described by the Normal function.

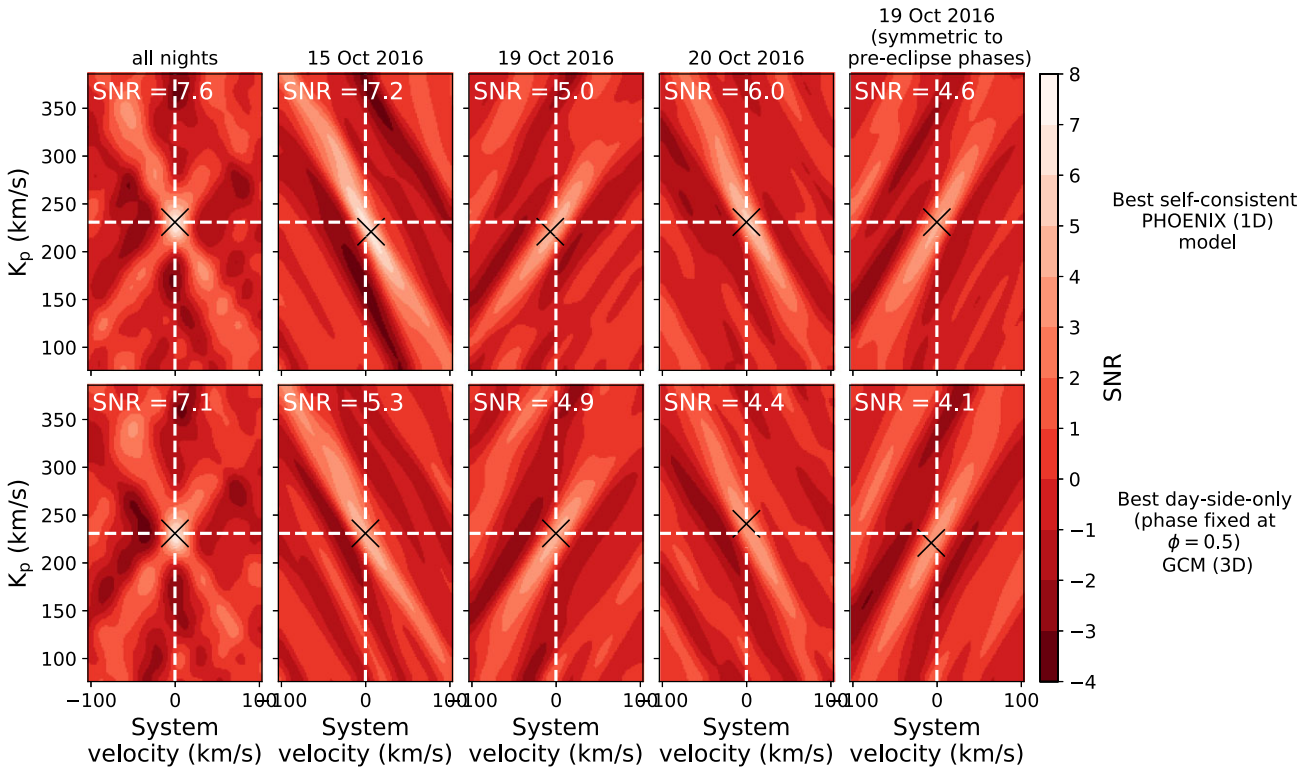


Figure A6. Same as Fig. 9 but shown for the best self-consistent PHOENIX model (top row), best day-side-only GCM (bottom row).

Downloaded from https://academic.oup.com/mnras/article/522/2/2145/7127706 by guest on 04 August 2023

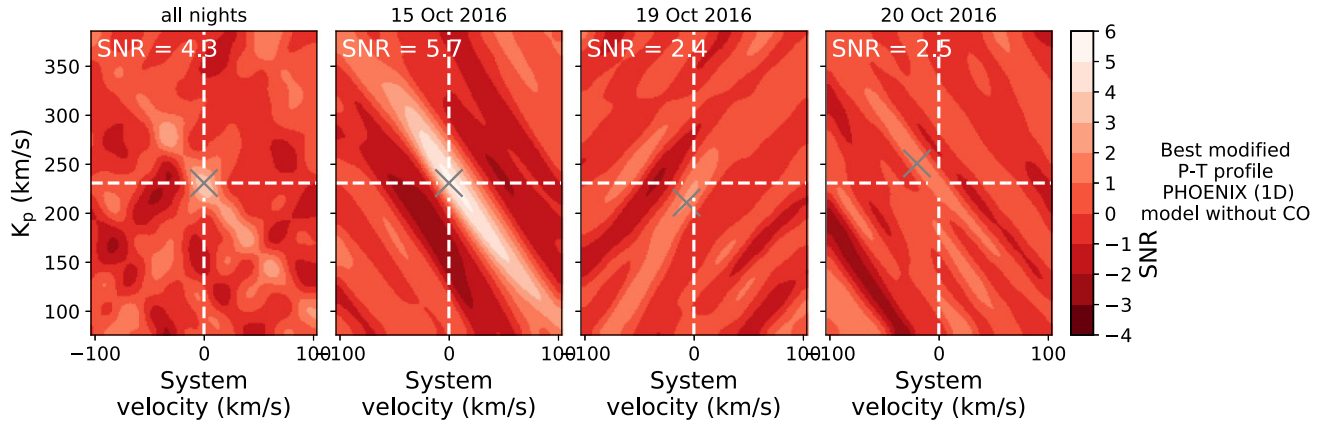


Figure A7. Similar to Fig. 9, but for the best modified P–T profile PHOENIX model with all opacity sources included except CO. The detection of 5.7σ for the first night suggests we are still sensitive to some of the other opacity sources in the atmosphere of WASP-33 b.

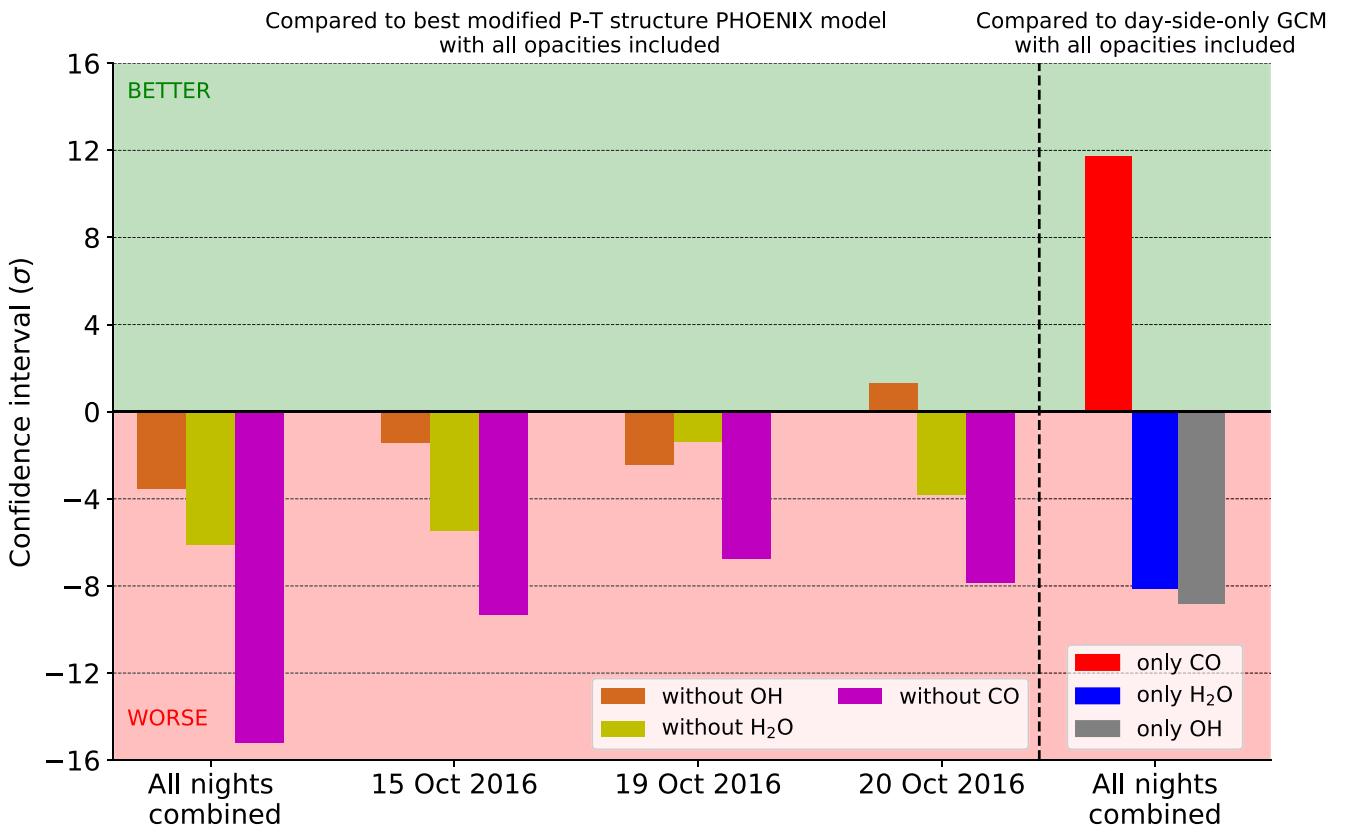


Figure A8. Comparison of the best modified P–T profile PHOENIX model (left-hand panel) and day-side-only GCM (right-hand panel) with all opacity sources included to models without/only certain opacity sources included. A negative confidence interval indicates the model is worse compared to the model where all opacity sources are included, whereas a positive confidence interval indicates the model is better. Confidence intervals are calculated using Wilks’ Theorem.

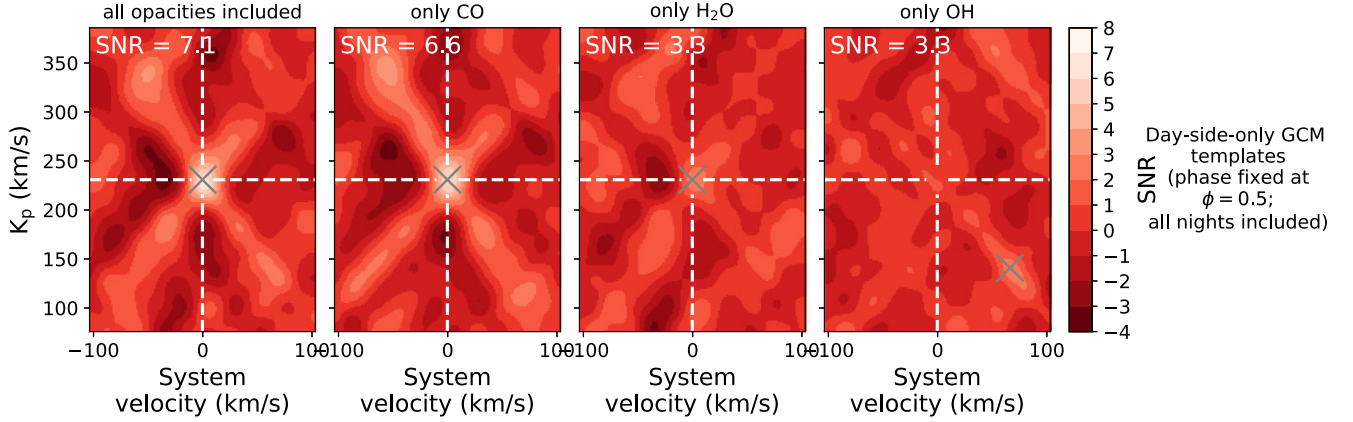


Figure A9. Similar to Fig. 9, but for the GCM day-side-only templates with respectively: all opacity sources, only CO, only H₂O, and only OH included. Amongst the opacity sources, CO is the only molecule that is robustly detected at 6.6σ .

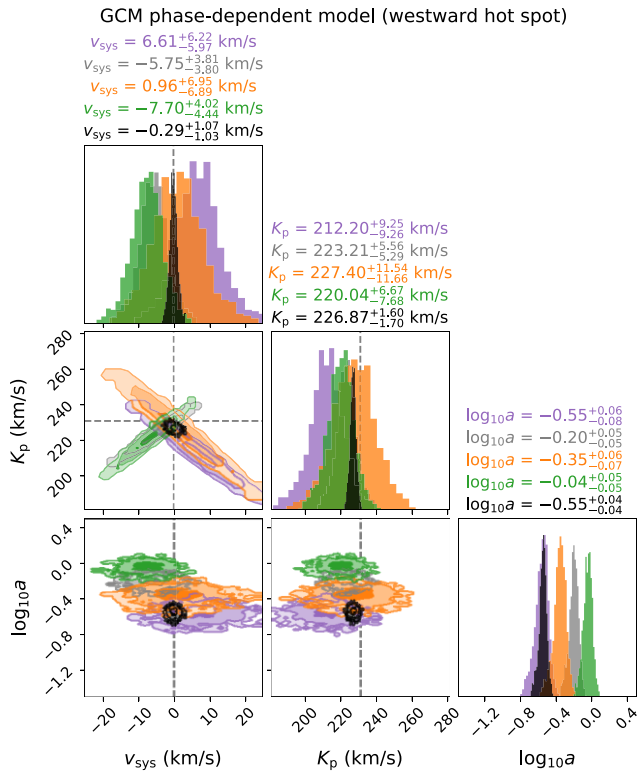


Figure A10. Same as the panels in Fig. 10, but results shown for the GCM phase-dependent model where all phases have been mirrored such that eastward offsets become westward, effectively turning our model into one with a westward hot spot offset.

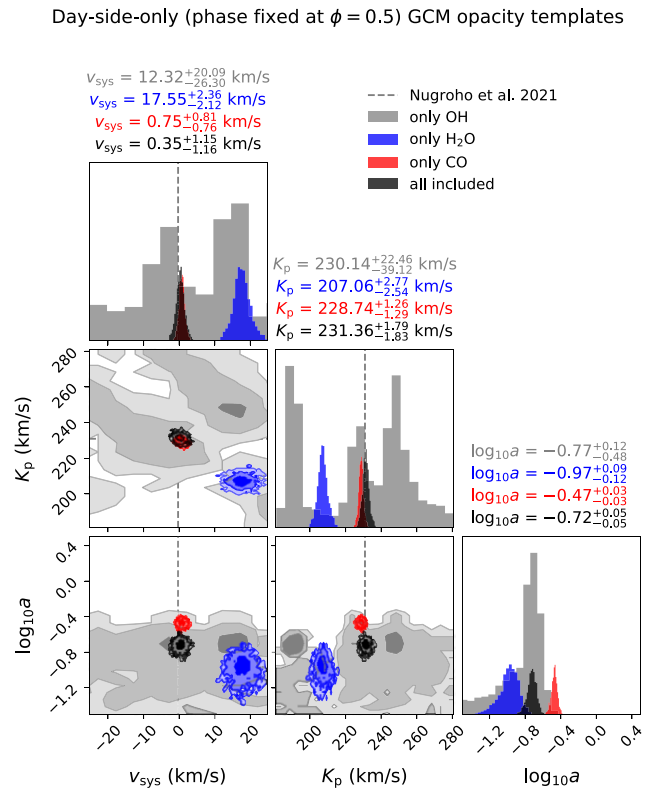


Figure A11. Same as the panels in Fig. 10, but results shown for the GCM day-side-only spectral templates for different opacity sources included. All results are shown for all nights combined.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.