nature astronomy

Supplementary information

https://doi.org/10.1038/s41550-025-02710-8

Continuous helium absorption from both the leading and trailing tails of WASP-107 b

In the format provided by the authors and unedited

Supplementary Information File

Data reduction pipelines: NAMELESS reduction and ExoTEP light curve fitting

In addition to `exoTEDRF`, we also use the `NAMELESS` pipeline [1–3] to reduce our *JWST* NIRISS-SOSS observations of WASP-107 b following a similar procedure to those presented in [4] and [3]. Starting from the uncalibrated data, we go through the Stages 1 and 2 steps of the STScI `jwst` v1.12.5 pipeline [5]. We subsequently correct for bad pixels, non-uniform background, remaining cosmic rays, and $1/f$ noise at the integration level using a series of custom steps. Bad pixels are flagged by looking for outliers in the second derivative of the frames, as described in [1], and their values are replaced using bicubic interpolation. The non-uniform background is subtracted by individually scaling the two regions of the model background provided by the STScI [1] which are separated by the jump in flux near x$\sim$700. We also correct for any cosmic rays remaining after the jump detection step of the `jwst` pipeline by computing the running median in time of all individual pixels and clipping any value that is more than 4σ away from its median. Finally, we correct for the $1/f$ noise following the method described in [3, 4], where each column of the first and second spectral orders is scaled individually, considering only pixels within a 30-pixel distance from the center of the traces to compute the $1/f$ noise. The light curves are then extracted from the first and second spectral orders using a box aperture with a width of 36 pixels.

We fit the white-light and spectroscopic light curves produced by our `NAMELESS` reduction using the ExoTEP framework [6, 7]. For the white-light curve fit, produced by summing all wavelengths of the first spectral order (0.85–2.85 μm), we keep the time of mid-transit ($T_0, \mathcal{U}[2460107, 2460108]$ BJD), planet-to-star radius ratio (R_p/R_s, $\mathcal{U}$[0.01,0.2]), semi-major axis (a/R_s, $\mathcal{U}$[10,30]), and impact parameter (b, $\mathcal{U}$[0,1]) as free parameters. We consider the quadratic limb-darkening law and fit directly for u_1 and u_2, assuming large uniform priors (U[-3,3]) to avoid the introduction of potential biases in the transmission spectrum [8]. For the systematics model, we consider a linear trend with slope v ($\mathcal{U}$[-10^9,10^9]) and intercept c ($\mathcal{U}$[-10^9,10^9]). ExoTEP models the transit light curves using the `batman` python package [9] and explores the parameter space using the Markov chain Monte Carlo (MCMC) sampler `emcee` [10]. We run `emcee` for 10,000 steps, discarding the first 6,000 steps as burn-in. We repeat the same procedure for the spectroscopic light curves, which we fit at the native resolution (one light curve per detector column), having fixed the orbital parameters of WASP-107 b to the best-fit values from the white-light curve ($T_0 = 2460107.5058791$ BJD, $a/Rs = 18.082$, $b = 0.087$). The resulting spectrum in shown in Supplementary Figure 4, in comparison with the spectra from `exoTEDRF` pipeline. We find that the overall shape of the two spectra is similar, with a mean absolute deviation of 0.76 σ. However, there is an increasing scatter between the two reductions at wavelengths below $\sim$0.85 μm (corresponding to the order 2 data) and above $\sim$2.2 μm, where the throughput of the instrument is significantly lower and the spectra are most sensitive to data reduction choices. These differences are most likely attributable to distinct treatments of

[1] https://jwst-docs.stsci.edu/

the $1/f$ noise and subtraction of the background, which impact the precision of the transmission spectra and dilute its amplitude.

To test the impact of the differences between the two spectra on the inferred atmospheric properties, we ran `pRT` retrievals on the `NAMELESS`-reduced dataset (Supplementary Figure 5). We find that the measured water and potassium abundances from the two reductions are consistent at less than $1\,\sigma$, and that they provide virtually the same NH_3 abundance upper limit. Moreover, we show that the extracted helium transit light curves from `NAMELESS` and `exoTEDRF` exhibit the same strong signal of pre- and post-transit absorption.

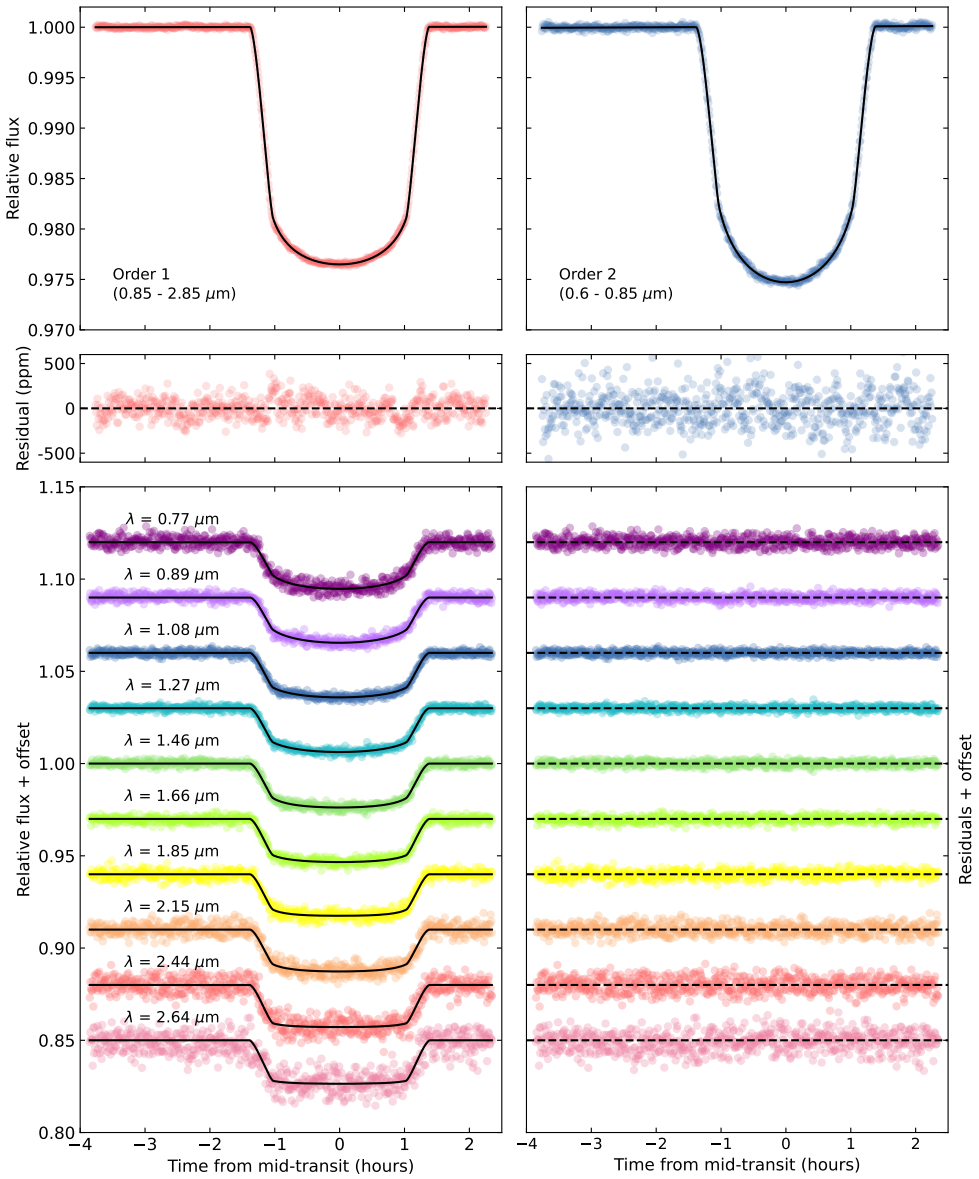
Significance of light curve fits

We performed linear de-trending while fitting the broadband order 2 light curve. However, the Bayesian Information Criterion (BIC) favors the model without linear de-trending, with a ΔBIC of 24.936. The corner plots for order 1 without de-trending and for order 2 with and without de-trending are shown in Supplementary Figures 6, 7, and 8, respectively. It is clear that the transit depths extracted from order 2 with and without de-trending are virtually identical, differing by only ~ 3 ppm ($30\times$ smaller than the standard deviation), and thus do not impact the nature of our results. This is particularly important in the context of pixel-resolution fits, where we aim to detect an extended helium atmosphere that begins well before and persists long after the optical transit duration. Therefore, we did not apply any de-trending algorithms.

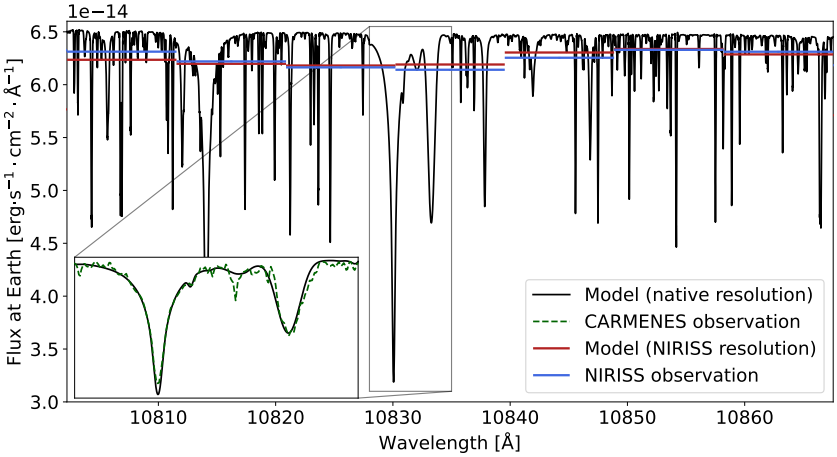
The pixel-resolution light curve fit centered on the helium line (λ: 1.0835 μm) is shown in Supplementary Figure 9. We can see the depth of the helium line is well-constrained and shows deeper transit depth than the one outside the line (see Supplementary Figure 10).

Retrieval frameworks

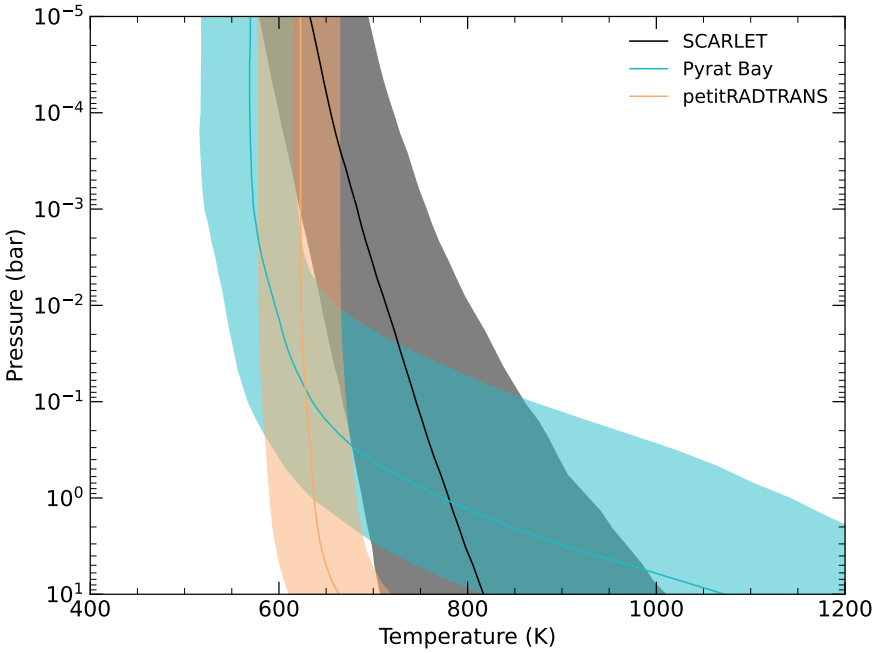
Our retrieval frameworks – `SCARLET`, `petitRADTRANS`, and `Pyrat Bay`—fit the overall trend of the SOSS data, with additional deviations at shorter wavelengths ($\lambda < 1\mu$m). Since we attribute the short-wavelength slope to stellar heterogeneity rather than hazes, the `pRT` and `Pyrat Bay` retrievals are unable to capture this feature. The thermal profile of our retrievals are presented in Supplementary Figure 3. The corner plots of the best-fit free-chemistry models from `SCARLET`, `pRT`, and `Pyrat Bay` are shown in Supplementary Figures 11, 12, and 13, respectively. Additionally we compare the spot vs. hazes with our `SCARLET` retrieval and show that spots are preferred with a $\Delta \ln Z = 11.62$ over the pure atmospheric hazes model. This corresponds to a $5.2\,\sigma$ detection of spots (see Supplementary Figure 14. The posterior corners from the `SCARLET` model without spots are highlighted in Supplementary Figure 15.



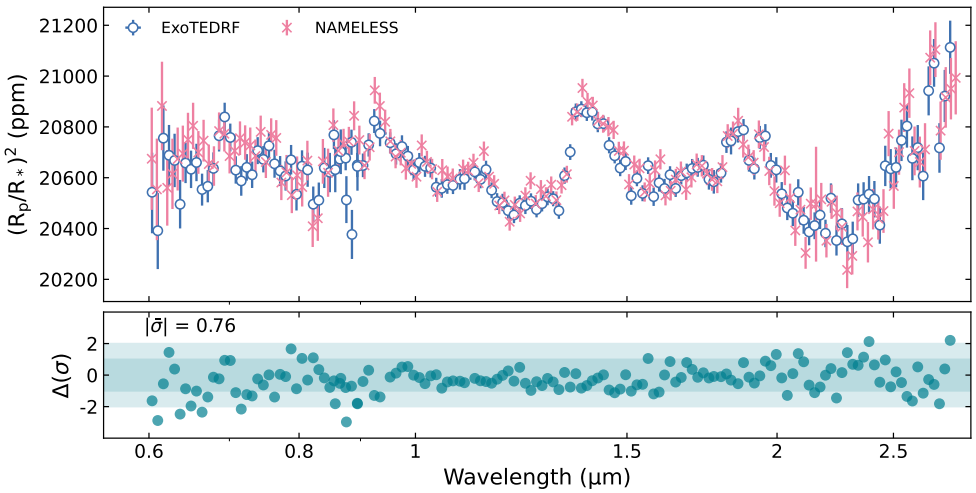
Supplementary Figure 1: *Top:* Broadband `exoTEDRF` light curves of NIRISS-SOSS order 1 (red) & 2 (blue) and best-fit transit models (black). *Middle:* The corresponding residuals. *Bottom:* Representative instrumental resolution light curves with their residuals on the right. No systematic trends have been removed from the data.



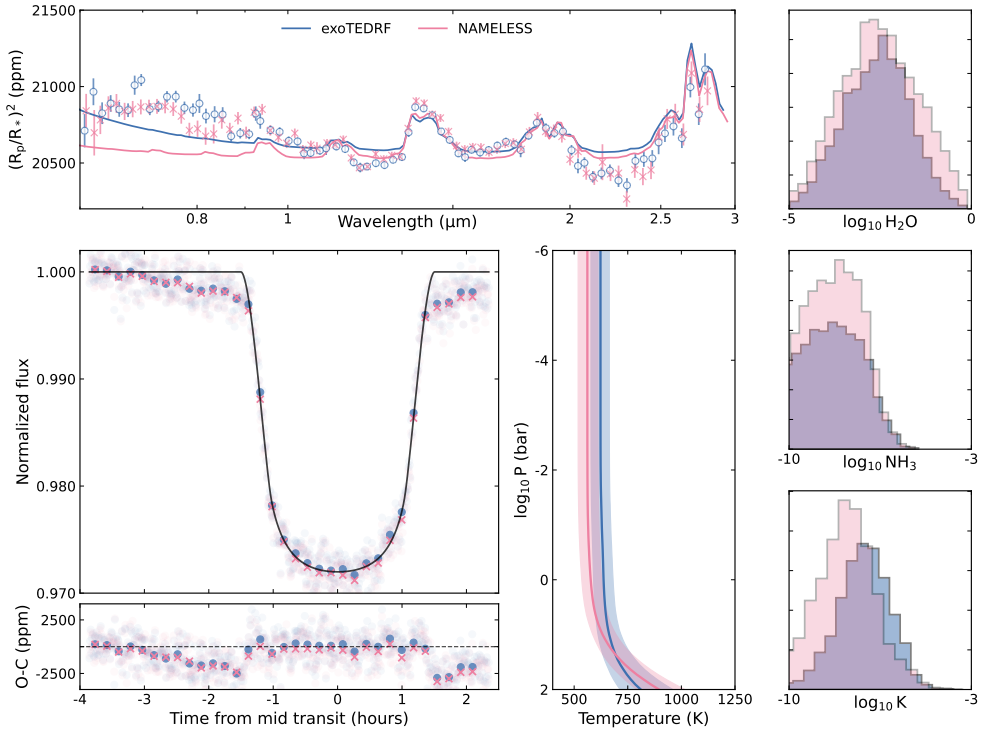
Supplementary Figure 2: Master-out model used for the helium modeling. We built our model spectrum by combining CARMENES observations and *Turbospectrum* simulations. The zoom-in window corresponds to the wavelength range where the stellar spectrum is synthetic, while outside it corresponds to the CARMENES observations of [11]. We compare our reconstruction to the disk-integrated stellar spectrum as seen by NIRISS-SOSS to adjust the level of the continuum.



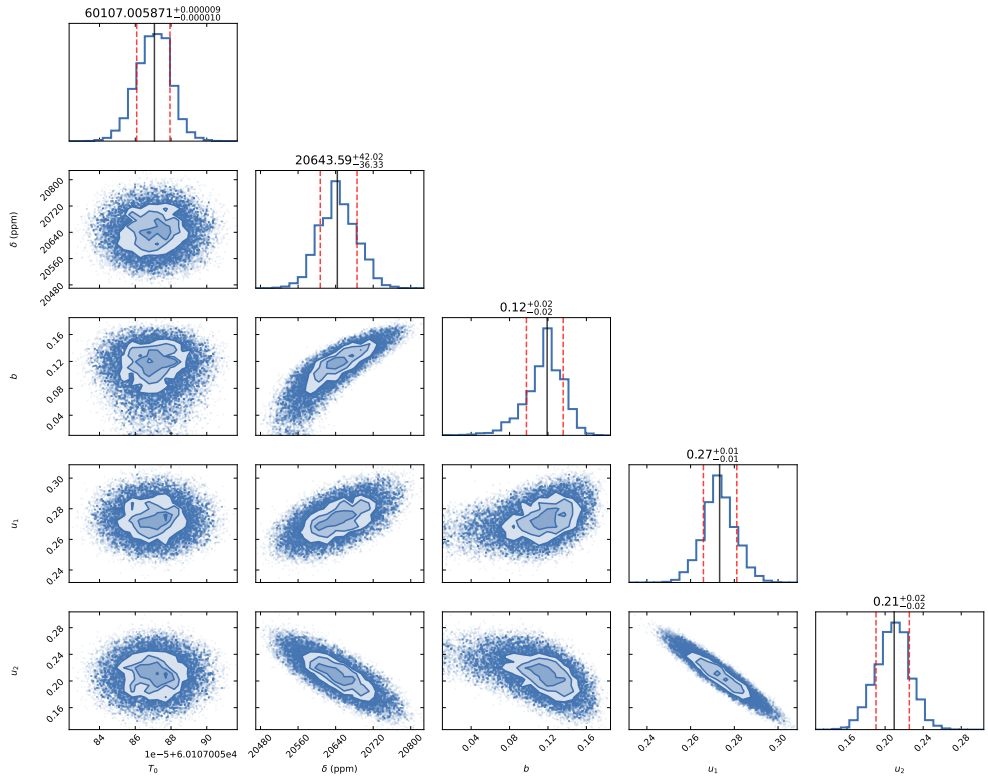
Supplementary Figure 3: Temperature-pressure profiles for the `SCARLET` (nominal setup), `Pyrat Bay` and `petitRADTRANS` ("definite" retrieval model) retrievals on the `exoTEDRF` dataset. The colors match Figure 4. For each retrieval, we show the median, $1\,\sigma$ contours from the posterior samples on the T-P profile. Although we use a non-uniform T-P profile with `SCARLET`, we retrieve a near-uniform temperature structure.



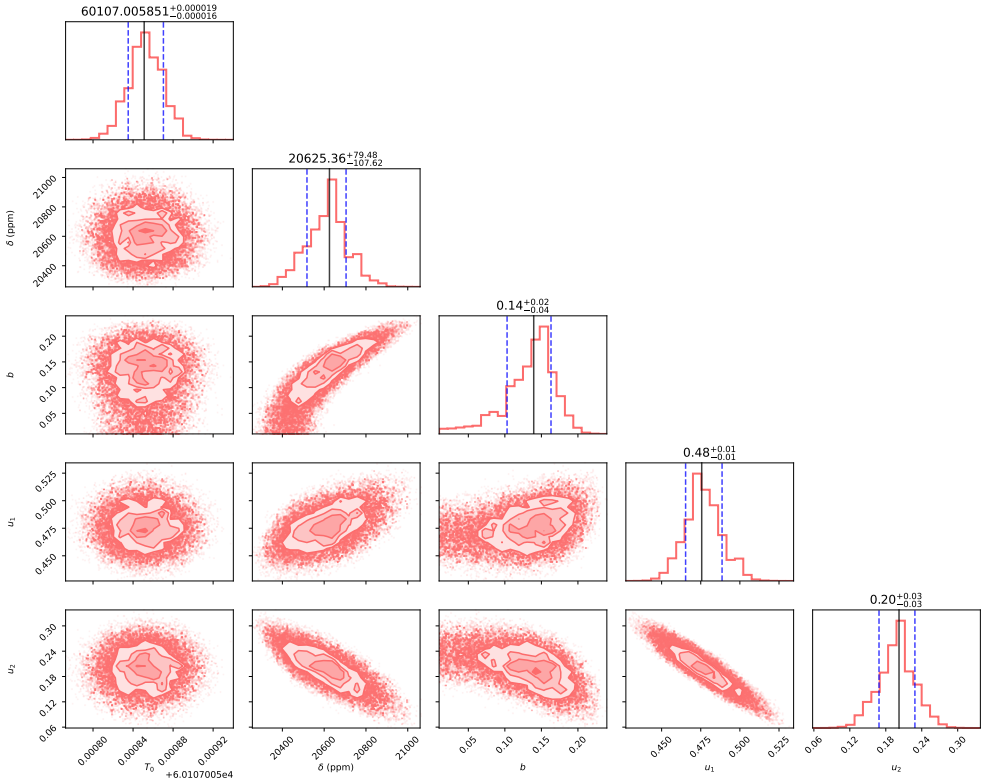
Supplementary Figure 4: Transmission spectra of WASP-107 b extracted from `NAMELESS` and `exoTEDRF` pipelines, binned down to R$\sim$100 and shown with 1 σ error bars. The bottom planel shows the difference in transit depths i.e (`exoTEDRF - NAMELESS`) scaled by the maximum difference. There is no trend in the differences, indicating the pipelines are in good agreement with an average scatter of 0.76 σ.



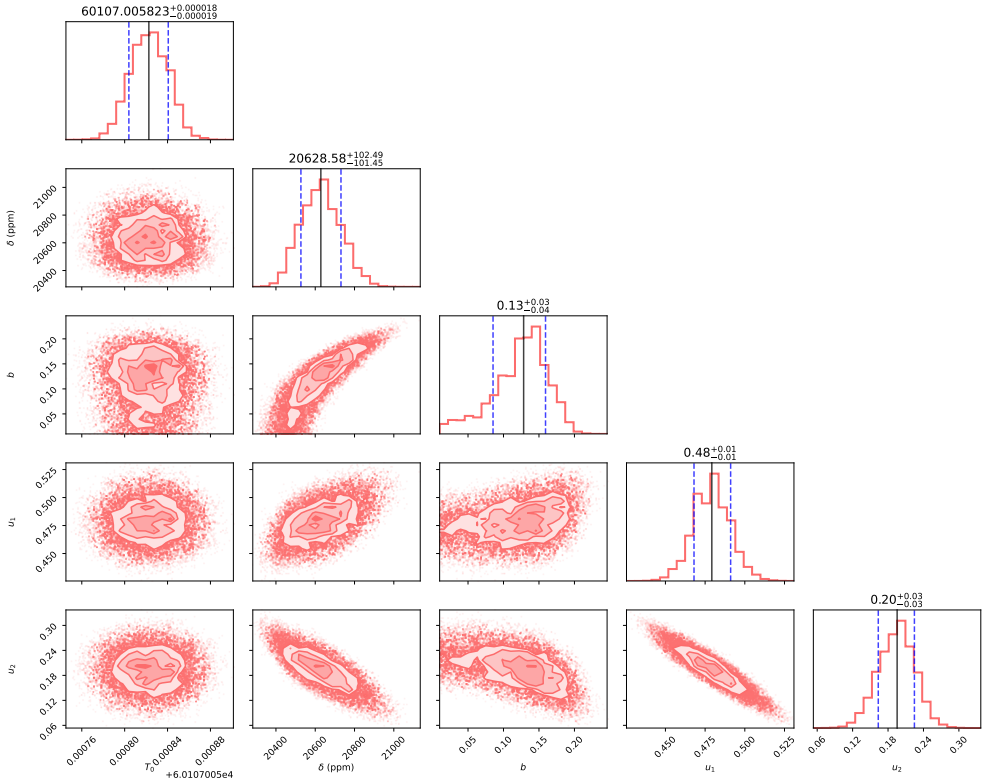
Supplementary Figure 5: Free retrieval comparison of the *JWST*/NIRISS-SOSS transmission spectrum from the `exoTEDRF` and `NAMELESS` pipelines (shown with $1\,\sigma$ error bars). The top panel shows the `pRT` fits for both datasets. Note that both datasets were run with the best-fit offset from `SCARLET`. The `NAMELESS` pipeline produces a shallower lower-wavelength slope than `exoTEDRF`, but the extracted thermal profiles (middle column) and volumetric abundances (right column) are consistent with each other. The helium light curves (pixels corresponding to λ: 1.0830216 – 1.0839538 μm) from both pipelines (left panels) are also consistent.



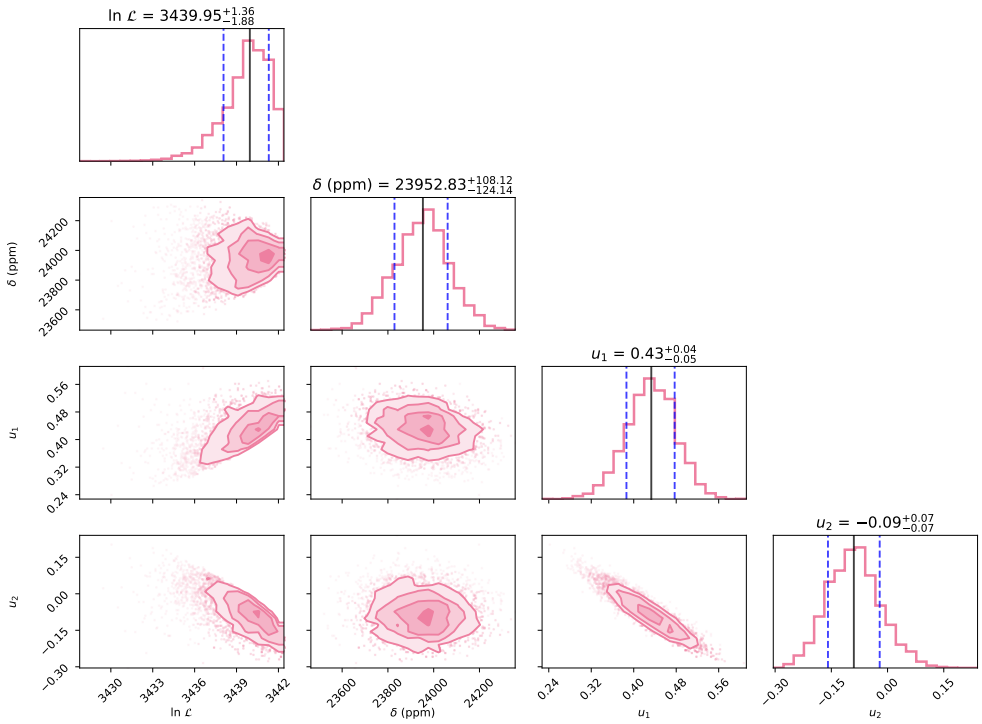
Supplementary Figure 6: Corner plots for NIRISS-SOSS Order 1 broadband fits using `juliet` without any detrending performed. The solid black line shows the 0.5 quantile with the upper and lower $1\,\sigma$ errors shown in red.



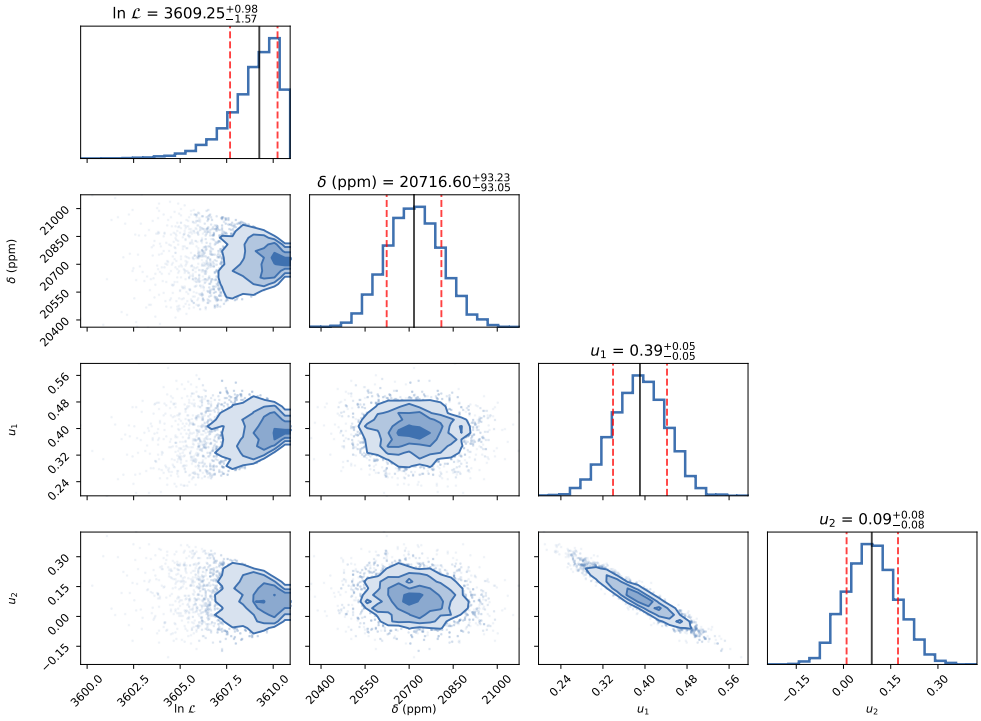
Supplementary Figure 7: Corner plots for NIRISS-SOSS Order 2 broadband fits using `juliet` with linear detrending performed on the `exoTEDRF` dataset. The solid black line shows the 0.5 quantile with the upper and lower 1 σ errors shown in blue.



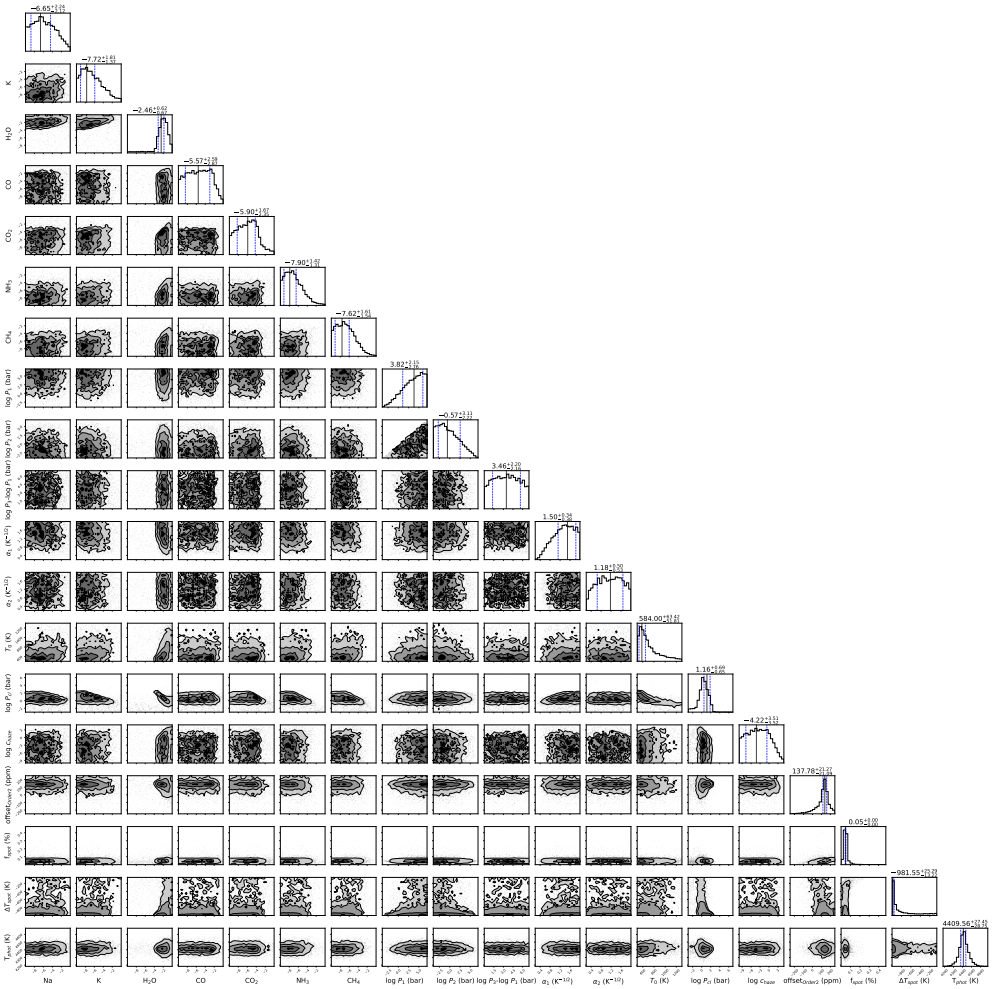
Supplementary Figure 8: Corner plots for NIRISS-SOSS Order 2 broadband fits using `juliet` without linear detrending performed on the `exoTEDRF` dataset. The solid black line shows the 0.5 quantile with the upper and lower $1\,\sigma$ errors shown in blue. The BIC indicates that this model is preferred over linear detrending model.



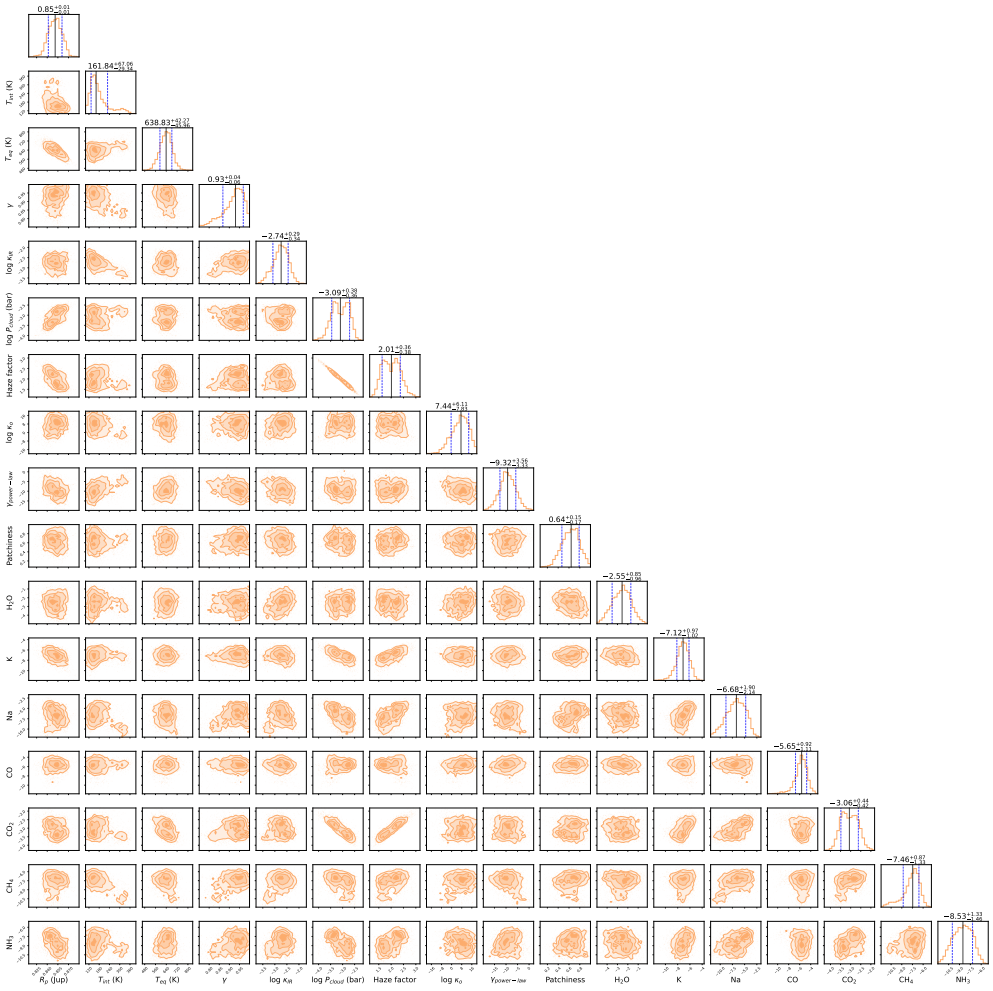
Supplementary Figure 9: Corner plots for NIRISS-SOSS pixel-resolution fit for λ: 1.0835 μm using `juliet` on the `exoTEDRF` dataset. The solid black line shows the 0.5 quantile with the upper and lower 1 σ errors shown in blue.



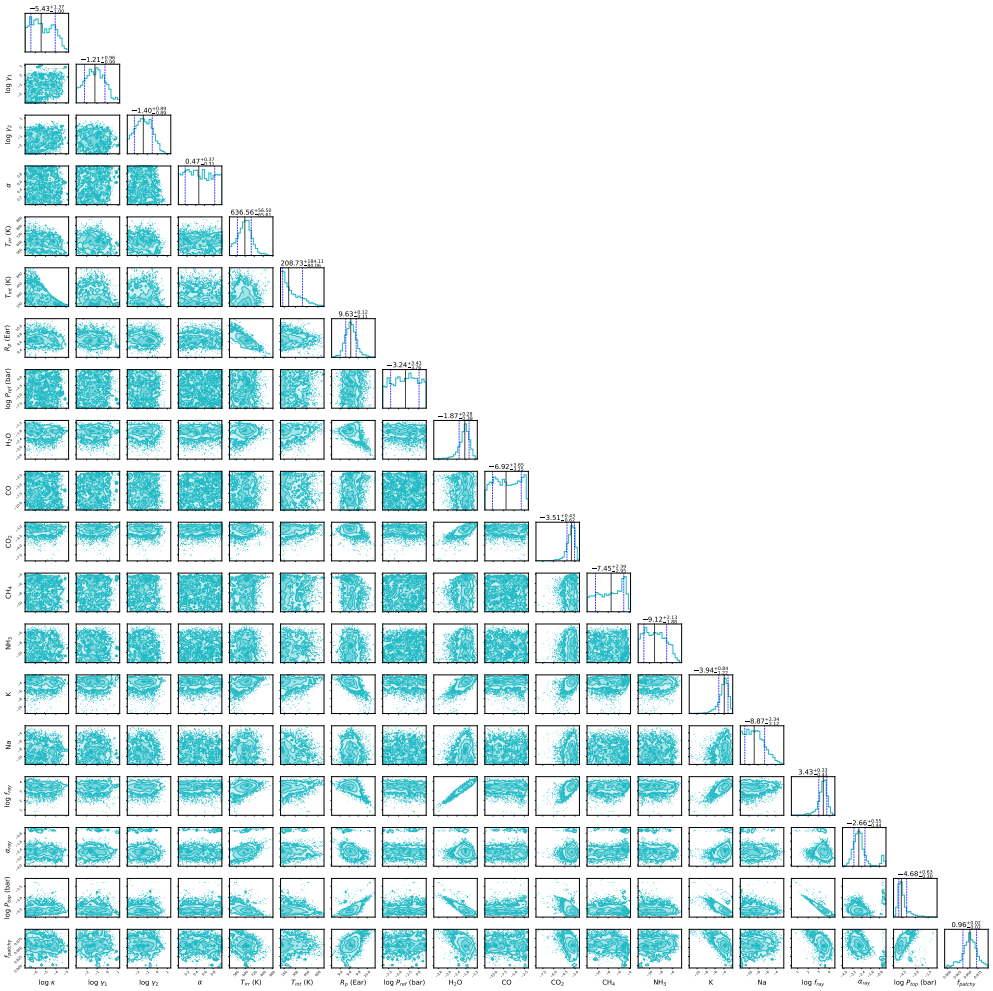
Supplementary Figure 10: Corner plots for NIRISS-SOSS pixel-resolution fit for region outside the helium line λ: 1.0861 μm using `juliet` on the `exoTEDRF` dataset. The solid black line shows the 0.5 quantile with the upper and lower 1 σ errors shown in blue.



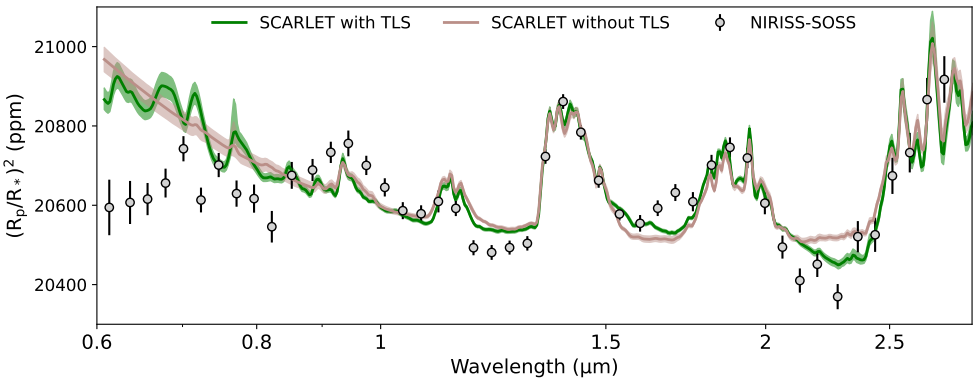
Supplementary Figure 11: Corner plots from the SCARLET free-chemistry retrievals of the NIRISS-SOSS dataset at R∼100. Median values with $1\,\sigma$ uncertainties are highlighted. Volume mixing ratios of the chemical species are shown in $\log_{10}$ scale.



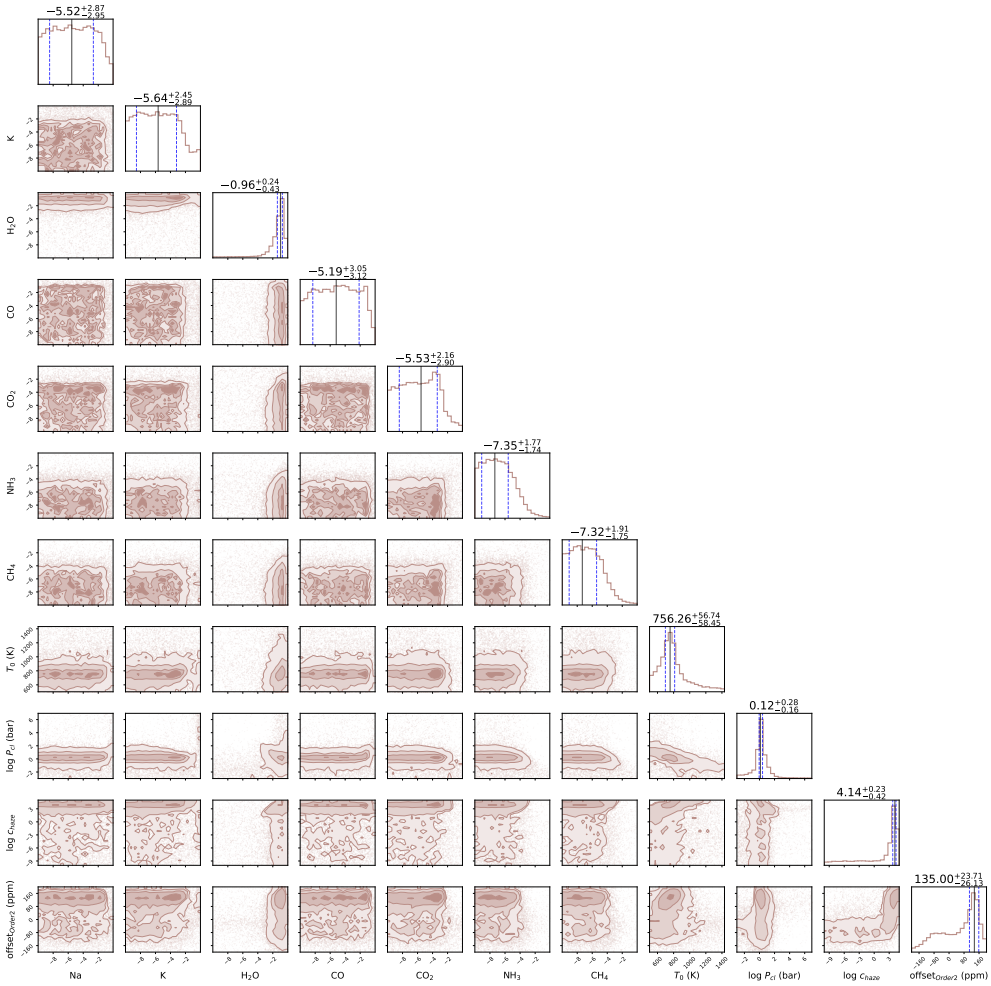
Supplementary Figure 12: Same as Supplementary Figure 11, but with `petitRADTRANS`.



Supplementary Figure 13: Same as Supplementary Figure 11, but with `Pyrat Bay`.



Supplementary Figure 14: Comparison of `SCARLET` models with (green) and without spots (brown). The model including spots and hazes is strongly preferred, with a $\Delta \ln Z = 11.62$ over the pure atmospheric hazes model. This corresponds to a $5.2\,\sigma$ detection of spots.



Supplementary Figure 15: Same as Supplementary Figure 11, but without spots and just hazes from `SCARLET`.

Molecule	Reference
H_2O	[12]
CO	[13]
CO_2	[14]
CH_4	[15]
HCN	[16]
H_2S	[17]
SO_2	[18]
NH_3	[19]

Supplementary Table 1: List of the references for the cross-sections used to compute the opacities of the molecules included in `SCARLET` retrievals.

Parameter	Posterior	Prior
Chemical Species		
$\log_{10}(X_{\mathrm{H_2O}})$	$-2.45^{+0.64}_{-0.67}$	$\mathcal{U}(-10.0, 0)$
$\log_{10}(X_{\mathrm{K}})$	< -4.89	$\mathcal{U}(-10.0, 0)$
$\log_{10}(X_{\mathrm{NH_3}})$	< -5.76	$\mathcal{U}(-10.0, 0)$
$\log_{10}(X_{\mathrm{CH_4}})$	< -5.14	$\mathcal{U}(-10.0, 0)$
$\log_{10}(X_{\mathrm{CO_2}})$	$-5.92^{+1.69}_{-2.36}$	$\mathcal{U}(-10.0, 0)$
$\log_{10}(X_{\mathrm{CO}})$	$-5.54^{+2.56}_{-2.92}$	$\mathcal{U}(-10.0, 0)$
$\log_{10}(X_{\mathrm{Na}})$	< -3.13	$\mathcal{U}(-10.0, 0)$
Temperature profile		
T_0 (K)	$584.86^{+63.39}_{-51.97}$	$\mathcal{U}(500, 1300)$
α_1 ($\mathrm{K}^{-1/2}$)	$1.49^{+0.33}_{-0.37}$	$\mathcal{U}(0.2, 2.0)$
α_2 ($\mathrm{K}^{-1/2}$)	$1.16^{+0.51}_{-0.51}$	$\mathcal{U}(0.2, 2.0)$
$\log_{10}(P_1)$ (bar)	$-1.18^{+2.16}_{-2.79}$	$\mathcal{U}(-9.0, 2.0)$
$\log_{10}(P_2)$ (bar)	$-5.53^{+3.06}_{-2.24}$	$\mathcal{U}(-9.0, 2.0)$
$\log_{10}(P_3\text{-}P_1)$ (bar)	$3.44^{+2.2}_{-2.17}$	$\mathcal{U}(0, 7.0)$
Aerosols		
$\log_{10}(c_{\mathrm{haze}})$	$-4.12^{+3.41}_{-3.56}$	$\mathcal{U}(-10, 5)$
$\log_{10}(P_{\mathrm{cloud}})$ (bar)	$-3.84^{+0.7}_{-0.65}$	$\mathcal{U}(-8.0, 2.0)$
Stellar contamination		
f_{spot} (%)	$4.74^{+0.35}_{-0.36}$	$\mathcal{U}(0, 50)$
T_{phot} (K)	$4409.63^{+26.79}_{-28.04}$	$\mathcal{N}(4425, 70)$
ΔT_{spot} (K)	$-981.58^{+25.48}_{-12.93}$	$\mathcal{U}(-1000, -50)$
Instrument offset		
Δ_{O2} (ppm)	$137.67^{+21.0}_{-21.3}$	$\mathcal{U}(-250, 250)$

Supplementary Table 2: Constraints on the atmosphere and stellar heterogeneity properties from the `SCARLET` free retrieval, along with the prior adopted on each parameter (uniform priors except on the photosphere temperature T_{phot}). We report the median and $\pm 1\sigma$ range for all parameters, or 2σ upper limits where applicable.

Parameter	Best Fit	Prior
Physical Properties		
R_p (R_{Jup})	$0.85^{+0.01}_{-0.01}$	$\mathcal{U}$[0.8, 1.1]
Thermal and pRT parameters		
T_{irr} (K)	$638.82^{+42.27}_{-45.96}$	$\mathcal{U}$[400, 900.0]
T_{int} (K)	< 314.89	$\mathcal{U}$[100.0, 600.0]
$\log_{10}(\kappa_{IR})$	$-2.74^{+0.29}_{-0.34}$	$\mathcal{U}$[-4.0, 2.0]
$\log_{10}(\gamma_{Guillot})$	$0.93^{+0.04}_{-0.06}$	$\mathcal{U}$[0.5, 1.0]
Chemical species		
$\log_{10}(X_{H_2O})$	$-2.55^{+0.85}_{-0.96}$	$\mathcal{U}$[-12.0, -0.1]
$\log_{10}(X_{CO})$	$-5.65^{+0.92}_{-1.11}$	$\mathcal{U}$[-12.0, -0.1]
$\log_{10}(X_{CO_2})$	$-3.06^{+0.44}_{-0.42}$	$\mathcal{U}$[-12.0, -0.1]
$\log_{10}(X_{CH_4})$	< -6.13	$\mathcal{U}$[-12.0, -0.1]
$\log_{10}(X_{K})$	< -5.57	$\mathcal{U}$[-12.0, -0.1]
$\log_{10}(X_{Na})$	< -3.94	$\mathcal{U}$[-12.0, -0.1]
$\log_{10}(X_{NH_3})$	< -6.55	$\mathcal{U}$[-12.0, -0.1]
Clouds and Hazes		
$\log_{10}P_{cloud}$ (bar)	$-3.09^{+0.38}_{-0.36}$	$\mathcal{U}$[-6.0,2.0]
Haze factor	$2.01^{+0.36}_{-0.38}$	$\mathcal{U}$[-4.0,10.0]
$\log_{10}(\kappa_o)$	$7.44^{+6.11}_{-7.83}$	$\mathcal{U}$[0.01, 20.0]
$\gamma_{power-law}$	$-9.32^{+3.56}_{-3.33}$	$\mathcal{U}$[-20, 2.0]
Patchiness	$0.64^{+0.15}_{-0.17}$	$\mathcal{U}$[0, 1.0]

Supplementary Table 3: Atmospheric constrains from the `petitRADTRANS` free retrieval, along with the prior adopted on each parameter. We report the median and $\pm 1\sigma$ range for all parameters, or 2σ upper limits where applicable.

Parameter	Median Fit	Prior
Physical Properties		
R_p ($R_\oplus$)	$9.63^{+0.12}_{-0.11}$	$\mathcal{U}$[8.4, 12.7]
Thermal and `Pyrat Bay` parameters		
T_{irr} (K)	$636.6^{+56.9}_{-65.9}$	$\mathcal{U}$[500.0, 1600.0]
T_{int} (K)	< 562.07	$\mathcal{U}$[100.0, 800.0]
log(κ')	$-5.43^{+1.37}_{-1.01}$	$\mathcal{U}$[-7.0, 1.0]
log(γ_1)	$-1.21^{+0.97}_{-0.99}$	$\mathcal{U}$[-3.0, 3.0]
log(γ_2)	$-1.40^{+0.90}_{-0.89}$	$\mathcal{U}$[-3.0, 3.0]
α	$0.469^{+0.368}_{-0.313}$	$\mathcal{U}$[0.0, 1.0]
log(P_{ref}) (bar)	$-3.236^{+3.433}_{-3.776}$	$\mathcal{U}$[-9.0, 2.0]
Chemical species		
$\log_{10}(X_{H_2O})$	$-1.868^{+0.285}_{-0.377}$	$\mathcal{U}$[-12.0, -0.05]
$\log_{10}(X_{CO})$	< -2.1	$\mathcal{U}$[-12.0, -0.05]
$\log_{10}(X_{CO_2})$	$-3.51^{+0.44}_{-0.62}$	$\mathcal{U}$[-12.0, -0.05]
$\log_{10}(X_{CH_4})$	< -4.36	$\mathcal{U}$[-12.0, -0.05]
$\log_{10}(X_{K})$	$-3.94^{+0.84}_{-1.23}$	$\mathcal{U}$[-12.0, -0.05]
$\log_{10}(X_{Na})$	< -4.20	$\mathcal{U}$[-12.0, -0.05]
$\log_{10}(X_{NH_3})$	< -5.63	$\mathcal{U}$[-12.0, -0.05]
Clouds and Hazes		
$\log_{10}(f_{ray})$	$3.43^{+0.33}_{-0.43}$	$\mathcal{U}$[0.0, 10.0]
α_{ray}	$-2.66^{+0.55}_{-0.44}$	$\mathcal{U}$[-6.0, 10.0]
$\log_{10}(P_{top})$ (bar)	$-4.68^{+0.64}_{-0.31}$	$\mathcal{U}$[-9.0, 2.0]
f_{patchy}	$0.955^{+0.016}_{-0.017}$	$\mathcal{U}$[0.0, 1.0]

Supplementary Table 4: Atmospheric constrains from the `Pyrat Bay` free retrieval, along with the prior adopted on each parameter. We report the best-fit and median with $\pm1\sigma$ range for all parameters, or 2σ upper limits where applicable..

References

[1] Coulombe, L.-P. *et al.* A broadband thermal emission spectrum of the ultra-hot Jupiter WASP-18b. *Nature* **620**, 292–298 (2023).

[2] Feinstein, A. D. *et al.* Early Release Science of the exoplanet WASP-39b with JWST NIRISS. *Nature* **614**, 670–675 (2023).

[3] Coulombe, L.-P. *et al.* Highly reflective white clouds on the western dayside of an exo-Neptune. *arXiv e-prints* arXiv:2501.14016 (2025).

[4] Benneke, B. *et al.* JWST Reveals CH_4, CO_2, and H_2O in a Metal-rich Miscible Atmosphere on a Two-Earth-Radius Exoplanet. *arXiv e-prints* arXiv:2403.03325 (2024).

[5] Bushouse, H. *et al.* Jwst calibration pipeline (2023). URL https://doi.org/10.5281/zenodo.8404029. If you use this software in your work, please cite it using the following metadata.

[6] Benneke, B. *et al.* A sub-Neptune exoplanet with a low-metallicity methane-depleted atmosphere and Mie-scattering clouds. *Nature Astronomy* **3**, 813–821 (2019). URL http://adsabs.harvard.edu/abs/2019NatAs...3..813B.

[7] Benneke, B. *et al.* Water Vapor and Clouds on the Habitable-zone Sub-Neptune Exoplanet K2-18b. *The Astrophysical Journal Letters* **887**, L14 (2019). URL https://dx.doi.org/10.3847/2041-8213/ab59dc. Publisher: The American Astronomical Society.

[8] Coulombe, L.-P., Roy, P.-A. & Benneke, B. Biases in Exoplanet Transmission Spectra Introduced by Limb-darkening Parametrization. *Astron. J.* **168**, 227 (2024).

[9] Kreidberg, L. batman: BAsic Transit Model cAlculatioN in Python. *Publ. Astron. Soc. Pac.* **127**, 1161 (2015).

[10] Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. ¡tt¿emcee¡/tt¿: The mcmc hammer. *Publications of the Astronomical Society of the Pacific* **125**, 306–312 (2013). URL http://dx.doi.org/10.1086/670067.

[11] Allart, R. *et al.* High-resolution confirmation of an extended helium atmosphere around WASP-107b. *Astron. Astrophys.* **623**, A58 (2019).

[12] Polyansky, O. L. *et al.* ExoMol molecular line lists XXX: a complete high-accuracy line list for water. *Mon. Not. R. Astron. Soc.* **480**, 2597–2608 (2018).

[13] Hargreaves, R. J. *et al.* Spectroscopic line parameters of NO, NO_2, and N_2O for the HITEMP database. *J. Quant. Spectrosc. Radiat. Transf.* **232**, 35–53 (2019).

[14] Yurchenko, S. N., Mellor, T. M., Freedman, R. S. & Tennyson, J. ExoMol line lists - XXXIX. Ro-vibrational molecular line list for CO_2. *Mon. Not. R. Astron. Soc.* **496**, 5282–5291 (2020).

[15] Hargreaves, R. J. *et al.* An Accurate, Extensive, and Practical Line List of Methane for the HITEMP Database. *Astrophys. J. Suppl. Ser.* **247**, 55 (2020).

[16] Harris, G. J., Tennyson, J., Kaminsky, B. M., Pavlenko, Y. V. & Jones, H. R. A. Improved HCN/HNC linelist, model atmospheres and synthetic spectra for WZ Cas. *Mon. Not. R. Astron. Soc.* **367**, 400–406 (2006).

[17] Azzam, A. A. A., Tennyson, J., Yurchenko, S. N. & Naumenko, O. V. ExoMol molecular line lists - XVI. The rotation-vibration spectrum of hot H_2S. *Mon. Not. R. Astron. Soc.* **460**, 4063–4074 (2016).

[18] Underwood, D. S. *et al.* ExoMol molecular line lists - XIV. The rotation-vibration spectrum of hot SO_2. *Mon. Not. R. Astron. Soc.* **459**, 3890–3899 (2016).

[19] Coles, P. A., Yurchenko, S. N. & Tennyson, J. ExoMol molecular line lists - XXXV. A rotation-vibration line list for hot ammonia. *Mon. Not. R. Astron. Soc.* **490**, 4638–4647 (2019).