

Parameter degeneracies associated with interpreting *HST* WFC3 transmission spectra of exoplanetary atmospheres

Aline Novais¹★ Chloe Fisher², Luan Ghezzi³, Daniel Kitzmann^{4,5}, Brian Thorsbro^{1,6}, and Kevin Heng^{7,8,9,10}

¹Division of Astrophysics, Department of Physics, Lund University, Box 118, 221 09, Lund, Sweden

²University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, United Kingdom

³Valongo Observatory, Federal University of Rio de Janeiro, Ladeira do Pedro Antonio, 43, 20080-090, Rio de Janeiro, Brazil

⁴Space Research and Planetary Sciences, Physics Institute, University of Bern, Gesellschaftsstrasse 6, CH-3012 Bern, Switzerland

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

⁶Observatoire de la Côte d'Azur, CNRS UMR 7293, BP4229, Laboratoire Lagrange, F-06304 Nice Cedex 4, France

⁷University Observatory Munich, Ludwig Maximilian University, Scheinerstrasse 1, D-81679 Munich, Germany

⁸ARTORG Center for Biomedical Engineering Research, University of Bern, Murtenstrasse 50, CH-3008, Bern, Switzerland

⁹University College London, Department of Physics & Astronomy, Gower St, London, WC1E 6BT, United Kingdom

¹⁰Astronomy & Astrophysics Group, Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom

Accepted 2025 March 5. Received 2025 March 4; in original form 2023 October 25

ABSTRACT

The Wide Field Camera 3 (WFC3) instrument on the *Hubble Space Telescope* has provided an abundance of exoplanet spectra over the years. These spectra have enabled analysis studies using atmospheric retrievals to constrain the properties of these objects. However, follow-up observations from the *James Webb Space Telescope* have called into question some of the results from these older datasets, and highlighted the need to properly understand the degeneracies associated with retrievals of WFC3 spectra. In this study, we perform atmospheric retrievals of 38 transmission spectra from WFC3 and use model comparison to determine the complexity required to fit the data. We explore the effect of retrieving system parameters such as the stellar radius and planet's surface gravity, and thoroughly investigate the degeneracies between individual model parameters – specifically the temperature, abundance of water, and cloud-top level. We focus on three case studies (HD 209458b, WASP-12b, and WASP-39b) in an attempt to diagnose some of the issues with these retrievals, in particular the low retrieved temperatures when compared to the equilibrium values. Our study advocates for the careful consideration of parameter degeneracies when interpreting retrieval results, as well as the importance of wider wavelength coverage to break these degeneracies, in agreement with previous studies. The combination of data from multiple instruments, as well as analysis from multiple data reductions and retrieval codes, will allow us to robustly characterise the atmosphere of these exoplanets.

Key words: planets and satellites: atmospheres, exoplanets

1 INTRODUCTION

Transmission spectra measure wavelength-dependent transit radii of exoplanets, which in principle encode information containing the cross sections and relative abundances of chemical species (Seager & Sasselov 2000; Brown 2001). Interpreting these spectra to extract chemical abundances is a process known as atmospheric retrieval (first introduced to the study of exoplanetary atmospheres by Madhusudhan & Seager 2009), which involves using modern Bayesian techniques to implement Occam's Razor (Benneke & Seager 2012). For a recent review of atmospheric retrieval, see Barstow & Heng (2020).

Several retrieval works in the past 30 years have attempted to characterise space-observed atmospheres using transmission data from *Hubble Space Telescope* (*HST*) and *Spitzer Space Telescope* (e.g. Kreidberg et al. 2014; Sing et al. 2016; Barstow et al. 2017; Tsiaras et al. 2018; Wakeford et al. 2018; Pinhas et al. 2019; Welbanks &

Madhusudhan 2019, amongst many others). Physical approximations such as isothermal and isobaric transit chords, and grey clouds (or even no clouds) were often used to estimate atmospheric properties. However, the low resolution and signal-to-noise ratio, together with the short wavelength range covered by *HST* and *Spitzer* data, have proven to be strong limitations against properly modelling the structure and composition of exoplanet atmospheres. As a result, retrieved atmospheric parameters have been shown to be strongly dependent on the wavelength of data and the choice of model. For instance, the Wide Camera Field 3 (WFC3) spectral range mainly covers H₂O features, failing to provide confident detections of other expected molecules, such as carbon-bearing species in particular. Therefore, retrievals can result in very high H₂O abundances, due to the lack of continuum opacities (Barstow & Heng 2020). Another example is the cloud treatment, as a model with clouds will “hide” part of the atmosphere, allowing the chemical abundances to be higher than in the cloud-free case (Seager & Sasselov 2000). For these reasons, retrievals from low-resolution transmission data often show degeneracies (i.e. ambiguity) between parameters (Benneke &

★ E-mail: aline.novais@fysik.lu.se

Seager 2012; Griffith 2014; Heng & Kitzmann 2017; Fisher & Heng 2018; Welbanks & Madhusudhan 2019).

Additionally, a constant issue in the exoplanet community is the strong disagreement between outcomes for the same planet, not only as a result of the degeneracies between parameters, but also due to different model implementations used by individual studies, as highlighted by Barstow & Heng (2020). Furthermore, it can be difficult to identify degeneracies, in particular when they occur between multiple parameters, which requires careful consideration of their posterior distributions.

Since the first studies on the equations governing the transmission spectra of exoplanets, degeneracies between atmospheric parameters have been considered. Brown (2001) noted the similar effects from lower clouds, higher temperatures, increased abundances and turbulent velocities, which all increase the strength of the spectral features. Lecavelier Des Etangs et al. (2008) elucidated a degeneracy between molecular abundances and total atmospheric pressure (i.e. a “normalization degeneracy”). This same degeneracy was explored in Benneke & Seager (2012) in the context of super-Earths, where they determine that the degeneracy can be broken by the inclusion of optical data covering the molecular Rayleigh scattering signature, assuming that N₂ and H₂+He are the only spectrally inactive species. Further work from Benneke & Seager (2013) showed how a degeneracy between a cloudy hydrogen-dominated atmosphere and a cloud-free high mean molecular weight atmosphere can be broken by precisely measuring the wings and depths of absorption features, given sufficiently small uncertainties on the data points. Line & Parmentier (2016) showed how this degeneracy means that, in wavelength regions with a single absorber (such as WFC3), one is only able to constrain very high or low molecular abundances, and are not sensitive to changes between these extremes.

Additional studies have also considered the planet’s bulk properties, such as mass and radius. de Wit & Seager (2013) investigated how the planet mass could be retrieved from transmission spectra, although they noted a degeneracy with high-altitude clouds. This is further explored in Batalha et al. (2017), who found this degeneracy to be detrimental when attempting to retrieve the masses of super-Earths and sub-Neptunes, in particular. In addition, Griffith (2014) found that small uncertainties in the planet’s radius can cause uncertainties in the gas mixing ratios of several orders of magnitude – another variation of the normalization degeneracy. In agreement with previous studies, they identify how the inclusion of optical data can mitigate this issue. Bétrémieux (2016) also investigate how refraction signatures can be used to break this degeneracy, and further studies explore the effects of accounting for a surface or optically thick clouds (Bétrémieux & Swain 2017, 2018). More recently, Welbanks & Madhusudhan (2019) performed an in-depth study of retrieval degeneracies and found that the combination of optical and infrared spectra can break the degeneracy between clouds and chemical abundances. They also highlighted the limitations of the isobaric assumption in analytical models (e.g. Heng & Kitzmann 2017), and discussed the importance of accurately accounting for collision-induced absorption (CIA) in correctly determining molecular abundances. The impact of CIA on transmission spectra was also explored in some of the previous studies (de Wit & Seager 2013; Line & Parmentier 2016; Bétrémieux & Swain 2017, 2018). In Welbanks & Madhusudhan (2019), they performed a case study of optical and infrared data of HD 209458b by implementing retrievals with increasingly complex physical models, and analysed the resulting degeneracies. However, for low-resolution data with limited wavelength coverage, higher model complexity is not always warranted, and can lead to overfitting.

1.1 *HST* in the *JWST* era

With the launch of the *James Webb Space Telescope (JWST)*, we are seeing a drastic improvement in the wavelength coverage and precision of exoplanet spectra over *HST*. *JWST* spectra of previously observed exoplanets are bringing to light the limitations of *HST*, with an array of new molecular detections and a deeper understanding of their atmospheric properties (e.g. Ahrer et al. 2023; Feinstein et al. 2023; Rustamkulov et al. 2023; Alderson et al. 2023; Tsai et al. 2023; Hammond et al. 2024). Despite the number of studies investigating degeneracies in transmission spectra retrievals, there are several cases where *JWST* measurements are demonstrating that our conclusions from *HST* data were incorrect, due to degenerate solutions within the *HST* wavelength range (Madhusudhan et al. 2023). In addition, Nixon et al. (2024) showed that uncertainties in the retrieved parameters from *HST* data have likely been underestimated, due to uncertainties on the correct model to use.

With *JWST* instruments such as NIRISS encompassing the entire WFC3 wavelength range at a higher resolution, and expanding further into the near infrared, future observations are likely to continue to challenge our interpretations from previous WFC3 exoplanet spectra. Lueber et al. (2024) using WASP-39b data from different *JWST* instruments (NIRCam, NIRISS, NIRSpec G395H, and NIR-Spec PRISM) from 0.5 to 5.5 μm have shown that both the increase in data resolution as well as the wide wavelength coverage were able to break the degeneracy between H₂O abundances and normalization parameters. Moreover, a comparison study between *HST* WFC3 and *JWST* NIRISS data for the same object by Fisher et al. (2024) has further demonstrated that WFC3 retrievals alone are not able to accurately constrain H₂O abundances, due to the difficulty to determine the continuum in such a narrow wavelength range.

However, even in the era of *JWST*, there is still a strong argument for using *HST* data, especially in the optical range, which is not covered by *JWST*. A recent study by Fairman et al. (2024) demonstrated the importance of optical data from STIS for constraining cloud properties in exoplanet atmospheres, and Fisher et al. (2024) explored the benefits of combining this with data from *JWST*’s NIRISS instrument. Furthermore, despite the anticipated removal of WFC3 from Hubble proposal cycles as of October 2025, there is an abundance of available WFC3 exoplanet spectra. This archival data still has much to offer for population studies, and understanding the challenges associated with its interpretations – both individually and collectively – provides a powerful lesson for our analysis of *JWST* spectra.

1.2 Structure of the study

The present study is a follow-up from Fisher & Heng (2018), analysing the same sample of 38 transmission spectra observed by *HST* WFC3, focusing on degeneracies and challenges associated with determining atmospheric parameters from WFC3 transmission spectra. Section 2 explains the retrieval code used in this work, the chosen atmospheric models, as well as the method of comparison between them. In Section 3 we investigate the effect of accounting for uncertainties in the stellar radius and planetary surface gravity, which are commonly considered fixed parameters. We analyse the individual influence of model properties (e.g. temperature, pressure, chemical abundance, cloud parameterization) and degeneracies between them in retrieval outcomes from WFC3 data. In Section 4 we compare our current H₂O abundance and temperatures with Fisher & Heng (2018) outcomes. In particular, we examine the temperature and its influence in chemical abundances, which are seen to be artificially low in retrievals from transmission spectra due to day-night asymmetries

Table 1. Prior ranges assumed in our retrievals. For R_{star} and $\log g_p$, μ and σ represent the mean and standard deviation of the Gaussian distribution, respectively, assumed to be the mean and uncertainty values listed in Table A1.

Parameter	Symbol	Range	Distribution	Units
Pressure	P	$[10^{-9}, 10^1]$	Log-uniform	bar
Temperature	T	$[200, 3100]$	Uniform	K
Water abundance	$X_{\text{H}_2\text{O}}$	$[10^{-12}, 10^{-1}]$	Log-uniform	–
Cloud-top pressure	$P_{\text{cloud-top}}$	$[10^{-4}, 10]$	Log-uniform	bar
Cloud composition parameter	Q_0	$[1, 100]$	Uniform	–
Cloud slope index	a_0	$[2, 13]$	Uniform	–
Spherical cloud particle radius	r_{cloud}	$[10^{-9}, 10^{-3}]$	Log-uniform	cm
Reference optical depth	τ_{ref}	$[10^{-5}, 10^2]$	Log-uniform	–
Cloud bottom pressure	$P_{\text{cloud-bottom}}$	$[10^0, 10^2]$	Log-uniform	bar
Reference transit radius	R_p	$[R_p - \sigma_{R_p}, R_p + \sigma_{R_p}]$	Uniform	R_{Jup}
Stellar radius	R_{star}	$\mu = R_{\text{star}}, \sigma = \sigma_{R_{\text{star}}}$	Gaussian	R_{\odot}
Planetary surface gravity	$\log g_p$	$\mu = \log g_p, \sigma = \sigma_{\log g_p}$	Gaussian	cm s^{-2}

Table 2. Free parameters retrieved by each model in our study.

Model	Number of parameters	T	$X_{\text{H}_2\text{O}}$	$P_{\text{cloud-top}}$	Q_0	a_0	r_{cloud}	τ_{ref}	$P_{\text{cloud-bottom}}$	R_p	R_{star}	$\log g_p$	line
cloud free	5	✓	✓	–	–	–	–	–	–	✓	✓	✓	–
grey clouds	6	✓	✓	✓	–	–	–	–	–	✓	✓	✓	–
non-grey clouds	11	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	–
flat line	1	–	–	–	–	–	–	–	–	–	–	–	✓

(MacDonald et al. 2020). In Section 5 we perform tests using three case studies (HD 209458b, WASP-12b, and WASP-39b) to further investigate the low temperature ‘‘problem’’. In Section 6 we compare our findings with previous literature results, highlighting how retrieved temperatures and chemical abundances are dependent on the choice of model. Finally, we perform retrievals using data reduced by two different works, in order to verify if model parameters are dependent on the reduction. Section 7 summarises the conclusions of our study.

2.1 HST WFC3 transmission spectra

To provide continuity with the study of Fisher & Heng (2018), we perform retrieval on the same set of 38 *HST* WFC3 transmission spectra. Spectra for GJ 1214b were obtained from Kreidberg et al. (2014), WASP-12b from Kreidberg et al. (2015), WASP-17b from Mandell et al. (2013), WASP-19b from Huitson et al. (2013), HD 97658b from Knutson et al. (2014), and TRAPPIST-1 objects from de Wit et al. (2018). The remaining WFC3 data were provided by Tsiaras et al. (2018).

2 METHODOLOGY

We perform atmospheric retrievals using the open-source, GPU-accelerated code BeAR (Bern Atmospheric Retrieval code¹, Kitzmann et al. 2020). BeAR uses the nested-sampling algorithm MultiNest (Skilling 2006; Feroz et al. 2009; Buchner et al. 2014; Buchner 2023) to fit exoplanet transmission or occultation spectra, as well as emission spectra of self-luminous objects. We refer the reader to Kitzmann et al. (2020), Lueber et al. (2022), and references therein for a full description of the BeAR framework, including temperature-pressure (T - P) profiles. Detailed descriptions of BeAR’s capabilities can also be found in its documentation.

2.2 Opacity sources

As most of our objects are presented in the literature as Jupiter-sized or Saturn-sized planets, our models assume atmospheres dominated by H₂ and He. Besides these background gases, H₂O is considered the main molecular opacity source, as the H₂O spectral feature at 1.4 μm is the strongest molecular signature in the *HST* WFC3 wavelength range. H₂ and He are initially set to solar abundances, which will vary according to the additional opacity sources included in the model (i.e. H₂O, CIA, Rayleigh scattering, and clouds).

For completeness, our work also contemplates four TRAPPIST-1 planets, for which we roughly define an Earth-like, N₂-dominated atmosphere. Therefore, we adopt the treatment of Fisher & Heng (2018), submitting the TRAPPIST-1 planets to two sets of retrievals, separately representing an H₂-dominated and an N₂-dominated atmosphere.

We use H₂O opacities from the DACE database (Grimm et al. 2021), which converts spectroscopic line lists from Polyansky et al. (2018) into opacities using the open-source opacity calculator HELIOS-K (Grimm & Heng 2015; Grimm et al. 2021). Our models also include

¹ BeAR is available at the GitHub repository <https://github.com/newstrangeworlds/bear>.

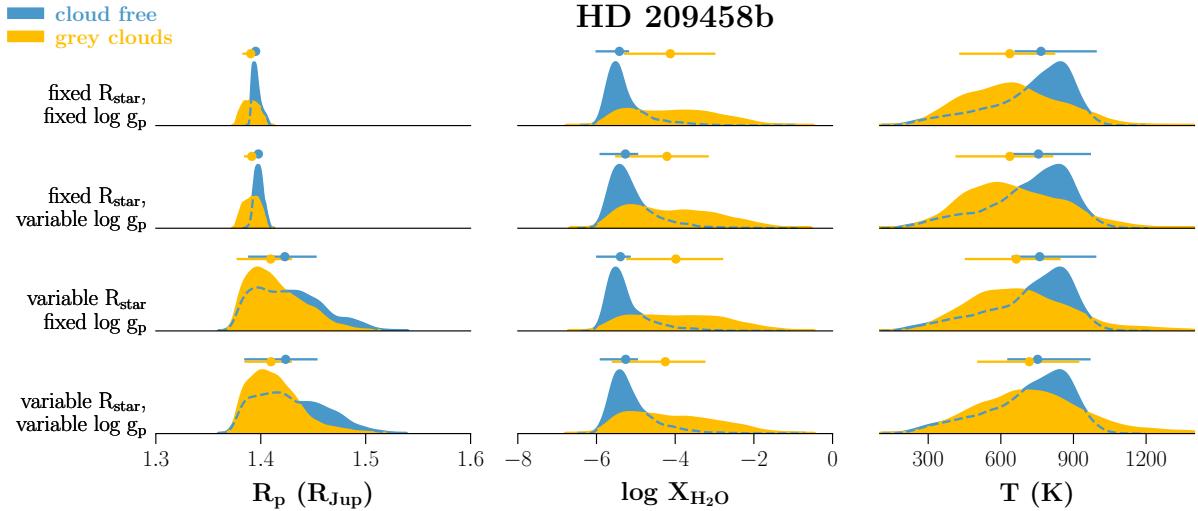


Figure 1. Posterior distributions of the reference planet radius R_p , water abundance $X_{\text{H}_2\text{O}}$, and temperature T retrieved using fixed or variable values of the stellar radius (R_\star) and planetary surface gravity ($\log g_p$) for HD 209458b. Cloud free (blue) and grey cloud (yellow) models were tested.

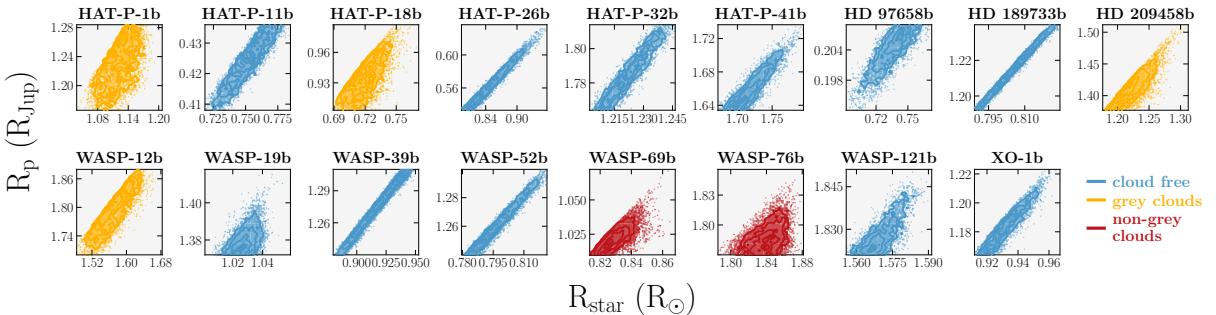


Figure 2. Correlation between retrieved stellar radii R_\star and reference transit radii R_p for all non-flat-line objects using their cloud free (blue), grey cloud (yellow), or non-grey cloud (red) best-fit model.

collision-induced absorption (CIA) opacities of $\text{H}_2\text{-H}_2$ from Abel et al. (2011) and $\text{H}_2\text{-He}$ from Abel et al. (2012), available within the HITRAN database (Karman et al. 2019), and Rayleigh scattering by H_2 (Cox 2000) and He (Sneep & Ubachs 2005; Thalman et al. 2014) in the total atmospheric opacity. We adopted an opacity resolution of 0.1 cm^{-1} for all runs.

Several of the planets in our sample do have detections of other species in their atmosphere. For example, the ultra-hot Jupiter WASP-12b shows evidence of VO from its STIS data (Evans et al. 2018), and H^- from the WFC3 thermal phase curve (Mikal-Evans et al. 2022). MacDonald & Madhusudhan (2017b) also found tentative evidence of NH_3 in the WFC3 transmission spectra of HD 209458b and WASP-31b, and HCN in WASP-63b (see also Kilpatrick et al. (2018)). However, to keep our analysis homogeneous across the sample and focus on degeneracies present with a single spectral feature, we choose to only include H_2O in our retrievals.

2.3 Retrieval code

The atmospheric retrieval code used by Fisher & Heng (2018) assumes isobaric transit chords, i.e. where opacities are assumed not to vary with pressure. Heng & Kitzmann (2017) showed that the isobaric approximation was sufficiently accurate in many cases for modelling WFC3 data, given its resolution and precision, but that inaccuracies could occur at lower temperatures in particular. In this work we adopt

a non-isobaric treatment using the retrieval code BeAR, to consider pressure-dependent transit chords. By assuming an isothermal atmosphere, some numerical approximations can be made to speed up the calculations. We consider an atmospheric pressure range from 10^{-9} to 10 bar, divided into 199 equal layers (200 levels) in log space. For comparison, isobaric retrievals by Fisher & Heng (2018) use a fixed pressure of 10^{-2} bar, assumed to be the WFC3 photospheric pressure.

Additionally, Fisher & Heng (2018) do not include Rayleigh scattering in their retrievals, which is included in this work. Further differences include the stellar radius R_\star and the planetary surface gravity $\log g_p$, which were previously considered fixed input parameters, and are now allowed to vary within a range of priors. The surface gravity is assumed to vary with pressure. Table A1 summarises the input values for these parameters for all objects analysed in this study. The uncertainty ranges of the reference transit radius R_p , R_\star , and $\log g_p$ from Table A1 are considered the prior ranges for each of these parameters, as listed in Table 1 along with their considered distributions.

2.4 Atmospheric models

Each planet is examined by a range of models, which consider different atmospheric parameters in the retrieval code. These parameters include the temperature, treated as constant, a reference transit ra-

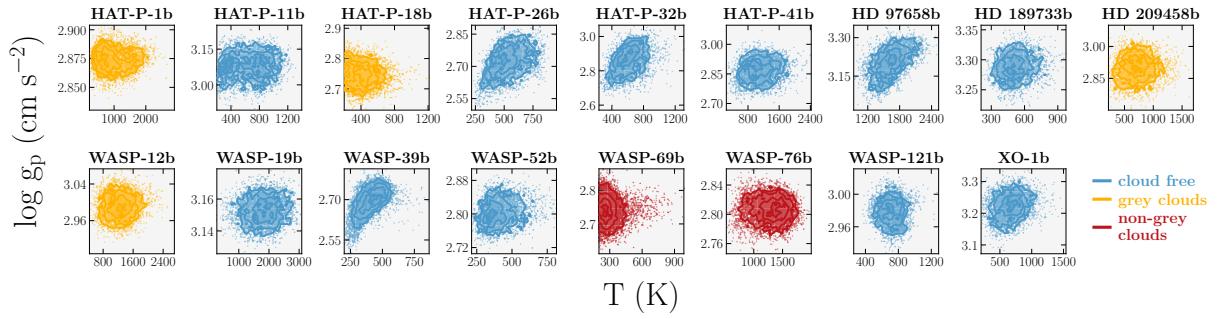


Figure 3. Same as Figure 2, but between temperatures T and planetary surface gravities $\log g_p$.

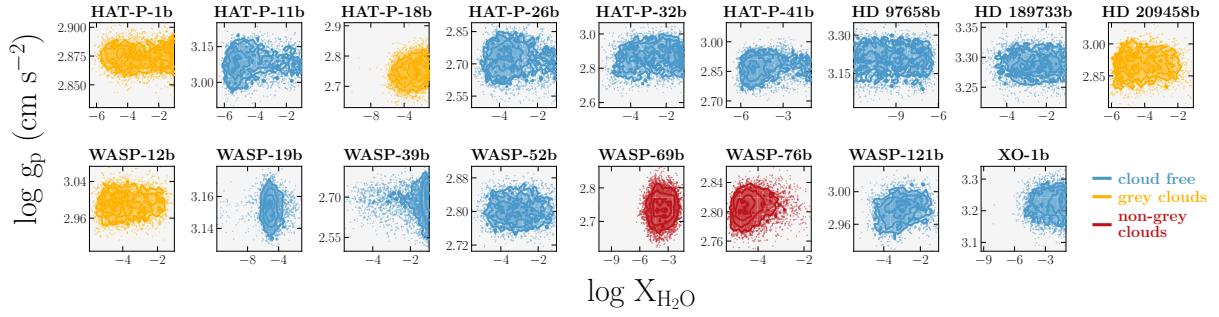


Figure 4. Same as Figure 2, but between H_2O abundances and planetary surface gravities $\log g_p$.

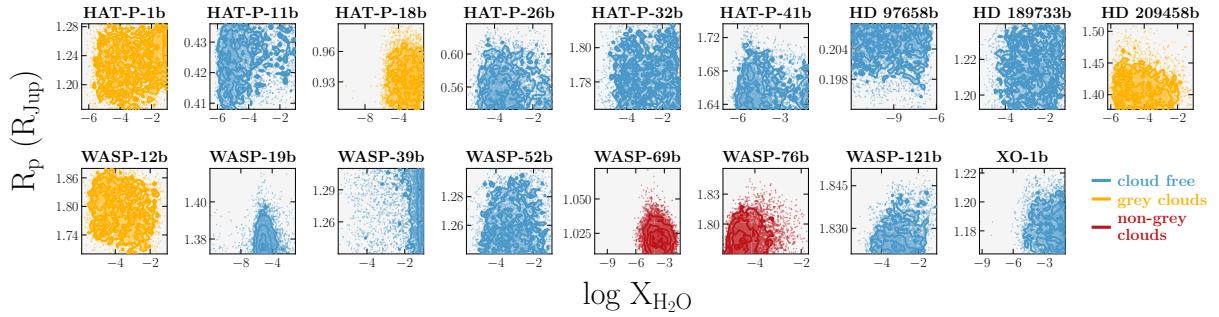


Figure 5. Same as Figure 2, but between H_2O abundances and reference transit radii R_p .

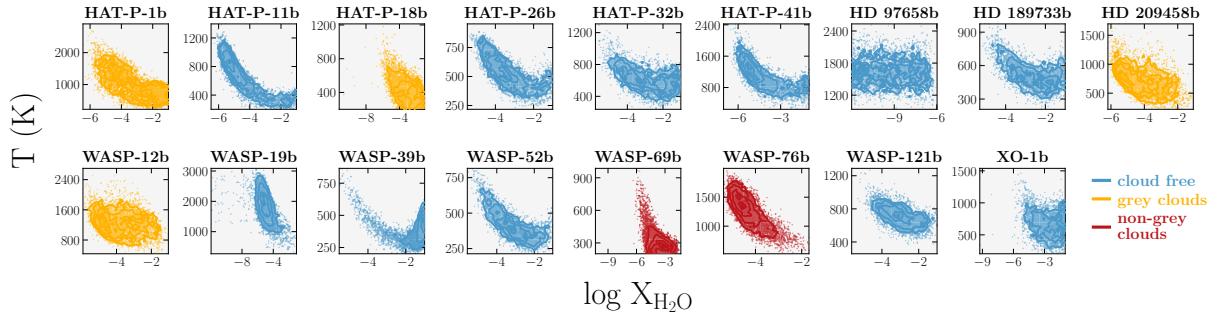


Figure 6. Same as Figure 2, but between H_2O abundances and temperatures T .

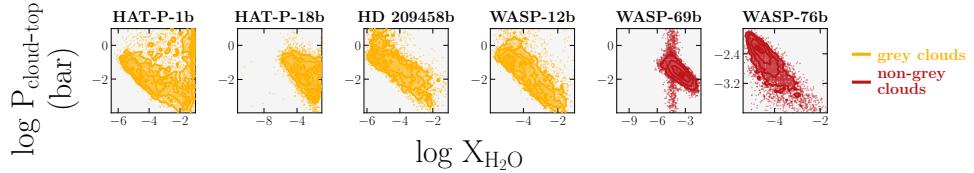


Figure 7. Same as Figure 2, but between H_2O abundances and cloud-top pressures $P_{\text{cloud-top}}$.

Table 3. Values used to test the spectral influence of each free parameter, as shown in Figure 8.

Planet	Value	T (K)	$\log g_p$ (cm s^{-2})	$\log X_{\text{H}_2\text{O}}$	R_p (R_{Jup})	R_{star} (R_{\odot})	$\log P_{\text{cloud-top}}$ (bar)
HD 209458b	low	700	2.68	-5.0	1.40	1.15	-6.0
	medium	1450	2.88	-3.0	1.45	1.20	-3.0
	high	2100	3.08	-1.0	1.50	1.25	0.0

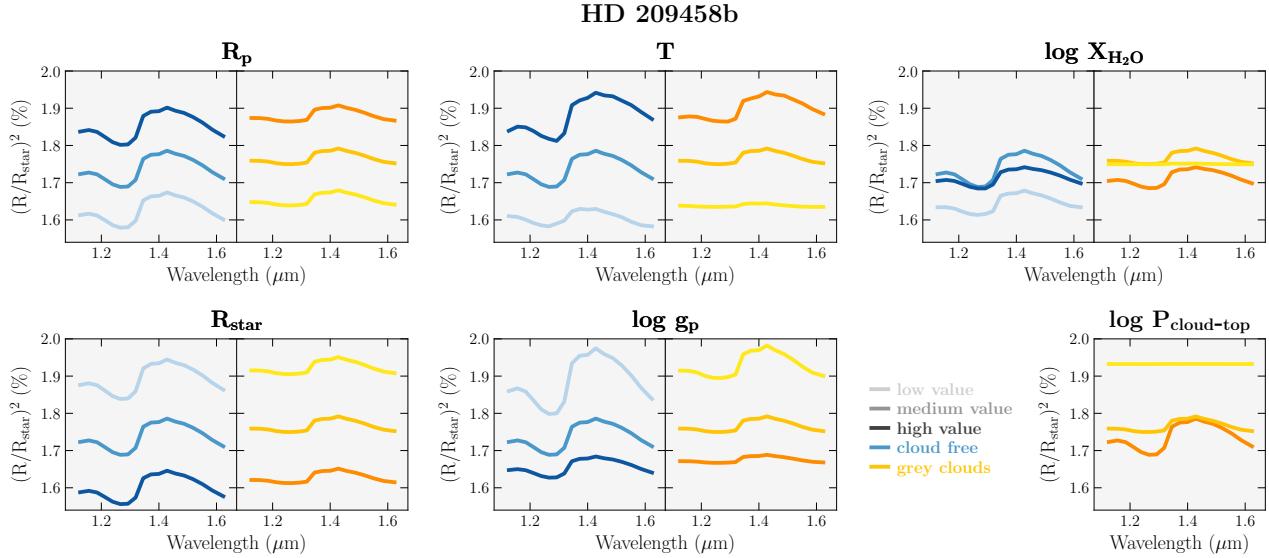


Figure 8. Forward models for HD 209458b data showing the difference when a low (light shades), medium (medium shades), or high (dark shades) value is considered for each free parameter. Cloud free (blue) and grey cloud (yellow) models were tested. For each panel, all values are fixed to the medium value except for the tested parameter (e.g. if the tested parameter is temperature, all other parameters are fixed to their medium values). Values used for each parameter (low, medium, high) are listed in Table 3.

dius, and opacity sources, such as water, CIA, Rayleigh scattering, and clouds. While H_2O , CIA, and Rayleigh scattering are included in all models, the inclusion of clouds is optional. We test the presence of clouds using four independent models: with no clouds (“cloud free”), with grey clouds, with non-grey clouds, and a flat line model. “Grey” clouds consist of particles with sizes greater than the wavelengths probed and result in a constant cloud opacity. Non-grey clouds consist of particles with sizes similar than the wavelengths probed and are associated with a variable opacity across these same wavelengths. Each model has a different number of free parameters, according to their complexity, as listed in Table 2. A “flat line” model is also tested, representing a spectrum with a flat continuum (i.e. no spectral features).

For our grey cloud model, we define a cloud-top pressure parameter $P_{\text{cloud-top}}$ which will establish a minimum altitude below which the atmosphere becomes 100% opaque. Accordingly, for pressures higher than the cloud-top pressure (i.e. altitudes below the level where $P = P_{\text{cloud-top}}$), the cloud optical depth is set to infinity.

For non-grey clouds, we follow Equation 5 of Lueber et al. (2022), motivated by Equation 32 of Kitzmann & Heng (2018):

$$\tau_{\text{nongrey}} = \tau_{\text{ref}} \frac{Q_0 x_{\lambda_{\text{ref}}}^{-a_0} + x_{\lambda_{\text{ref}}}^{0.2}}{Q_0 x_{\lambda}^{-a_0} + x_{\lambda}^{0.2}}, \quad (1)$$

where $x_{\lambda} = 2\pi r_{\text{cloud}}/\lambda$ and $x_{\lambda_{\text{ref}}} = 2\pi r_{\text{cloud}}/\lambda_{\text{ref}}$, λ is the wavelength, λ_{ref} is the reference wavelength fixed at 1 μm , and τ_{ref} is the vertical optical depth at the reference wavelength. Q_0 , a_0 , and r_{cloud} are the composition parameter, small particle slope, and particle ra-

dius, respectively. The cloud-top pressure is a free parameter, and the cloud bottom pressure is fixed to the bottom of the atmosphere, i.e. at 10 bar, because this choice has no consequences for transmission spectra as long as the atmosphere becomes optically thick already at lower pressures. Consequently, τ_{nongrey} is added to the total atmospheric optical depth for pressures higher than the cloud-top pressure, extending the cloud from the cloud-top pressure to the assumed bottom of the atmosphere.

2.5 Non-isothermal retrievals

In order to explore temperature variations with altitude, we also implement non-isothermal retrievals. BeAR has the option to include a variety of different T - P profiles, and in this work we select the polynomial profile. For this, an n -parameter profile sets the boundaries of the $n - 1$ atmospheric layers, and a polynomial is fit between them. In our case, we use linear profiles between the boundaries (see Kitzmann et al. 2020 for details on the T - P profile parametrisation). All non-isothermal runs in this work utilise a 4-parameter T - P profile, which is enough to show the differences between the upper and lower layers, yet not as complex to require long calculations (Kitzmann et al. 2020; Schleich et al. 2024).

2.6 Bayesian model comparison

To effectively explore the multi-dimensional parameter space, we use the method of nested sampling (Skilling 2006) as implemented

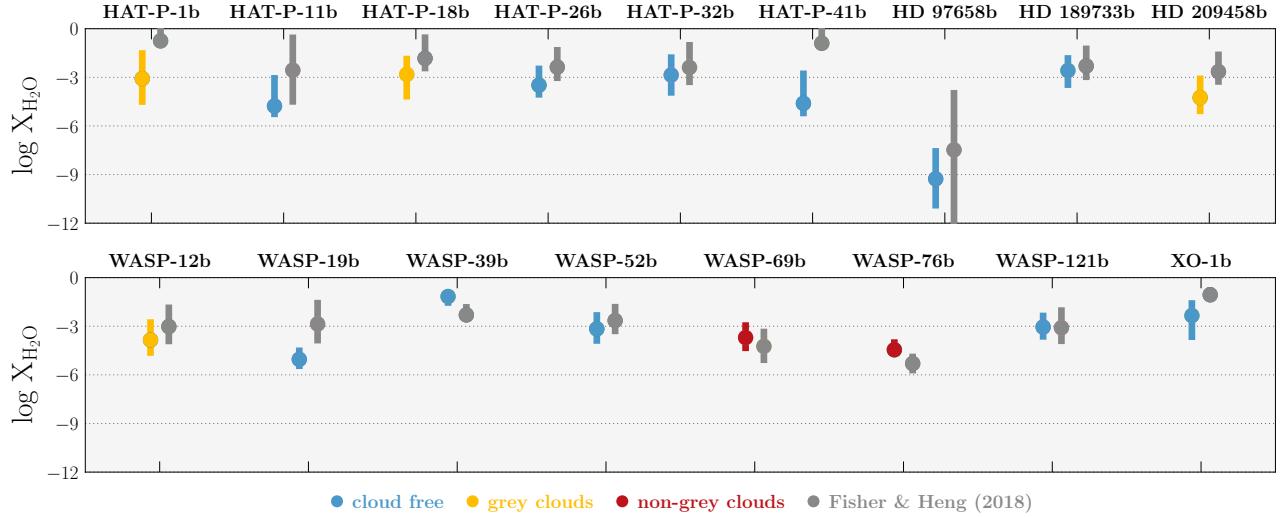


Figure 9. Comparison between retrieved H_2O abundances for all non-flat-line object using the best-fit model for each object. Cloud-free, grey-cloud, and non-grey-cloud models are represented in blue, yellow, and red, respectively. Grey circles correspond to isobaric values found by Fisher & Heng (2018).

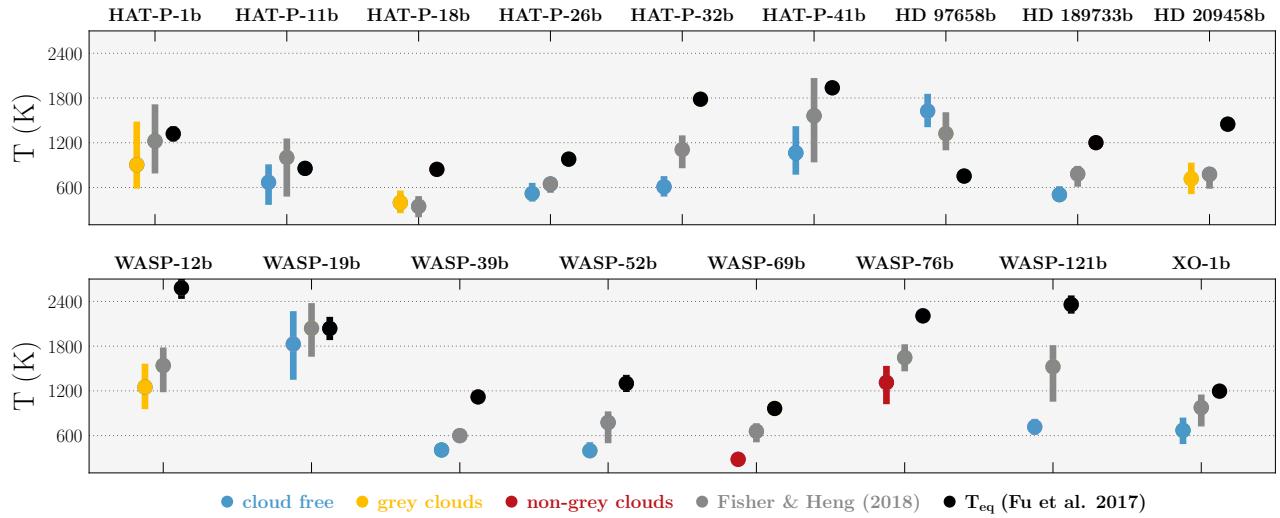


Figure 10. Same as Figure 9, but for retrieved temperatures. Equilibrium temperatures calculated by Fu et al. (2017) are represented in black.

via the MULTINEST software package (Feroz et al. 2009). Nested sampling allows the practitioner to formally implement Occam’s Razor via Bayesian model comparison by calculating the Bayesian evidence (marginalized likelihood) associated with a pair of models (Trotta 2008). Table 1 of Trotta (2008), which is reproduced in Table 2 of Benneke & Seager (2013), lists the correspondence between the so-called Bayes factor (ratio of Bayesian evidences) and number of standard deviations (σ) for which a model is disfavoured relative to the best-fit model.

The first step after running retrievals using our four non-isobaric models (cloud free, grey clouds, non-grey clouds, and flat line) is to compare these models by their corresponding Bayes factors. Bayesian comparison for all 38 objects (42 sets of retrievals) is displayed in Figures A1 and A2, where the natural logarithm of the Bayes factor is calculated relative to the model with the highest Bayesian evidence. Hence, the highest Bayesian evidence model is denoted by a null Bayes factor, and all models with Bayes factors below unity are considered favoured.

According to Figures A1 and A2, a large number of the objects can

be sufficiently fit by a flat line model, i.e. the flat-line model has log Bayes factor of less than unity: GJ 436b, GJ 1214b, GJ 3470b, HAT-P-3b, HAT-P-12b, HAT-P-17b, HAT-P-38b, HD 149026b, WASP-17b, WASP-29b, WASP-31b, WASP-43b, WASP-63b, WASP-67b, WASP-74b, WASP-80b, WASP-101b, TRAPPIST-1d, TRAPPIST-1e, TRAPPIST-1f, and TRAPPIST-g. We continue our study excluding these objects from our analyses, as no further model complexity is required to fit the data. The next sections include analyses of the remaining 17 planets.

We find that most of the non-flat-line objects present at least two models with log Bayes factor < 1 . The cloud complexity defines which are the best-fit models of each object, meaning the simplest model with log Bayes factor < 1 is considered the most favoured model. For example, if the cloud-free and one of the cloudy models both have log Bayes factor < 1 (e.g. HAT-P-11b, HAT-P-26b, HAT-P-32b, HAT-P-41b, HD 97658b, HD 189733b, WASP-19b, WASP-39b, WASP-52b, XO-1b), it means that clouds are not required to fit the data. If only both cloudy models have log Bayes factor < 1 (e.g. HAT-P-1b, HAT-P-18b, HD 209458b, WASP-12b), then the simpler,

grey-cloud model is favoured, as the extra cloud parameters are not required to fit the spectrum. Tables A2 and A3 show our retrieval outcomes for all models, except flat lines (due to the lack of spectral features). Posterior distributions for each object using all models (except flat lines) are provided online (see Data Availability).

3 PARAMETER DEGENERACIES

Degeneracies between multiple parameters are commonly seen in retrievals from low-resolution data. These degeneracies have been previously investigated by a number of studies (e.g. Benneke & Seager 2012; Griffith 2014; Heng & Kitzmann 2017; Fisher & Heng 2018; Welbanks & Madhusudhan 2019). In this section we explore the specific degeneracies in our own retrievals, and particularly focus on the effects of our additional parameters, the stellar radius and planet gravity.

3.1 Correlations between parameters

Figure 1 illustrates the posterior distributions of H_2O abundance, temperature, and reference transit radius (for an example case – HD 209458b) when R_{star} and/or $\log g_p$ are fixed or allowed to vary in the retrievals. We demonstrate that X_{H_2O} and T are barely affected by variable R_{star} and $\log g_p$ assumption, which mainly affects R_p posteriors. A strong degeneracy arises between R_{star} and R_p as seen for all non-flat-line objects in Figure 2, which is expected from the calculation of the transit depth. Figure 3 shows the relationship between the retrieved surface gravity and temperature values. We expect a degeneracy between these parameters due to their opposing effects on the atmospheric scale height, and this can be seen more clearly in some planets than others, such as WASP-39b. However, no clear degeneracy is observed between $\log g_p$ and water abundance for any of the planets, as shown in Figure 4.

The degeneracy between R_{star} and R_p likely dominates the spectral normalization, suppressing the so-called normalization degeneracy expected between R_p and H_2O abundance, as seen in Fisher & Heng (2018). Therefore, Figure 5 shows R_p and X_{H_2O} are not correlated in our retrievals. However, note that the water abundance is quite widely constrained across several orders of magnitude, a result of the underlying normalization degeneracy (Line & Parmentier 2016). The lack of impact of $\log g_p$ and R_{star} on the water abundance is reassuring, as these parameters have regularly been fixed in previous retrieval studies, despite large uncertainties on their values.

Figure 6 shows an additional degeneracy between H_2O abundance and temperature in most cases. This often presents as a “banana-shaped” joint posterior distribution – at lower H_2O abundances, we see temperature decreasing with increasing H_2O abundance, whereas at higher H_2O abundances, the temperature increases with increasing abundance (e.g. HAT-P-11b, HAT-P-26b, HAT-P-32b, HAT-P-41b, HD 189733b, WASP-52b). This behaviour can be explained by the relationship between the spectral feature and the atmospheric scale height. At lower abundances, only the temperature is affecting the scale height, causing an increase in the amplitude of the feature at higher temperatures due to the larger scale height. At very high abundances, the increasing H_2O abundance starts to increase the mean molecular weight of the atmosphere, causing a reduction in the scale height when $\log X_{H_2O} \gtrsim -2$, and thus muting the H_2O spectral features. To compensate this effect, the temperature increases to increase the scale height again.

In retrievals where grey or non-grey clouds are included, there is a further degeneracy between cloud-top pressure and H_2O abundance

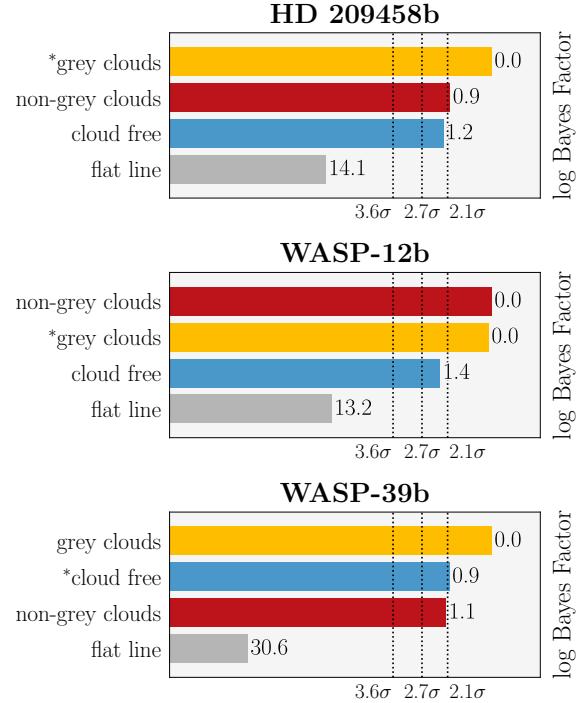


Figure 11. Bayesian comparison between cloud-free (blue), grey-cloud (yellow), and non-grey cloud (red) models for HD 209458b, WASP-12b, and WASP-39b. For each object, models are ordered from highest (top) to lowest (bottom) Bayesian evidence, and consequently from most favoured (highest evidence) to least favoured in comparison to the highest evidence model. Values on the right side of each bar show the corresponding natural logarithm of the Bayes factor relative to the highest-evidence model. Vertical dotted lines show significances of 3.6σ , 2.7σ , and 2.1σ , corresponding to log Bayes factors of 5.0, 2.5, and 1.0, considered “strong”, “moderate”, and “weak” evidences compared to the highest-evidence model, according to Trotta (2008). Therefore, models with $\log \text{Bayes factor} < 1$ have inconclusive evidence compared to the highest-evidence model, and cannot be ruled out as favoured, while models with $\log \text{Bayes factor} \geq 1$ are not favoured in comparison to the highest-evidence model. All runs are calculated at a spectral resolution of 0.1 cm^{-1} .

(Figure 7). This is because the cloud-top pressure can be shifted up and down to mute or extend the spectral feature, balancing the change in abundance.

3.2 Influence of each free parameter

Due to the ambiguity caused by degeneracies, in order to reliably fit the spectral data, retrieval models may artificially compensate the lack of one (or more) parameter(s) with the lack or excess of other(s). In other words, when one limits a parameter to higher values, the model compensates by adjusting other free parameters to higher (or lower) values. Therefore, it is useful to understand the effect of each individual parameter on the transmission spectra. We note that several studies have investigated the effects of atmospheric parameters on the shape of transmission spectra (e.g. Line & Parmentier 2016), but in this work we perform additional tests for our particular model setup. Figure 8 demonstrates how our cloud free, grey cloud, and non-grey cloud modelled spectra behave when parameters (T , X_{H_2O} , R_p , R_{star} , $P_{\text{cloud-top}}$, and $\log g_p$) are fixed to low, medium, or high values, listed in Table 3. We used spectra of HD 209458b as an example, considering input parameters listed in Table A1.

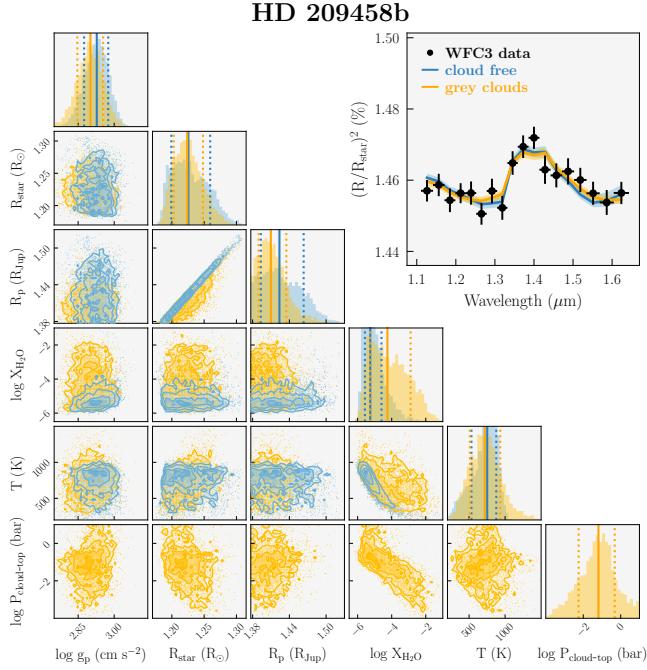


Figure 12. Retrieval outcomes for HD 209458b using our cloud free (blue) and grey cloud (yellow) models. Histograms represent the posterior distribution for each free parameter, with their median value (solid lines) and 1σ uncertainty range (dashed lines). Top-right panel in each figure shows the fitted spectrum and its associated 1σ uncertainty region, calculated at a resolution of 0.1 cm^{-1} and binned to the data resolution for better visualisation. Black points correspond to data observed by *HST* WFC3.

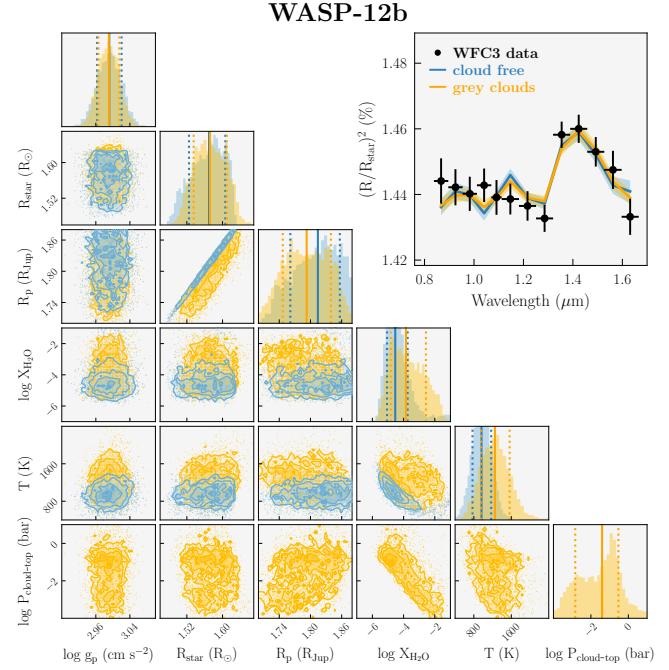


Figure 13. Same as Figure 12 but for WASP-12b.

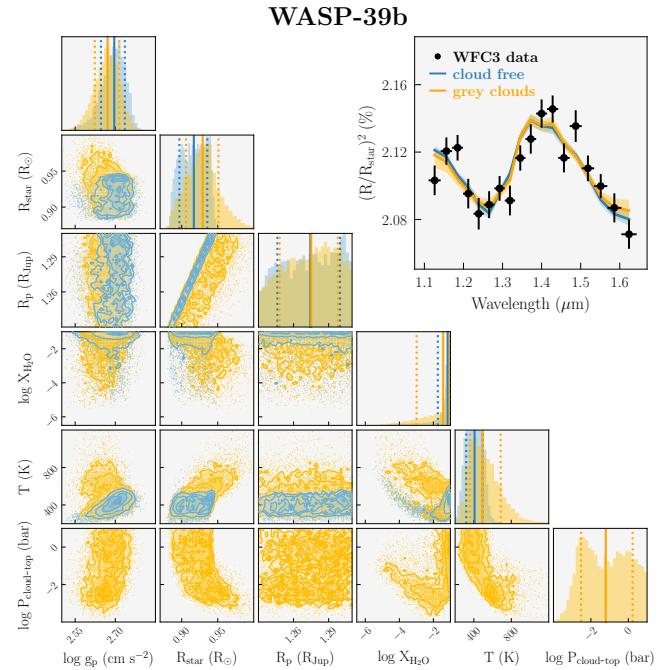


Figure 14. Same as Figures 12 and 13 but for WASP-39b.

4 POPULATION STUDY: FULL SAMPLE RETRIEVALS

4.1 Water abundances

As H_2O is the main absorber in the *HST* WFC3 wavelength range, it is also the most sensitive parameter in our models. Therefore, we examine the H_2O abundances retrieved for the favoured models (cloud free, grey clouds, or non-grey clouds) for each object. These values are represented in Figure 9, along with values retrieved by Fisher & Heng (2018) using an isobaric model.

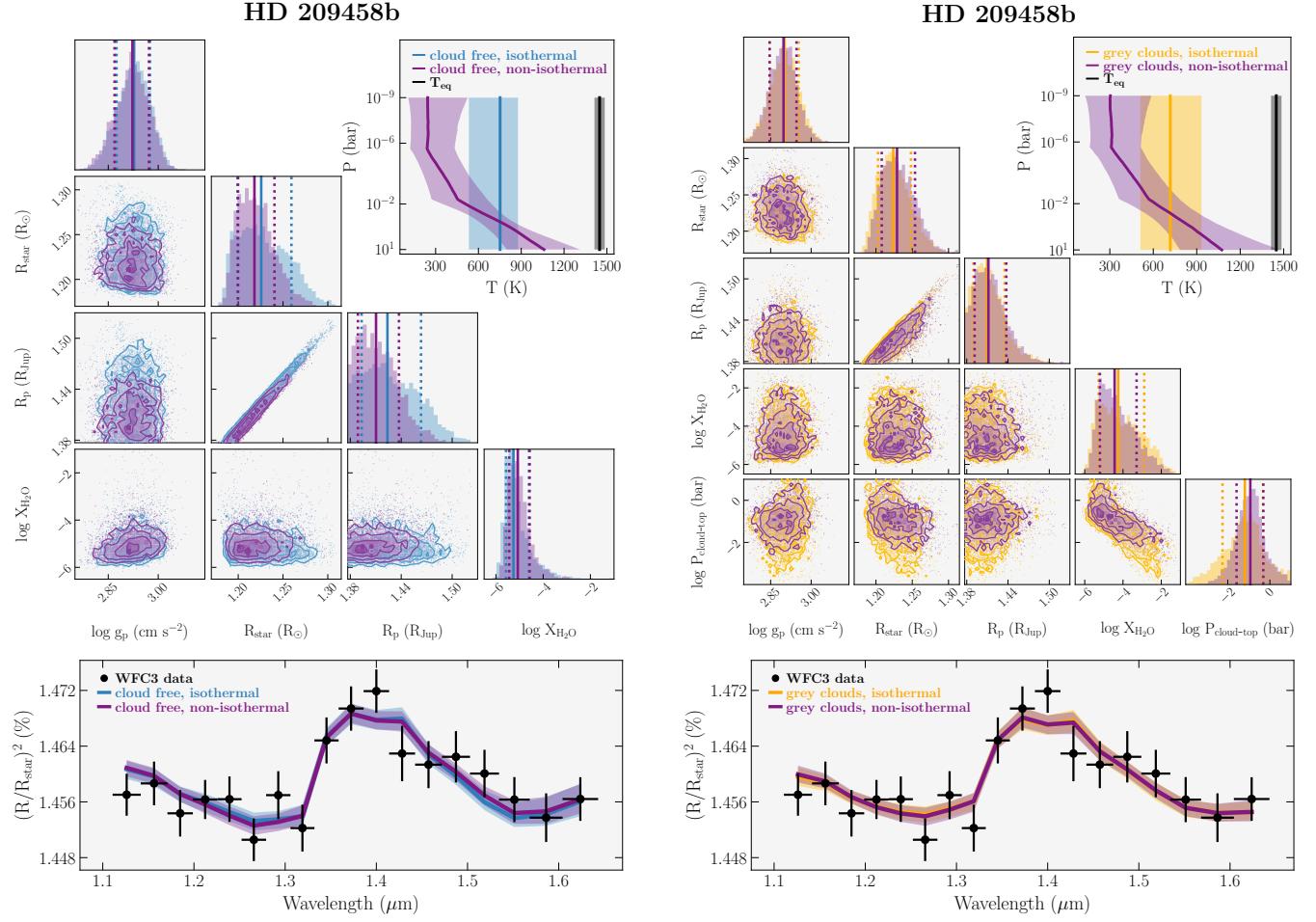


Figure 15. Top figures: HD 209458b isothermal full cloud-free retrieval (blue, left) and grey-cloud retrieval (yellow, right) vs. non-isothermal cloud-free (purple, left) and grey-cloud (purple, right) outcomes. Top-right panel in each figure shows the temperature-pressure profiles, with the equilibrium temperature (T_{eq}) and its 1σ uncertainty calculated by Fu et al. (2017, see Table A1 for each value). Bottom figures: best-fit spectra and its associated 1σ uncertainty region for the same object, binned to the data resolution. Black points correspond to data observed by *HST* WFC3.

Isobaric H₂O abundances retrieved by Fisher & Heng (2018) often present larger uncertainties, and are sometimes inconsistent with our non-isobaric values (e.g. HAT-P-1b and HAT-P-41b). An explanation for these large uncertainties and inconsistencies is the absolute spectral continuum sourced by CIA. In Fisher & Heng (2018), even though CIA is included in all models, their isobaric approach neglects the effects of CIA in their retrievals. As CIA strongly depends on pressure, our pressure-gradient models correctly calculate CIA in each pressure layer, providing more accurate H₂O abundances (as shown in Welbanks & Madhusudhan 2019) and highlighting the need for a non-isobaric assumption.

Abnormally high H₂O abundances close to 10% (i.e. $\log X_{\text{H}_2\text{O}} \approx -1$) are retrieved for WASP-39b, which favours a cloud free model. This is because WFC3 mainly identifies H₂O features, but fails to cover spectral features in wider wavelength ranges, such as molecular and cloud signatures. With the lack of additional opacity features, accurately constraining the water abundance is challenging (as shown in Line & Parmentier 2016) and our outcomes may result in unphysical values (Barstow & Heng 2020). This will be further examined in the next section. Furthermore, H₂O abundances that are too high (or too low) could be compensated within the retrievals by higher or lower values for the other parameters, such as temperature or clouds.

4.2 Temperatures

After water abundance, temperature is the most sensitive parameter in our models. Analogous to Figure 9, Figure 10 examines temperature values retrieved for the favoured models (cloud free, grey clouds, or non-grey clouds) for each object. Isobaric values retrieved by Fisher & Heng (2018) and equilibrium temperatures calculated by Fu et al. (2017) are also shown.

Isobaric temperatures retrieved by Fisher & Heng (2018) are overall comparable or higher than this work's values. As explained in the previous section, our models sets the CIA continuum in each pressure layer, as opposed to Fisher & Heng (2018), which affects the retrieved H₂O abundances, and consequently the scale height, and thus the retrieved temperatures. This effect is also demonstrated in Figure 8.

Low temperatures are a well-known puzzle in atmospheric retrievals of transmission spectra. MacDonald et al. (2020) showed that 1D retrievals often obtain temperatures ~ 1000 K cooler than the planet's equilibrium temperature. They attribute these unphysically low temperatures to the 1D model's inability to reproduce chemical differences between the morning and evening terminators. This effect is found to be strongest for ultra-hot Jupiters. Figure 10 shows that, for nearly all objects, we retrieve a temperature below the equilibrium

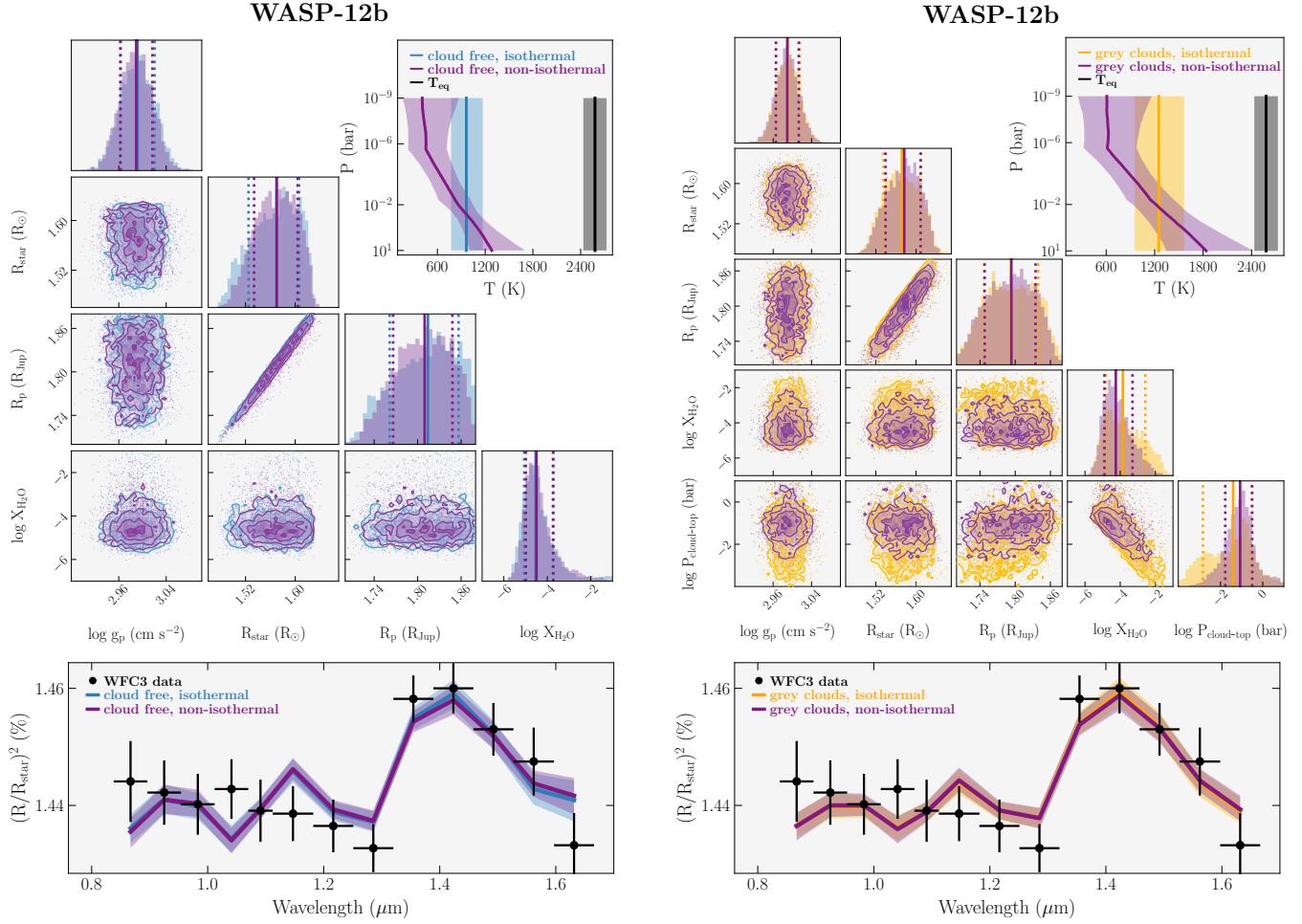


Figure 16. Same as Figure 15 but for WASP-12b.

value of Fu et al. (2017), with differences of sometimes 1000 K or higher.

In addition to limb asymmetries, there are other issues that could cause erroneous temperatures. Firstly, our model assumes an isothermal atmosphere, so not only do we assume longitudinal symmetry, but also vertical symmetry. Since transmission spectra probe a relatively high-altitude region, this could account for some retrieved temperatures being cooler than expected. Secondly, the temperature only arises in two computations in the model – in the calculation of the opacities and the atmospheric scale height. Since the effect of temperature on the opacities is fairly small with respect to the resolution of *HST* data, this is unlikely to be constraining the temperature. Therefore, the main effect of temperature is on the scale height of the atmosphere. With only two spectral features covered by WFC3 (a strong one at $\sim 1.4 \mu\text{m}$ and a weaker one at $\sim 1.2 \mu\text{m}$), and little baseline for the continuum, it can be challenging to retrieve an accurate scale height from this data alone. For example, one of the biggest temperature discrepancies occurs for the ultra-hot Jupiter WASP-121b, which, as previously mentioned, has detections of additional species in its atmosphere that could be affecting this wavelength range (Evans et al. 2018; Mikal-Evans et al. 2022). This is a potential case where neglecting other molecular species could be affecting the retrieval results.

These issues of low retrieved temperatures will be further investigated in Section 5.

5 CASE STUDIES: INVESTIGATING THE LOW TEMPERATURE ‘‘PROBLEM’’

As discussed in Section 4.2, our retrievals obtain substantially lower temperatures than expected from the equilibrium values, in most cases. We speculated that this could be caused by limb asymmetries (as investigated in MacDonald et al. 2020), the narrow and relatively high-altitude pressure region probed by transmission, or challenges retrieving an accurate scale height from WFC3 data. The latter causes further issues in retrieving the abundance of H_2O , demonstrated by the abnormally high abundance values retrieved for WASP-39b.

In order to further investigate the low temperature problem, we choose three planets to study in more detail: HD 209458b, WASP-12b, and WASP-39b. These objects were selected to represent our full dataset due to being three of the best-studied giant exoplanets in the literature. Additionally, WASP-12b and WASP-39b represent some extremes of the previously mentioned cases of low temperature and high H_2O abundance, respectively. HD 209458b provides a more intermediate case study. The next sections provide additional tests on these three objects.

Figure 11 shows that, amongst the models with $\log \text{Bayes factor} < 1$, grey clouds are the simplest model for HD 209458b and WASP-12b, while the cloud-free model is the simplest for WASP-39b. Thus following Occam’s razor, we consider these the best-fit models for each planet (see 2.6 for definition of ‘‘best-fit model’’). In order to

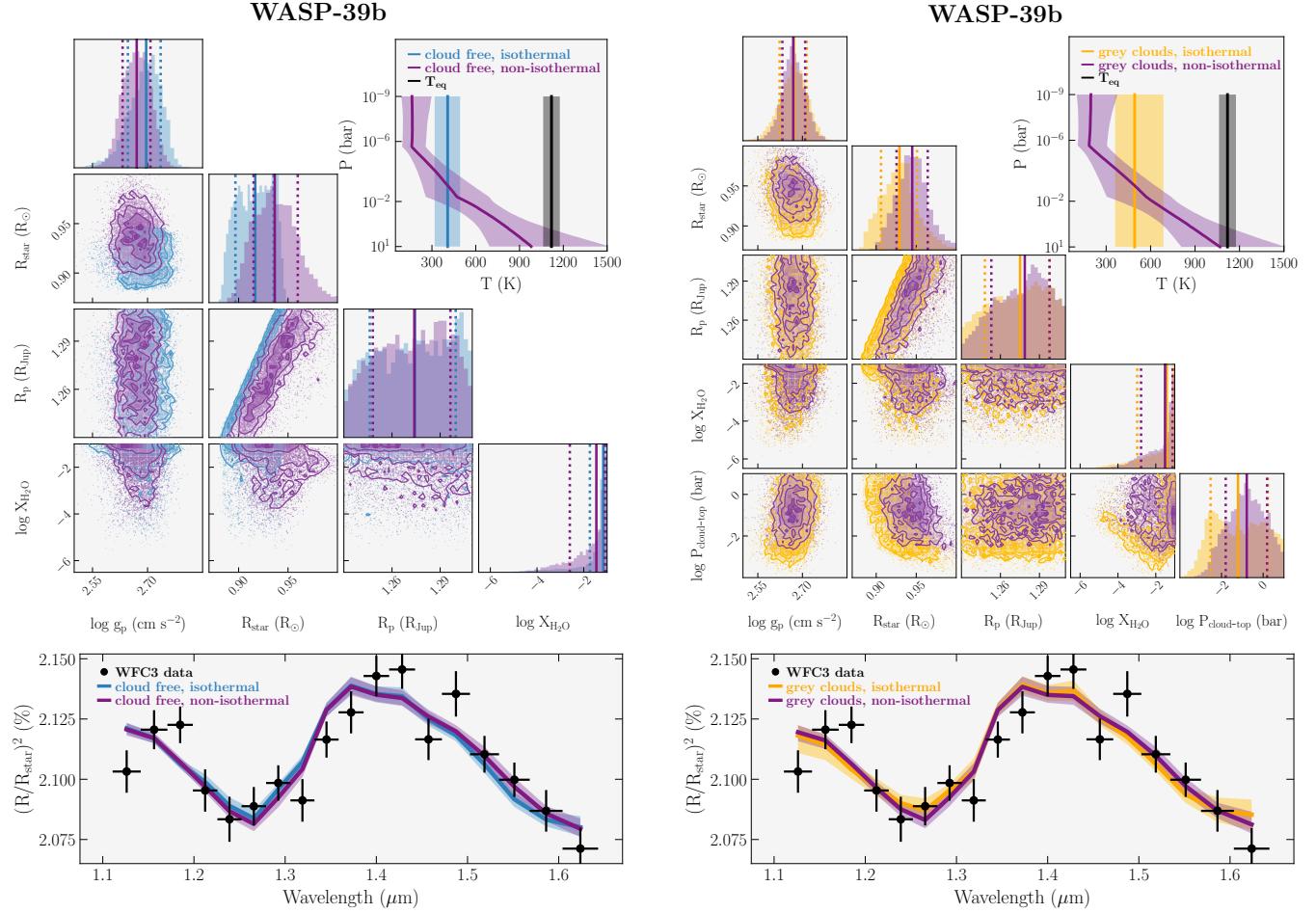


Figure 17. Same as Figures 15 and 16 but for WASP-39b.

compare the presence or absence of clouds in our additional analyses, we perform tests using both the cloud-free and grey-cloud models.

5.1 HD 209458b

Figure 12 highlights the differences between HD 209458b cloud-free and grey-cloud retrieval outcomes. The inclusion of clouds leads to a higher H_2O abundance with a wider posterior, due to the degeneracy between X_{H_2O} and the cloud-top pressure $P_{\text{cloud-top}}$. As previously discussed in Section 3, secondary normalization degeneracies are also present between R_p and R_{\star} and between H_2O and temperature. With the inclusion of clouds, the H_2O abundance is allowed to reach higher values, as expected.

5.2 WASP-12b

As expected, Figure 13 of WASP-12b displays a similar behaviour for X_{H_2O} and R_p posteriors as for HD 209458b, where the addition of clouds leads to a wider X_{H_2O} posterior. $X_{H_2O}-P_{\text{cloud-top}}$, $X_{H_2O}-T$, and R_p-R_{\star} degeneracies are also present. In particular, although the median and 1σ cloud-free and grey-cloud retrieved temperatures remain substantially lower than the T_{eq} value of 2580 K, the grey-cloud model retrieves higher temperatures, and the grey-cloud posterior distribution encompasses higher temperatures in its posterior tail.

5.3 WASP-39b

Figure 14 shows cloud-free and grey-cloud retrieval outcomes for WASP-39b provide very high H_2O abundances and very low temperatures. As the cloud-free model is favoured, and therefore clouds are not required to fit the data, the inclusion of clouds in the case of WASP-39b has a more modest effect in comparison to HD 209458b and WASP-12b. The H_2O posterior has a stronger left tail tending to slightly lower values when clouds are included. The cloud-top pressure is fairly unconstrained, demonstrating its lack of effect on the spectrum due to the extremely high retrieved water abundance. The $X_{H_2O}-T$ degeneracy for the cloud-free model displayed here is representative of the high abundance section of the banana shape discussed in Section 3, where the high H_2O abundance increases the mean molecular weight, which is balanced by an increase in temperature.

5.4 Temperature tests

As previously mentioned, most spectra in our sample retrieve temperatures much lower than their equilibrium values. In this sub-section we explore some of the possible explanations for these low temperatures. Firstly, we test the effect of the isothermal assumption on our case studies, to investigate if the narrow and high-altitude pressure region probed by transmission could explain the low temperatures. Secondly, we repeat our retrievals with constrained prior on tempera-

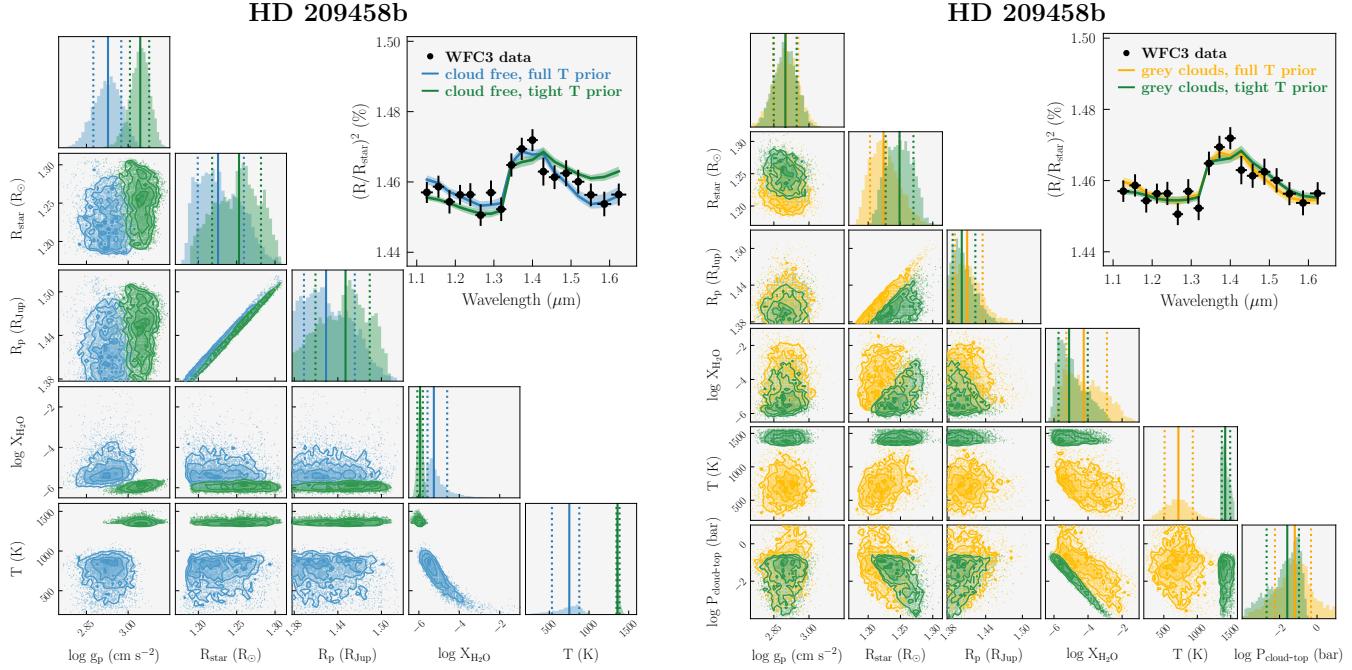


Figure 18. Retrievals for HD 209458b using a full temperature prior range (200–3100 K) vs. using a tight prior range around the equilibrium temperature (1341–1557 K, i.e. $T_{\text{eq}} \pm 3\sigma$). Blue and yellow represent cloud free and grey cloud models using full priors, while green represents the same models but using tight priors. T_{eq} and σ values are listed in Table A1. Top-right panel in each corner plot shows the fitted spectrum and its associated 1σ uncertainty region, binned to the data resolution. Black points correspond to data observed by *HST* WFC3.

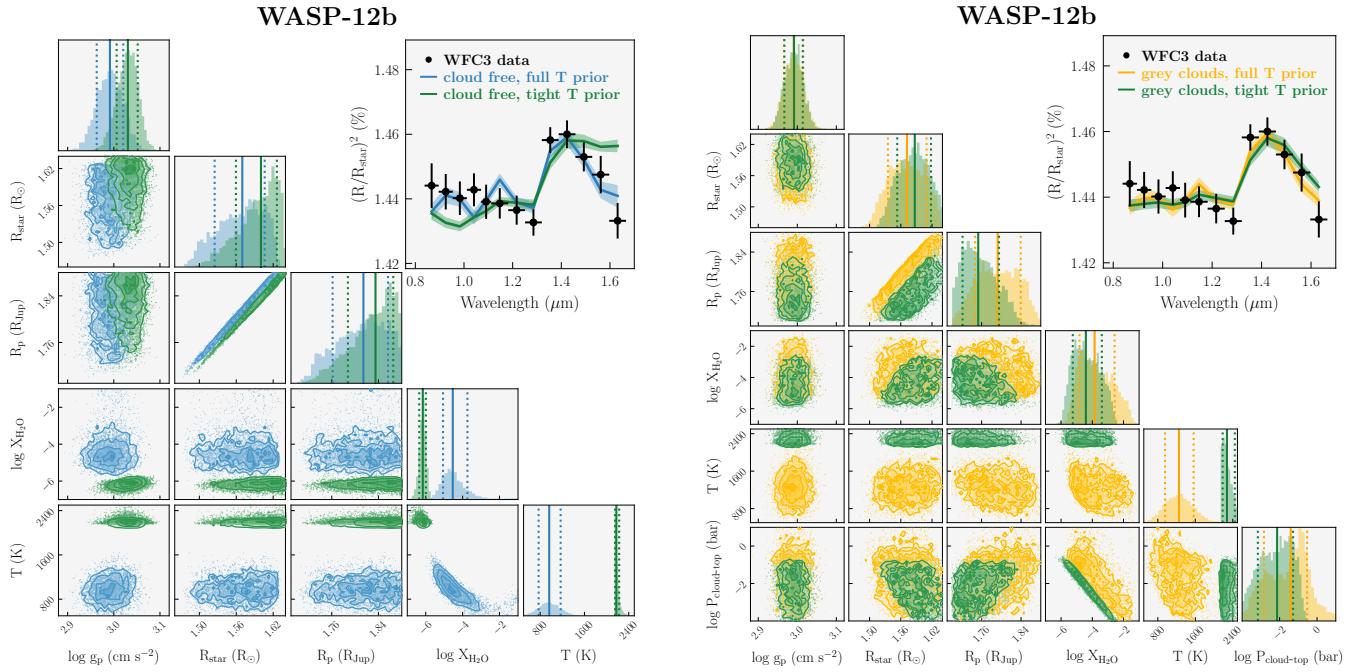


Figure 19. Same as Figure 18 but for WASP-12b, using a tight temperature prior range of 2142–3018 K (i.e. $T_{\text{eq}} \pm 3\sigma$).

ture around the equilibrium value, forcing the temperature posteriors to higher values, and compare the results with the wider prior retrievals.

5.4.1 Isothermal vs. non-isothermal retrievals

In order to test the effect of the isothermal assumption, we performed retrievals using cloud-free and grey-cloud non-isothermal models, with the setup described in Section 2.5. Temperature-pressure profiles, posterior distributions, and modelled spectra are displayed in Figures 15, 16, and 17 for HD 209458b, WASP-12b, and WASP-39b,

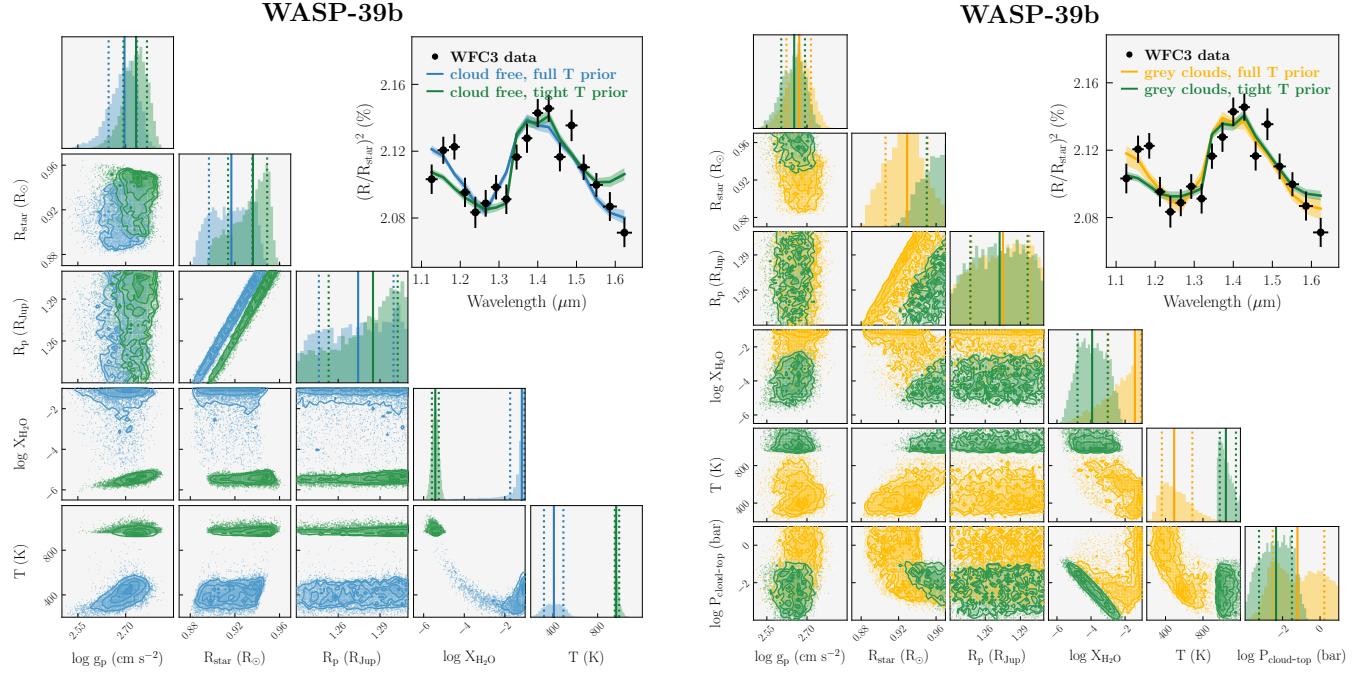


Figure 20. Same as Figures 18 and 19 but for WASP-39b, using a tight temperature prior range of 948–1290 K (i.e. $T_{\text{eq}} \pm 3\sigma$).

Planet	Cloud Free		Grey Cloud	
	Full Prior	Tight Prior	Full Prior	Tight Prior
HD 209458b	$1.07^{+0.24}_{-0.26}$	$2.34^{+0.33}_{-0.27}$	$1.05^{+0.22}_{-0.11}$	$1.14^{+0.19}_{-0.10}$
WASP-12b	$1.71^{+0.39}_{-0.20}$	$4.8^{+0.49}_{-0.37}$	$1.54^{+0.42}_{-0.22}$	$1.98^{+0.39}_{-0.22}$
WASP-39b	$2.27^{+0.37}_{-0.24}$	$3.85^{+0.32}_{-0.25}$	$2.58^{+0.32}_{-0.28}$	$2.94^{+0.21}_{-0.14}$

Table 4. Reduced χ^2 values for the retrievals of HD 209458b, WASP-12b, and WASP-39b using the two different temperature priors – the full prior and the tight prior around the T_{eq} value.

respectively. Retrieved parameter values are listed in Table A4. The isothermal retrievals are the same as previously presented (Figures 12 to 14).

The first aspect to notice is the temperature profile on the top-right corner of each figure (15, 16, and 17). For all three planets, the temperature profiles show a steep decrease from 10 to 10^{-5} – 10^{-6} bar, and remain approximately isothermal above this region. At altitudes around 10^{-1} – 10^{-3} bar, where the *HST* WFC3 transmission spectra are expected to probe, the non-isothermal temperature values are slightly lower than isothermal, yet they are generally consistent within the uncertainty range. All isothermal and non-isothermal profiles are substantially lower than the equilibrium temperatures. For WASP-39b, the non-isothermal profile does overlap with the T_{eq} at the very bottom of the atmosphere (~ 10 bar), however it is highly unconstrained as we do not expect transmission spectra to probe these deep pressures.

We call attention to the sharp posteriors and elevated values for H₂O abundances retrieved for WASP-39b, which confirms high H₂O abundances are not caused by our isothermal assumption. The agreement between the isothermal and non-isothermal temperatures at the WFC3 photospheric region also suggests that the isothermal assumption is sufficient for interpreting these low-resolution spectra.

In fact, despite the wider wavelength coverage and increased precision of *JWST*, Lueber et al. (2024) demonstrated that the isothermal model is sometimes favoured for their retrievals on *JWST* transmission spectra of WASP-39b, in particular for the NIRISS and PRISM instruments.

In contrast, for WASP-12b even the deep atmosphere is inconsistent with the equilibrium temperature. However, as investigated in MacDonald et al. (2020), this ultra-hot Jupiter is a likely candidate for strong limb asymmetries, which can cause substantially lower temperatures in 1D retrievals.

Whilst we use these non-isothermal retrievals to test for variable temperature structures in the atmospheres, it is worth noting that the molecular abundances are kept constant with altitude. This could introduce variations in the temperature profile that compensate for the underlying water abundance profile. Although testing these non-constant molecular profiles is out of the scope of this paper, it is an important caveat to bear in mind.

5.4.2 Tight temperature priors

So far, all our retrievals used a temperature prior range of 200–3100 K. We now perform a test in which we narrow the uniform temperature prior range to values around the equilibrium temperature for each object, forcing the retrieved temperature to higher values. Our priors for this test range from $T_{\text{eq}} - 3\sigma$ to $T_{\text{eq}} + 3\sigma$, where T_{eq} are the equilibrium temperatures calculated by Fu et al. (2017, assuming zero albedo and uniform re-distribution of heat) and σ is the standard deviation. These values are listed in Table A1.

Figures 18, 19, and 20 show how our cloud-free and grey-cloud retrievals are affected by this test. Retrieved parameter values are listed in Table A5. Full retrievals are the same as in Figures 12 to 14. Full retrievals show a degeneracy between temperature and H₂O abundance, and tight prior retrievals exhibit the same behaviour, presenting lower X_{H₂O} due to the forced high temperatures. In all

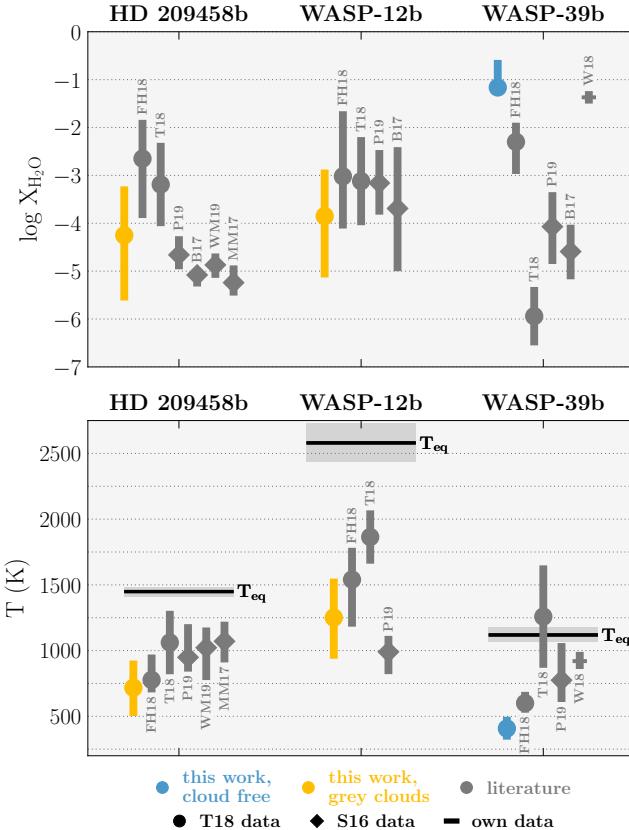


Figure 21. Comparison between retrieved H_2O abundances (top) and temperatures (bottom) for the best-fit cloud-free (blue) and grey-cloud (yellow) models using BeAR, and for best-fit models found by different works (grey). Abbreviations in the plot correspond to the works of Barstow et al. 2017 (B17), Fisher & Heng 2018 (FH18), Tsiaras et al. 2018 (T18), Wakeford et al. 2018 (W18), Pinhas et al. 2019 (P19), Welbanks & Madhusudhan 2019 (WM19), and MacDonald & Madhusudhan 2017a (MM17). Values retrieved using data reduced by Tsiaras et al. (2018) and Sing et al. 2016 (S16) are represented by circles and diamonds, respectively. Works that use their own data are denoted by a horizontal line. For non-isothermal retrievals (e.g. P19, WM19, MM17), the temperatures at the bottom of the atmosphere. T_{eq} represents equilibrium temperatures by Fu et al. (2017).

cases, the temperature posterior is up against the lower limit of the prior, as expected.

When tight priors are used, the cloud-free best-fit spectra in Figures 18, 19, and 20 (left panels) reveal an unusually high CIA continuum in comparison to the tight prior, grey-cloud fits (right panels). This is likely due to the tight prior retrievals being forced to extremely low H_2O abundances, to compensate for the higher temperatures, and the lack of additional absorbers in our models. This results in a visibly worse fit to the data, particularly at the reddest wavelength points.

WASP-39b is an extreme case, as the very high X_{H_2O} median value decreases by about 6 orders of magnitude in the cloud-free tight-prior results, and 3 orders of magnitude in the grey-cloud tight-prior results. In the grey-cloud tight temperature prior retrievals (right panels of Figure 20), the increased scale height from the higher temperature values can be compensated for by a combination of a lower X_{H_2O} , decreasing the mean molecular weight, and a lower $P_{\text{cloud-top}}$, bringing up the continuum level to reduce the size of the spectral feature. As for the cloud-free tight-prior retrievals (left panels of Figure 20), a different effect occurs. Besides lower X_{H_2O} values,

the high scale height from the high temperatures are additionally compensated by an increase in the planetary surface gravity, which is not seen in the grey-cloud retrievals.

This tight temperature priors test supplements our previous conclusions: as clouds are an extra free parameter to normalize the spectral continuum, they are necessary to compensate for too high or too low X_{H_2O} and temperature values, providing a more accurate fit to WFC3 spectra. When clouds are not included, the continuum can be compensated for by a variable, pressure-dependent planetary surface gravity. Whilst this tight temperature prior test demonstrated an alternative to accurately fit WFC3 transmission spectra, a careful consideration of model degeneracies and data limitations is still indispensable. Additionally, these parameters are not enough to cover missing molecular features, accentuating the necessity of a broader wavelength coverage.

Although these tight prior tests offer an alternative fit to the data that is perhaps more physically realistic, their fits remain worse than the cooler-temperature solutions found by the wide prior retrievals. A reduced χ^2 test reveals values consistently closer to 1 for the wide prior retrievals for HD 2095458b and WASP-12b (see Table 4), as expected since the higher temperatures were rejected in the wide prior retrievals. For WASP-39b, the reduced χ^2 values reveal that the fit was already poor in the wide prior cases, indicating that our model struggles to fit this dataset. This could be due to missing physics, or additional noise in the data.

Whilst Section 5.4.1 demonstrated that these low temperatures could be explained by the high altitudes probed by transmission, it could also be due to the quality of the spectra and the data-driven nature of atmospheric retrievals that could favour an incorrect solution. For example, additional noise effects that are unaccounted for in the error bars could lead to biased retrieval results. This serves as another warning for relying on the results from WFC3 spectra without careful consideration of model degeneracies and data limitations.

6 COMPARISON WITH OTHER WORKS

As H_2O is the main absorber in the WFC3 wavelength range, it is also the most sensitive parameter in our models. Whilst the error bars are wide, several planets do get quite different H_2O abundances for different models. Moreover, the number of degeneracies regarding retrieve abundance values, in agreement with previous studies (e.g. Benneke & Seager 2012; Griffith 2014; Line & Parmentier 2016; Welbanks & Madhusudhan 2019; Fairman et al. 2024).

Although the WFC3 range encompasses two water features, it fails to cover additional opacity features in wider spectral ranges, such as molecular and haze signatures, which may result in unreliable retrieved H_2O abundances (Barstow & Heng 2020). Figure 21 presents our cloud-free and grey-cloud H_2O abundance and temperature values found for HD 209458b, WASP-12b, and WASP-39b, along with values retrieved by a number of studies (Barstow et al. 2017; Fisher & Heng 2018; Tsiaras et al. 2018; Wakeford et al. 2018; Pinhas et al. 2019; Welbanks & Madhusudhan 2019; MacDonald & Madhusudhan 2017a).

6.1 Dependence on model properties

Retrieved H_2O abundances in the literature may vary in a range between $-6 < \log X_{H_2O} < -2$ (including uncertainties) for HD 209458b and for WASP-12b. This range encompasses the values

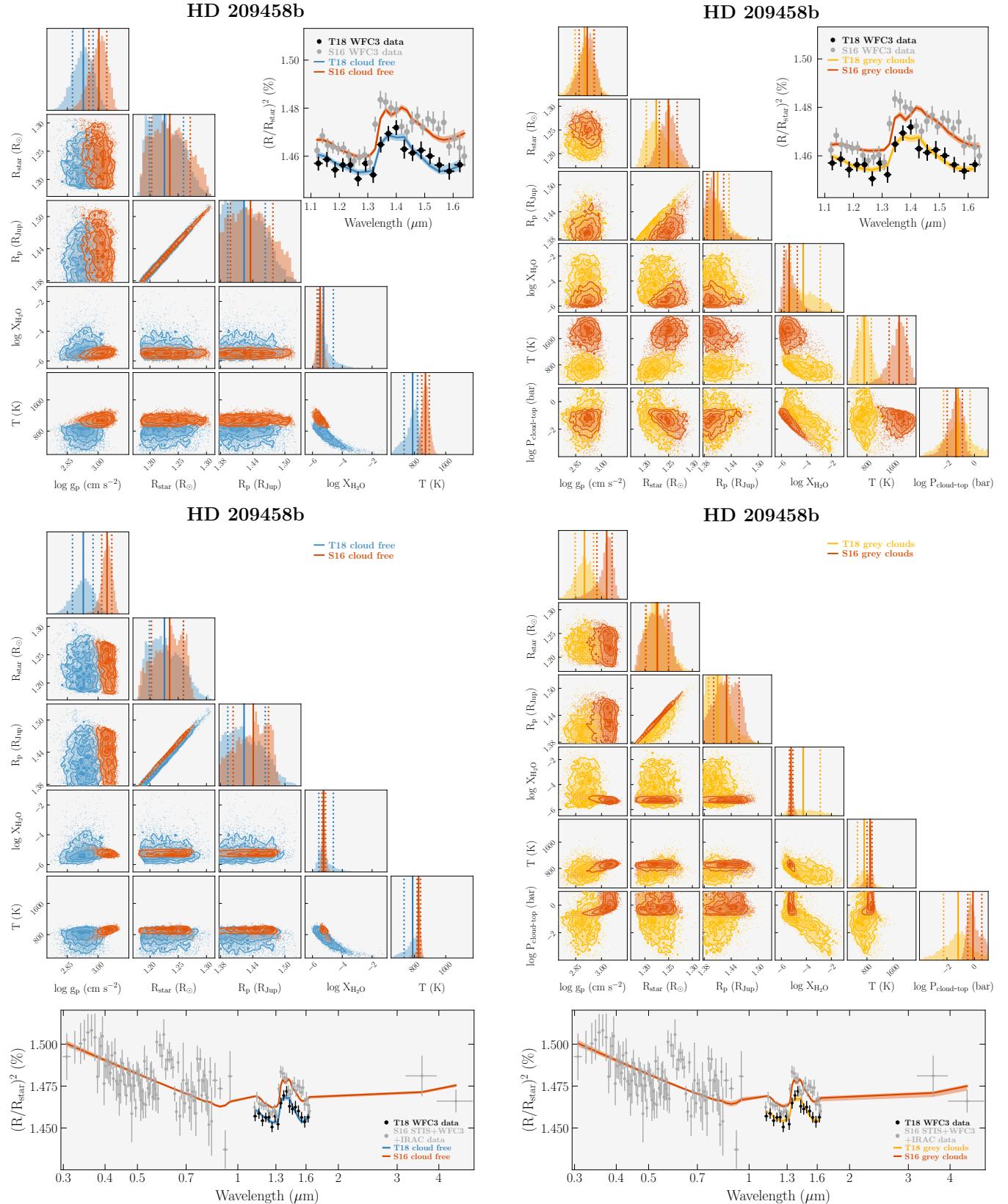


Figure 22. HD 209458b full cloud-free (blue, left) and grey-cloud (yellow, right) retrievals using data reduced by Tsiaras et al. 2018 (T18) vs. cloud-free (orange, left) and grey-cloud (orange, right) retrievals using data reduced by Sing et al. 2016 (S16). Black points and grey points correspond to data reduced by Tsiaras et al. (2018) and Sing et al. (2016), respectively. Top corner plots show retrievals using Tsiaras et al. (2018) and Sing et al. (2016) *HST* WFC3 data only, and middle corner plots show retrievals using Tsiaras et al. (2018) *HST* WFC3 and Sing et al. (2016) *HST* STIS, *HST* WFC3, and *Spitzer* IRAC data. Bottom spectra are the best-fit models and their associated 1σ uncertainty regions corresponding to the middle corner plots. All spectra are calculated at a resolution of 0.1 cm^{-1} and binned to the data resolution.

found by our current work. Moreover, retrieved temperatures also diverge between different studies, and are generally not consistent with their associated equilibrium temperatures (1449 ± 36 K, 2580 ± 146 K, and 1119 ± 57 K for HD 209458b, WASP-12b, and WASP-39b, respectively; [Fu et al. 2017](#)). In order to fairly compare these values, individual assumptions of each work have to be taken into account. For HD 209458b and WASP-12b data, [Fisher & Heng \(2018\)](#) apply an isobaric approximation which has little to no effect on the CIA, resulting in higher H_2O abundances. [Tsiaras et al. \(2018\)](#) retrieve values from WFC3 spectra assuming an isothermal temperature-pressure profile that includes CIA and Rayleigh scattering, in agreement with our analyses. However, their cloud parameterization is notably different from ours, which may cause discrepancies in retrieved H_2O abundances ([Barstow 2020](#)). [Pinhas et al. \(2019\)](#) consider a non-isothermal temperature-pressure profile, and their temperature values in Figure 21 correspond to the temperature at the top of the atmosphere (minimum pressure). Additionally, [Pinhas et al. \(2019\)](#) also include inhomogeneous clouds. Some of these different modelling assumptions could explain the discrepancies in our results for WASP-39b (Figure 21), though the results for HD 209458b and WASP-12b are consistent within the error bars. [Barstow et al. \(2017\)](#) use a fixed temperature profile for their models, based on assumptions about the planets' Bond albedos. Hence we do not include their temperatures in Figure 21.

Studies which focused on retrievals for an individual planet rather than a population study (e.g. [Welbanks & Madhusudhan 2019](#) and [MacDonald & Madhusudhan 2017a](#) for HD 209458b, and [Wakeford et al. 2018](#) for WASP-39b) were also represented in Figure 21. [Welbanks & Madhusudhan \(2019\)](#) and [MacDonald & Madhusudhan \(2017a\)](#) include a non-isothermal T - P profile and inhomogeneous cloud coverage, as well as optical data from [Sing et al. 2016](#), resulting in narrower uncertainty limits. Yet, our results for $\log X_{\text{H}_2\text{O}}$ remain consistent with theirs within the error bars. For HD 209458b, [MacDonald & Madhusudhan 2017b](#) find tentative evidence for NH_3 in the spectra, which could also explain some differences in our results.

6.2 Dependence on data reduction

Some of the differences in Figure 21 can be explained by the use of different datasets or different data reductions. We highlight the case of WASP-39b, for which our models present abnormally high H_2O values due to the limited range of WFC3, which only probes water features. [Wakeford et al. \(2018\)](#) also find an elevated H_2O abundance for WASP-39b using WFC3 data combined with data from *HST* STIS, VLT FORS2, and *Spitzer* IRAC, which shows these elevated abundance values are not exclusive to our models. On the other hand, [Pinhas et al. \(2019\)](#) find lower H_2O abundances while also considering spectral data from *HST* STIS, *HST* WFC3 (using different data reductions than [Wakeford et al. 2018](#)), and *Spitzer* IRAC. The wide scatter in H_2O abundance across the different studies for WASP-39b demonstrates the necessity of wider wavelength coverage, from individual or multiple instruments, in order to robustly constrain molecular abundances. Further studies using *JWST* spectra with data from $0.5 \mu\text{m}$ up to $5 \mu\text{m}$ (e.g. [Ahrer et al. 2023](#); [Alderson et al. 2023](#); [Feinstein et al. 2023](#); [Rustamkulov et al. 2023](#); [Lueber et al. 2024](#)) are able to infer the presence of clouds, and constrain abundances of H_2O , CO_2 , SO_2 , and possibly other species with features in the infrared region. However, a recent retrieval study using *JWST* spectra still retrieves high H_2O abundance values from NIRISS data, where H_2O is the main opacity source ([Lueber et al. 2024](#)).

Finally, we briefly explore different data reductions and how they can affect our retrieval outcomes. We perform retrievals for HD 209458b using the two cited data reductions: by [Sing et al. \(2016\)](#), used by e.g. [Barstow et al. 2017](#), [Pinhas et al. 2019](#), [Welbanks & Madhusudhan 2019](#), and [MacDonald & Madhusudhan 2017a](#), and by [Tsiaras et al. \(2018\)](#), used by e.g. [Fisher & Heng 2018](#) and this work. Since [Sing et al. \(2016\)](#) additionally include *HST* STIS and *Spitzer* IRAC data aside from WFC3 data, we chose to run separate retrievals using only WFC3 and using STIS+WFC3+IRAC combined data. These data sets were tested using the same cloud-free and grey-cloud models presented throughout this work. Retrieval outcomes from the [Sing et al. \(2016\)](#) data are presented together with the [Tsiaras et al. \(2018\)](#) data in Figure 22 and Table A6. By comparing the [Tsiaras et al. \(2018\)](#) and [Sing et al. \(2016\)](#) WFC3 data, it becomes clear that the data have a significant vertical offset in this wavelength range, with the [Sing et al. \(2016\)](#) data presenting higher transit depths in comparison to [Tsiaras et al. \(2018\)](#). To first order, one would expect the retrieval to adjust the reference transit radius R_p in order to compensate for the spectral offset. However, the WFC3 spectra also present different shapes at wavelengths above $1.3 \mu\text{m}$, changing the size and shape of the $1.4 \mu\text{m}$ H_2O feature. The consequence is a higher scale height for the [Sing et al. \(2016\)](#) spectrum, which leads to higher temperatures in the retrieval of their data. Yet, in the cloud-free case both temperatures retrieved from the [Sing et al. \(2016\)](#) and [Tsiaras et al. \(2018\)](#) data are still low compared to the expected T_{eq} of 1450 K ([Fu et al. 2017](#)). Interestingly, for the grey cloud retrieval of the [Sing et al. \(2016\)](#) data, the temperature posterior now encompasses the equilibrium value, suggesting that the data reductions themselves could be an explanation for the low retrieved temperatures in our results. The retrieved abundance of H_2O is more tightly constrained to the lower values for the [Sing et al. \(2016\)](#) data, especially in the retrieval including grey clouds. The H_2O abundances retrieved from the [Sing et al. \(2016\)](#) data, presented in Figure 22 and Table A6, are fully consistent with results from other studies that use the same data, as exhibited in Figure 21.

When STIS and IRAC data are included, we verify the wider wavelength coverage is able to provide sharper posteriors for H_2O abundance and temperature, consequently breaking degeneracies between $X_{\text{H}_2\text{O}}$, T , and $P_{\text{cloud-top}}$, in agreement with [Pinhas et al. \(2019\)](#) and [Fairman et al. \(2024\)](#). We note that our models include H_2O as the only species besides background gases H_2 and He , therefore no additional molecular features are seen in the retrieved spectra. The grey cloud retrieval on the combined data displays a different behaviour to the WFC3-only case. Now, the cloud-top pressure is constrained to deep in the atmosphere, indicating a lack of cloud coverage, likely due to the strong Rayleigh slope provided by the STIS data, or to the non-inclusion of optical species such as Na and K. This then results in a similar temperature posterior as the cloud-free case, as expected, tightly constrained around cooler values. Our retrieved H_2O abundances are consistent with other studies that use [Sing et al. \(2016\)](#) data (see Figure 21), and with results for Sing WFC3 data only. Overall, the agreement between independent results using spectra reduced by [Sing et al. \(2016\)](#) indicates chemical abundances and temperature posterior constraints are dependent on the data reduction.

It is evident that results are extremely sensitive to each step of the characterisation, from data reduction and calibration to retrieval model assumptions (e.g. if they include Rayleigh scattering, CIA, non-isothermal profile, which chemical species are considered, and how clouds are parameterized). This means even small changes in any step of the procedure may result in inconsistencies. Furthermore, we acknowledge that retrievals using data from WFC3 alone may be unreliable for some planets, highlighting the need to cover a wider

wavelength range to better constrain chemical composition and cloud features (Barstow & Heng 2020; Fairman et al. 2024). The *JWST* follow-up of planets previously observed with WFC3 will bring to light the reliability of the constraints from these measurements.

7 SUMMARY

Our main findings include:

- A non-isobaric treatment is important in establishing chemical abundances in retrievals from *HST* WFC3 transmission data due to the effects of collision-induced absorption (CIA), corroborating the conclusion of Welbanks & Madhusudhan (2019);
- The inclusion of the stellar radius and the planetary surface gravity as free parameters in the retrievals does not have a strong effect on the other retrieved parameters, with the exception of planetary radius, as expected;
- The large number of degeneracies present in WFC3 retrievals can make accurate constraints a challenge. Strong degeneracies exist between H₂O abundance, temperature, and cloud-top pressure, as well as an expected degeneracy between stellar and planetary radii. The limited wavelength coverage of WFC3 makes it difficult to measure the atmospheric scale height, which in turn affects the retrieved parameters;
- Retrievals of WFC3 data often obtain temperatures substantially lower than the equilibrium values. This has a number of possible explanations, such as 3D effects (MacDonald et al. 2020), the above-mentioned challenges in measuring the scale height, or the narrow and high-altitude pressure region probed by transmission. These low temperatures could lead to unreliable constraints on chemical abundances. Yet, non-isothermal retrievals from the same data also result in cool temperature profiles.
- Forcing the temperature priors to higher values can lead to lower H₂O abundances. This may be compensated by the inclusion of clouds, or by a pressure-dependent planetary surface gravity when clouds are not included. However, in this case the temperature posterior hits the lower limit of the prior, and the fit to the data is generally worse. It is possible that the cause of some of these low temperatures could be the data quality, since the low-resolution of WFC3 makes the retrievals highly sensitive to shifts in the data points.
- Differences in data reductions and the inclusion of wider-wavelength data can cause major differences in the retrieval results. This indicates that singular retrievals on WFC3 data alone are unlikely to be reliable.

In conclusion, our study advocates for the importance of wide wavelength coverage for accurately characterising the atmospheres of even the most observable exoplanetary atmospheres. However, we cannot rely on *JWST* alone. Fairman et al. (2024) already demonstrated the importance of optical data for constraining atmospheric cloud properties, which in turn will affect spectra in the wavelengths covered by *JWST*. We also stress the importance of understanding the many degeneracies that exist in retrievals. Additionally, Nixon et al. (2024) discuss the need for accounting for model uncertainty when reporting constraints, which has a further effect of increasing abundance uncertainties. By combining data from multiple instruments, and fully understanding the limitations of our atmospheric retrievals, we can hope to build the most accurate picture of an exoplanet's atmosphere.

ACKNOWLEDGEMENTS

The authors thank the anonymous referee for their helpful insights. We thank Jens Hoeijmakers and Bibiana Prinoth for setting up the GPU at Lund University, allowing us to run BeAR. We thank Johannes Buchner for the Nested Sampling Monte Carlo library. We thank Anna Lueber for useful feedback. A.N. acknowledges financial support from The Fund of the Walter Gyllenberg Foundation. C.F. acknowledges financial support from the Swiss National Science Foundation (SNSF) Mobility Fellowship under grant no. P500PT_203110 and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement no 805445. L.G. acknowledges financial support from Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), through the ARC research grant E-26/211.386/2019. B.T. acknowledges the financial support from the Wenner-Gren Foundation (WGF2022-0041). K.H. acknowledges partial funding support from an ERC Consolidator grant (EXOKLEIN) and an ERC Synergy grant (Geoastronomy).

DATA AVAILABILITY

The open-source atmospheric code BeAR (Kitzmann et al. 2020) is publicly available at <https://github.com/newstrangeworlds/bear>. The open-source nested-sampling package PYMULTINEST (Buchner et al. 2014) is available at <https://johannesbuchner.github.io/PyMultiNest/>. Retrieval outcomes for the 42 *HST* WFC3 spectra using BeAR are available online at <https://doi.org/10.5281/zenodo.1498277>.

REFERENCES

- Abel M., Frommhold L., Li X., Hunt K. L. C., 2011, *The Journal of Physical Chemistry A*, 115, 6805
 Abel M., Frommhold L., Li X., Hunt K. L. C., 2012, *J. Chem. Phys.*, 136, 044319
 Ahrer E.-M., et al., 2023, *Nature*, 614, 653
 Alderson L., et al., 2023, *Nature*, 614, 664
 Anderson D. R., et al., 2011, *A&A*, 531, A60
 Anderson D. R., et al., 2014, *MNRAS*, 445, 1114
 Anglada-Escudé G., Rojas-Ayala B., Boss A. P., Weinberger A. J., Lloyd J. P., 2013, *A&A*, 551, A48
 Bakos G. Á., et al., 2010, *ApJ*, 710, 1724
 Barstow J. K., 2020, *MNRAS*, 497, 4183
 Barstow J. K., Heng K., 2020, *Space Sci. Rev.*, 216, 82
 Barstow J. K., Aigrain S., Irwin P. G. J., Sing D. K., 2017, *ApJ*, 834, 50
 Batalha N. E., Kempton E. M. R., Mbarek R., 2017, *ApJ*, 836, L5
 Benneke B., Seager S., 2012, *ApJ*, 753, 100
 Benneke B., Seager S., 2013, *ApJ*, 778, 153
 Bétrémieux Y., 2016, *MNRAS*, 456, 4051
 Bétrémieux Y., Swain M. R., 2017, *MNRAS*, 467, 2834
 Bétrémieux Y., Swain M. R., 2018, *ApJ*, 865, 12
 Biddle L. I., et al., 2014, *MNRAS*, 443, 1810
 Boyajian T., et al., 2015, *MNRAS*, 447, 846
 von Braun K., et al., 2012, *ApJ*, 753, 171
 Brown T. M., 2001, *ApJ*, 553, 1006
 Buchner J., 2023, *Statistics Surveys*, 17, 169
 Buchner J., et al., 2014, *A&A*, 564, A125
 Chan T., Ingemyr M., Winn J. N., Holman M. J., Sanchis-Ojeda R., Esquerdo G., Everett M., 2011, *AJ*, 141, 179
 Cox A. N., 2000, *Allen's astrophysical quantities*. Springer
 Delrez L., et al., 2016, *MNRAS*, 458, 4025
 Esposito M., et al., 2014, *A&A*, 564, L13
 Evans T. M., et al., 2018, *AJ*, 156, 283
 Fairman C., Wakeford H. R., MacDonald R. J., 2024, *AJ*, 167, 240

- Feinstein A. D., et al., 2023, *Nature*, **614**, 670
- Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, **398**, 1601
- Fisher C., Heng K., 2018, *MNRAS*, **481**, 4698
- Fisher C., et al., 2024, *MNRAS*, **535**, 27
- Fu G., Deming D., Knutson H., Madhusudhan N., Mandell A., Fraine J., 2017, *ApJ*, **847**, L22
- Griffith C. A., 2014, *Philosophical Transactions of the Royal Society of London Series A*, **372**, 20130086
- Grimm S. L., Heng K., 2015, *ApJ*, **808**, 182
- Grimm S. L., et al., 2018, *A&A*, **613**, A68
- Grimm S. L., et al., 2021, *ApJS*, **253**, 30
- Hammond M., et al., 2024, *AJ*, **168**, 4
- Hartman J. D., et al., 2009, *ApJ*, **706**, 785
- Hartman J. D., et al., 2011a, *ApJ*, **728**, 138
- Hartman J. D., et al., 2011b, *ApJ*, **742**, 59
- Hartman J. D., et al., 2012, *AJ*, **144**, 139
- Hebb L., et al., 2009, *ApJ*, **693**, 1920
- Hébrard G., et al., 2013, *A&A*, **549**, A134
- Hellier C., et al., 2010, *ApJ*, **723**, L60
- Hellier C., et al., 2011, *A&A*, **535**, L7
- Hellier C., et al., 2012, *MNRAS*, **426**, 739
- Hellier C., et al., 2014, *MNRAS*, **440**, 1982
- Hellier C., et al., 2015, *AJ*, **150**, 18
- Heng K., Kitzmann D., 2017, *MNRAS*, **470**, 2972
- Howard A. W., et al., 2012, *ApJ*, **749**, 134
- Huitson C. M., et al., 2013, *MNRAS*, **434**, 3252
- Johnson J. A., et al., 2008, *ApJ*, **686**, 649
- Karman T., et al., 2019, *Icarus*, **328**, 160
- Kilpatrick B. M., et al., 2018, *AJ*, **156**, 103
- Kitzmann D., Heng K., 2018, *MNRAS*, **475**, 94
- Kitzmann D., Heng K., Oreshenko M., Grimm S. L., Apai D., Bowler B. P., Burgasser A. J., Marley M. S., 2020, *ApJ*, **890**, 174
- Knutson H. A., et al., 2014, *ApJ*, **794**, 155
- Kreidberg L., et al., 2014, *Nature*, **505**, 69
- Kreidberg L., et al., 2015, *ApJ*, **814**, 66
- Lecavelier Des Etangs A., Pont F., Vidal-Madjar A., Sing D., 2008, *A&A*, **481**, L83
- Line M. R., Parmentier V., 2016, *ApJ*, **820**, 78
- Lueber A., Kitzmann D., Bowler B. P., Burgasser A. J., Heng K., 2022, *ApJ*, **930**, 136
- Lueber A., Novais A., Fisher C., Heng K., 2024, *A&A*, **687**, A110
- MacDonald R. J., Madhusudhan N., 2017a, *MNRAS*, **469**, 1979
- MacDonald R. J., Madhusudhan N., 2017b, *ApJ*, **850**, L15
- MacDonald R. J., Goyal J. M., Lewis N. K., 2020, *ApJ*, **893**, L43
- Maciejewski G., et al., 2016, *Acta Astron.*, **66**, 55
- Madhusudhan N., Seager S., 2009, *ApJ*, **707**, 24
- Madhusudhan N., Sarkar S., Constantinou S., Holmberg M., Piette A. A. A., Moses J. I., 2023, *ApJ*, **956**, L13
- Mandell A. M., Haynes K., Sinukoff E., Madhusudhan N., Burrows A., Deming D., 2013, *ApJ*, **779**, 128
- Mikal-Evans T., et al., 2022, *Nature Astronomy*, **6**, 471
- Nikolov N., et al., 2014, *MNRAS*, **437**, 46
- Nixon M. C., Welbanks L., McGill P., Kempton E. M. R., 2024, *ApJ*, **966**, 156
- Pinhas A., Madhusudhan N., Gandhi S., MacDonald R., 2019, *MNRAS*, **482**, 1485
- Polyansky O. L., Kyuberis A. A., Zobov N. F., Tennyson J., Yurchenko S. N., Lodi L., 2018, *MNRAS*, **480**, 2597
- Rustamkulov Z., et al., 2023, *Nature*, **614**, 659
- Sato B., et al., 2012, *PASJ*, **64**, 97
- Schleich S., Bora Saikia S., Changeat Q., Güdel M., Voigt A., Waldmann I., 2024, *A&A*, **690**, A336
- Seager S., Sasselov D. D., 2000, *ApJ*, **537**, 916
- Sing D. K., et al., 2016, *Nature*, **529**, 59
- Skilling J., 2006, *Bayesian Analysis*, **1**, 833
- Sneep M., Ubachs W., 2005, *J. Quant. Spectrosc. Radiative Transfer*, **92**, 293
- Southworth J., et al., 2012, *MNRAS*, **426**, 1338
- Thalman R., Zarzana K. J., Tolbert M. A., Volkamer R., 2014, *J. Quant. Spectrosc. Radiative Transfer*, **147**, 171
- Torres G., Winn J. N., Holman M. J., 2008, *ApJ*, **677**, 1324
- Tregloan-Reed J., Southworth J., Tappert C., 2013, *MNRAS*, **428**, 3671
- Triaud A. H. M. J., et al., 2015, *MNRAS*, **450**, 2279
- Trotta R., 2008, *Contemporary Physics*, **49**, 71
- Tsai S.-M., et al., 2023, *Nature*, **617**, 483
- Tsiaras A., et al., 2018, *AJ*, **155**, 156
- Van Grootel V., et al., 2014, *ApJ*, **786**, 2
- Van Grootel V., et al., 2018, *ApJ*, **853**, 30
- Wakeford H. R., et al., 2018, *AJ*, **155**, 29
- Welbanks L., Madhusudhan N., 2019, *AJ*, **157**, 206
- West R. G., et al., 2016, *A&A*, **585**, A126
- de Wit J., Seager S., 2013, *Science*, **342**, 1473
- de Wit J., et al., 2018, *Nature Astronomy*, **2**, 214

APPENDIX A: SUPPLEMENTARY RESULTS

The following figures and tables display the retrieval results using BeAR cloud free (blue), grey cloud (yellow), non-grey cloud (red), and flat line (grey) models for each of the 38 *HST* WFC3 transmission spectra examined in this work. All retrievals were calculated using a spectral resolution of 0.1 cm^{-1} .

Table A1 lists the input values used in our retrievals. Figures A1 and A2 show the Bayesian comparison between the four (cloud free, grey clouds, non-grey clouds, and flat line) models for all objects. Tables A2 and A3 present the values for the free parameters retrieved by the cloud free, grey cloud, and non-grey cloud models for all objects. Tables A4 and A5 display the values for the free parameters retrieved by the cloud free, grey cloud, and non-grey cloud models for our case studies: HD 209458b, WASP-12b, and WASP-39b using non-isothermal models and tight temperature priors, respectively. Table A6 shows values for the free parameters retrieved by the cloud free and grey cloud models using data reduced by Sing et al. (2016).

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

Table A1. Input parameters used in our retrievals. References for R_p , R_{star} , and $\log g_p$ are listed in the “Reference” column. T_{eq} for all objects are from Fu et al. (2017).

Planet	T_{eq} (K)	R_p (R _{Jup})	R_{star} (R _⊕)	$\log g_p$ (cm s ⁻²)	Reference
GJ 436b	633 ± 58	0.369 ± 0.015	0.455 ± 0.018	3.120 ± 0.030	von Braun et al. (2012)
GJ 1214b	573 ± 35	0.243 ± 0.021	0.211 ± 0.011	2.885 ± 0.067	Anglada-Escudé et al. (2013)
GJ 3470b	692 ± 101	0.346 ± 0.029	0.480 ± 0.040	2.830 ± 0.110	Biddle et al. (2014)
HAT-P-1b	1320 ± 103	1.225 ± 0.059	1.1150 ± 0.050	2.873 ± 0.010	Johnson et al. (2008); Nikolov et al. (2014)
HAT-P-3b	1127 ± 68	0.827 ± 0.055	0.799 ± 0.039	3.330 ± 0.058	Chan et al. (2011)
HAT-P-11b	856 ± 37	0.422 ± 0.014	0.750 ± 0.020	3.050 ± 0.060	Bakos et al. (2010)
HAT-P-12b	958 ± 28	0.959 ^{+0.029} _{-0.021}	0.701 ^{+0.017} _{-0.012}	2.750 ± 0.030	Hartman et al. (2009)
HAT-P-17b	780 ± 34	1.010 ± 0.029	0.838 ± 0.021	3.110 ± 0.020	Howard et al. (2012)
HAT-P-18b	843 ± 35	0.947 ± 0.044	0.717 ± 0.026	2.734 ± 0.044	Esposito et al. (2014)
HAT-P-26b	980 ± 56	0.565 ^{+0.072} _{-0.032}	0.788 ^{+0.098} _{-0.043}	2.650 ± 0.090	Hartman et al. (2011a)
HAT-P-32b	1784 ± 58	1.789 ± 0.025	1.219 ± 0.016	2.820 ± 0.080	Hartman et al. (2011b)
HAT-P-38b	1080 ± 78	0.825 ^{+0.092} _{-0.063}	0.923 ^{+0.096} _{-0.067}	2.990 ± 0.080	Sato et al. (2012)
HAT-P-41b	1937 ± 74	1.685 ^{+0.076} _{-0.051}	1.683 ^{+0.058} _{-0.036}	2.840 ± 0.060	Hartman et al. (2012)
HD 97658b	753 ± 33	0.200 ^{+0.009} _{-0.008}	0.741 ^{+0.024} _{-0.023}	3.166 ± 0.060	Van Grootel et al. (2014)
HD 149026b	1627 ± 83	0.654 ^{+0.060} _{-0.045}	1.368 ^{+0.12} _{-0.083}	3.360 ± 0.066	Torres et al. (2008)
HD 189733b	1201 ± 51	1.216 ± 0.024	0.805 ± 0.016	3.290 ± 0.020	Boyajian et al. (2015)
HD 209458b	1449 ± 36	1.451 ± 0.074	1.203 ± 0.061	2.880 ± 0.070	Boyajian et al. (2015)
WASP-12b	2580 ± 146	1.790 ± 0.090	1.570 ± 0.070	2.990 ± 0.030	Hebb et al. (2009)
WASP-17b	1632 ± 126	1.932 ± 0.053	1.583 ± 0.041	2.500 ± 0.027	Southworth et al. (2012)
WASP-19b	2037 ± 156	1.395 ± 0.023	1.004 ± 0.016	3.152 ± 0.008	Tregloan-Reed et al. (2013)
WASP-29b	963 ± 69	0.792 ^{+0.056} _{-0.035}	0.808 ± 0.044	2.950 ± 0.050	Hellier et al. (2010)
WASP-31b	1576 ± 58	1.549 ± 0.050	1.252 ± 0.033	2.659 ± 0.036	Anderson et al. (2011)
WASP-39b	1119 ± 57	1.270 ± 0.040	0.918 ^{+0.022} _{-0.019}	2.617 ± 0.065	Maciejewski et al. (2016)
WASP-43b	1374 ± 147	1.060 ± 0.050	0.670 ± 0.040	3.672 ± 0.070	Hellier et al. (2011)
WASP-52b	1300 ± 115	1.270 ± 0.030	0.790 ± 0.020	2.810 ± 0.030	Hébrard et al. (2013)
WASP-63b	1508 ± 69	1.430 ^{+0.100} _{-0.060}	1.880 ^{+0.100} _{-0.060}	2.620 ± 0.050	Hellier et al. (2012)
WASP-67b	1026 ± 59	1.400 ^{+0.300} _{-0.200}	0.870 ± 0.040	2.700 ± 0.015	Hellier et al. (2012)
WASP-69b	964 ± 38	1.057 ± 0.047	0.813 ± 0.028	2.726 ± 0.046	Anderson et al. (2014)
WASP-74b	1915 ± 116	1.560 ± 0.060	1.640 ± 0.050	2.950 ± 0.020	Hellier et al. (2015)
WASP-76b	2206 ± 95	1.830 ^{+0.040} _{-0.060}	1.730 ± 0.040	2.800 ± 0.020	West et al. (2016)
WASP-80b	824 ± 58	0.999 ^{+0.030} _{-0.031}	0.586 ^{+0.017} _{-0.018}	3.145 ± 0.015	Triaud et al. (2015)
WASP-101b	1552 ± 81	1.410 ± 0.050	1.290 ± 0.040	2.760 ± 0.040	Hellier et al. (2014)
WASP-121b	2358 ± 122	1.865 ± 0.044	1.458 ± 0.030	2.973 ± 0.017	Delrez et al. (2016)
XO-1b	1196 ± 60	1.206 ^{+0.047} _{-0.042}	0.934 ^{+0.037} _{-0.032}	3.211 ± 0.040	Torres et al. (2008)
TRAPPIST-1d (N ₂)	288	0.070 ± 0.002	0.121 ± 0.003	2.676 ± 0.045	Van Grootel et al. (2018); Grimm et al. (2018)
TRAPPIST-1d (H ₂)	288	0.070 ± 0.002	0.121 ± 0.003	2.676 ± 0.045	Van Grootel et al. (2018); Grimm et al. (2018)
TRAPPIST-1e (N ₂)	251	0.081 ± 0.002	0.121 ± 0.003	2.960 ± 0.031	Van Grootel et al. (2018); Grimm et al. (2018)
TRAPPIST-1e (H ₂)	251	0.081 ± 0.002	0.121 ± 0.003	2.960 ± 0.031	Van Grootel et al. (2018); Grimm et al. (2018)
TRAPPIST-1f (N ₂)	219	0.093 ± 0.003	0.121 ± 0.003	2.923 ± 0.020	Van Grootel et al. (2018); Grimm et al. (2018)
TRAPPIST-1f (H ₂)	219	0.102 ± 0.003	0.121 ± 0.003	2.923 ± 0.020	Van Grootel et al. (2018); Grimm et al. (2018)
TRAPPIST-1g (N ₂)	199	0.102 ± 0.003	0.121 ± 0.003	2.931 ± 0.020	Van Grootel et al. (2018); Grimm et al. (2018)
TRAPPIST-1g (H ₂)	199	0.102 ± 0.003	0.121 ± 0.003	2.931 ± 0.020	Van Grootel et al. (2018); Grimm et al. (2018)

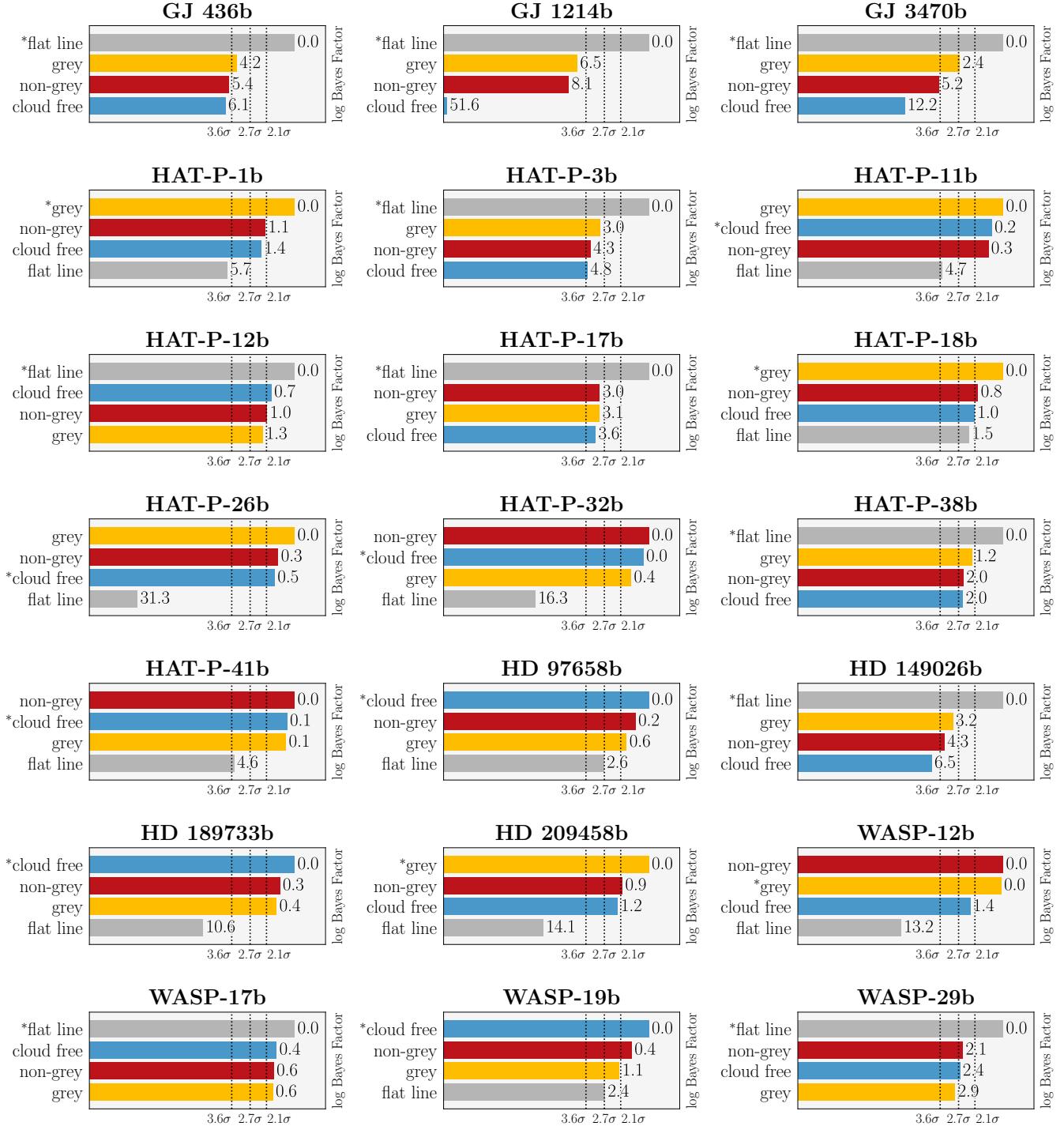


Figure A1. Bayesian comparison between cloud-free (blue), grey-cloud (yellow), non-grey cloud (red), and flat-line (grey) models. For each object, models are ordered from highest (top) to lowest (bottom) Bayesian evidence, and consequently from most favoured (highest evidence) to least favoured in comparison to the highest-evidence model. Values on the right side of each bar show the corresponding natural logarithm of the Bayes factor relative to the highest-evidence model. Vertical dotted lines show significances of 3.6σ , 2.7σ , and 2.1σ , corresponding to log Bayes factors of 5.0, 2.5, and 1.0, considered “strong”, “moderate”, and “weak” evidences compared to the highest-evidence model, according to [Trotta \(2008\)](#). Therefore, the simplest models with $\log \text{Bayes factor} < 1$ are considered favoured, and are denoted by *. All runs are calculated at a spectral resolution of 0.1 cm^{-1} .

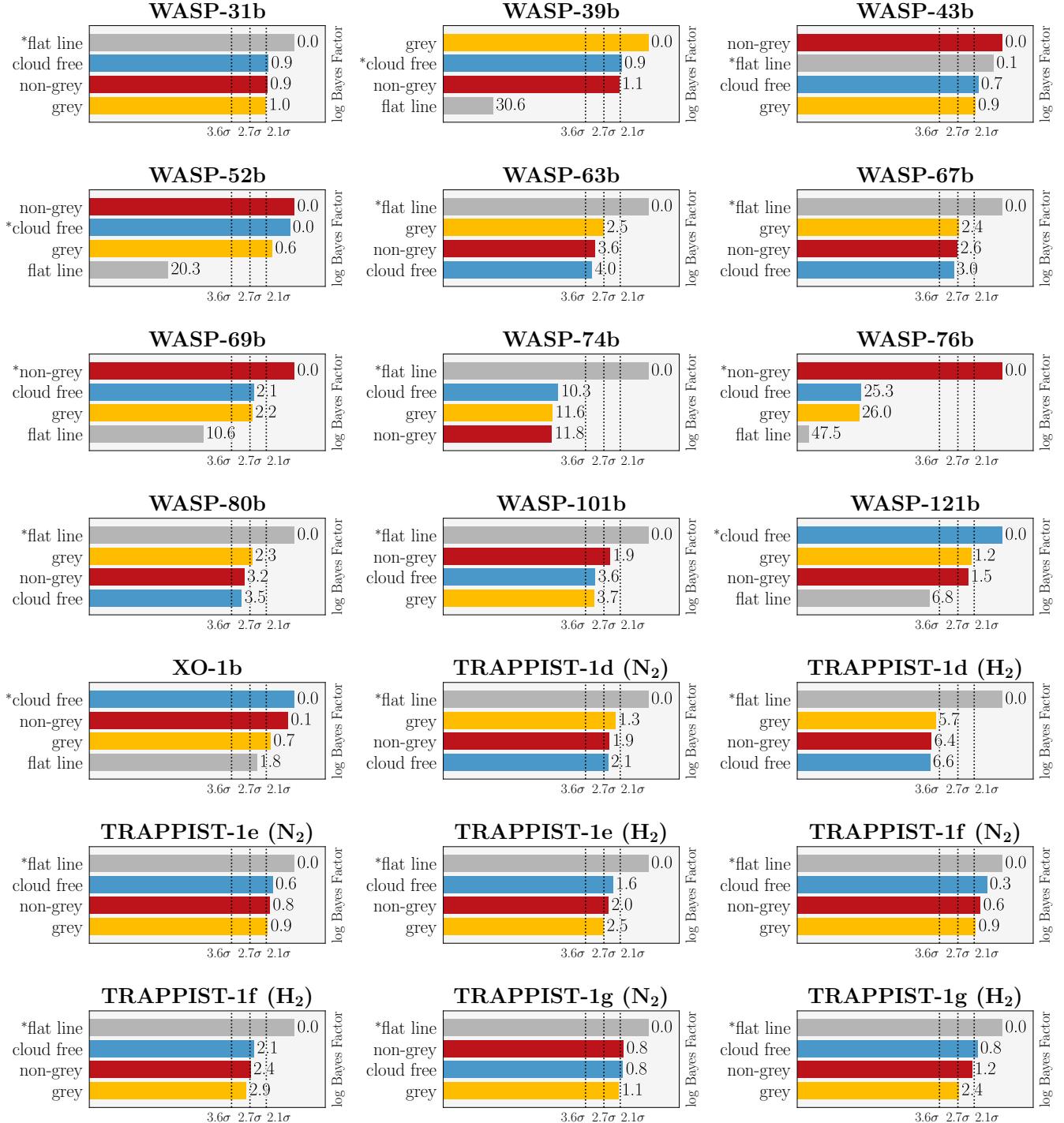


Figure A2. Continuation of Figure A1.

Table A2. Free parameters retrieved by our BeAR cloud free, grey cloud, and non-grey cloud models, for all objects. If the best-fit model for a planet is not a flat line, the model is denoted by * (see Figures A1 and A2). Values correspond to the median of the posterior distribution, and uncertainties correspond to the 16th and 84th quantiles of the distribution (i.e. $\pm 1\sigma$).

Planet	Model	$\log g_p$ (cm s $^{-2}$)	R_{\star} (R $_{\odot}$)	R_p (R $_{Jup}$)	$\log X_{H_2O}$	T (K)	$\log P_{\text{cloud-top}}$ (bar)
GJ 436b	cloud free	3.13 $^{+0.03}_{-0.03}$	0.459 $^{+0.011}_{-0.010}$	0.369 $^{+0.009}_{-0.008}$	-6.84 $^{+2.59}_{-0.70}$	285 $^{+62}_{-167}$	-
	grey clouds	3.12 $^{+0.03}_{-0.02}$	0.458 $^{+0.008}_{-0.011}$	0.363 $^{+0.005}_{-0.007}$	-8.42 $^{+2.16}_{-2.24}$	397 $^{+14}_{-211}$	-0.82 $^{+1.62}_{-0.58}$
	non-grey clouds	3.13 $^{+0.02}_{-0.02}$	0.460 $^{+0.009}_{-0.009}$	0.368 $^{+0.007}_{-0.007}$	-7.45 $^{+2.48}_{-1.19}$	343 $^{+98}_{-183}$	-0.82 $^{+1.38}_{-0.65}$
GJ 1214b	cloud free	3.29 $^{+0.03}_{-0.03}$	0.226 $^{+0.009}_{-0.004}$	0.255 $^{+0.010}_{-0.005}$	-5.86 $^{+0.08}_{-0.28}$	209 $^{+7}_{-5}$	-
	grey clouds	2.91 $^{+0.06}_{-0.05}$	0.215 $^{+0.005}_{-0.006}$	0.229 $^{+0.004}_{-0.006}$	-7.14 $^{+2.87}_{-2.00}$	286 $^{+55}_{-92}$	-1.87 $^{+1.09}_{-0.58}$
	non-grey clouds	2.91 $^{+0.04}_{-0.05}$	0.216 $^{+0.006}_{-0.005}$	0.230 $^{+0.005}_{-0.007}$	-7.43 $^{+2.35}_{-1.68}$	301 $^{+62}_{-101}$	-1.62 $^{+0.86}_{-0.44}$
GJ 3470b	cloud free	3.10 $^{+0.06}_{-0.05}$	0.489 $^{+0.020}_{-0.014}$	0.361 $^{+0.015}_{-0.010}$	-5.92 $^{+2.02}_{-1.40}$	811 $^{+554}_{-107}$	-
	grey clouds	2.88 $^{+0.07}_{-0.07}$	0.492 $^{+0.021}_{-0.020}$	0.342 $^{+0.018}_{-0.018}$	-4.49 $^{+1.47}_{-1.37}$	496 $^{+165}_{-227}$	-2.12 $^{+0.89}_{-0.82}$
	non-grey clouds	2.88 $^{+0.07}_{-0.07}$	0.490 $^{+0.021}_{-0.018}$	0.351 $^{+0.019}_{-0.012}$	-4.56 $^{+1.37}_{-1.33}$	412 $^{+101}_{-177}$	-1.66 $^{+0.65}_{-0.54}$
HAT-P-1b	cloud free	2.88 $^{+0.01}_{-0.01}$	1.105 $^{+0.023}_{-0.024}$	1.240 $^{+0.020}_{-0.025}$	-1.91 $^{+2.27}_{-0.79}$	532 $^{+132}_{-205}$	-
	*grey clouds	2.87 $^{+0.01}_{-0.01}$	1.125 $^{+0.023}_{-0.023}$	1.232 $^{+0.020}_{-0.030}$	-3.08 $^{+1.75}_{-1.75}$	905 $^{+200}_{-579}$	-1.69 $^{+1.11}_{-1.31}$
	non-grey clouds	2.87 $^{+0.01}_{-0.01}$	1.115 $^{+0.027}_{-0.019}$	1.249 $^{+0.030}_{-0.023}$	-2.53 $^{+2.04}_{-1.23}$	540 $^{+135}_{-305}$	-1.05 $^{+1.26}_{-1.20}$
HAT-P-3b	cloud free	3.35 $^{+0.06}_{-0.05}$	0.802 $^{+0.022}_{-0.017}$	0.855 $^{+0.024}_{-0.018}$	-7.34 $^{+3.11}_{-3.32}$	624 $^{+332}_{-591}$	-
	grey clouds	3.34 $^{+0.05}_{-0.05}$	0.803 $^{+0.025}_{-0.019}$	0.837 $^{+0.029}_{-0.025}$	-7.67 $^{+2.55}_{-2.54}$	1191 $^{+69}_{-1095}$	-2.06 $^{+1.11}_{-1.30}$
	non-grey clouds	3.34 $^{+0.05}_{-0.04}$	0.803 $^{+0.027}_{-0.016}$	0.851 $^{+0.031}_{-0.017}$	-7.66 $^{+2.55}_{-2.22}$	865 $^{+177}_{-594}$	-1.35 $^{+1.06}_{-1.22}$
HAT-P-11b	*cloud free	3.08 $^{+0.05}_{-0.05}$	0.752 $^{+0.014}_{-0.012}$	0.424 $^{+0.008}_{-0.007}$	-4.77 $^{+0.68}_{-1.91}$	669 $^{+302}_{-242}$	-
	grey clouds	3.08 $^{+0.05}_{-0.04}$	0.755 $^{+0.016}_{-0.012}$	0.419 $^{+0.009}_{-0.009}$	-4.31 $^{+1.01}_{-1.02}$	722 $^{+274}_{-295}$	-0.65 $^{+0.85}_{-0.79}$
	non-grey clouds	3.08 $^{+0.04}_{-0.05}$	0.752 $^{+0.012}_{-0.011}$	0.423 $^{+0.008}_{-0.007}$	-4.86 $^{+0.56}_{-1.28}$	705 $^{+276}_{-200}$	-0.90 $^{+1.80}_{-1.16}$
HAT-P-12b	cloud free	2.76 $^{+0.03}_{-0.03}$	0.718 $^{+0.004}_{-0.004}$	0.946 $^{+0.005}_{-0.005}$	-4.84 $^{+0.59}_{-0.69}$	308 $^{+75}_{-137}$	-
	grey clouds	2.76 $^{+0.03}_{-0.02}$	0.718 $^{+0.004}_{-0.005}$	0.944 $^{+0.006}_{-0.006}$	-4.68 $^{+0.64}_{-0.64}$	310 $^{+73}_{-118}$	-0.08 $^{+0.77}_{-0.64}$
	non-grey clouds	2.76 $^{+0.02}_{-0.02}$	0.717 $^{+0.004}_{-0.005}$	0.945 $^{+0.006}_{-0.007}$	-4.78 $^{+0.55}_{-0.58}$	312 $^{+71}_{-107}$	-0.93 $^{+1.66}_{-1.10}$
HAT-P-17b	cloud free	3.11 $^{+0.02}_{-0.02}$	0.850 $^{+0.009}_{-0.011}$	0.996 $^{+0.010}_{-0.014}$	-4.84 $^{+1.05}_{-1.51}$	370 $^{+120}_{-222}$	-
	grey clouds	3.11 $^{+0.02}_{-0.02}$	0.851 $^{+0.008}_{-0.011}$	0.994 $^{+0.008}_{-0.011}$	-4.97 $^{+3.35}_{-1.87}$	441 $^{+153}_{-286}$	-1.02 $^{+1.23}_{-0.96}$
	non-grey clouds	3.11 $^{+0.02}_{-0.02}$	0.849 $^{+0.010}_{-0.010}$	0.993 $^{+0.010}_{-0.010}$	-4.90 $^{+1.47}_{-1.37}$	405 $^{+152}_{-193}$	-1.38 $^{+1.22}_{-1.12}$
HAT-P-18b	cloud free	2.77 $^{+0.04}_{-0.04}$	0.726 $^{+0.019}_{-0.016}$	0.957 $^{+0.021}_{-0.021}$	-1.93 $^{+2.06}_{-0.81}$	232 $^{+22}_{-49}$	-
	* grey clouds	2.75 $^{+0.04}_{-0.04}$	0.719 $^{+0.019}_{-0.015}$	0.933 $^{+0.027}_{-0.029}$	-2.82 $^{+2.15}_{-1.15}$	396 $^{+141}_{-163}$	-1.56 $^{+0.94}_{-0.72}$
	non-grey cloud	2.76 $^{+0.03}_{-0.03}$	0.720 $^{+0.013}_{-0.014}$	0.947 $^{+0.019}_{-0.019}$	-3.15 $^{+1.25}_{-1.77}$	249 $^{+45}_{-140}$	-1.21 $^{+0.20}_{-1.10}$
HAT-P-26b	*cloud free	2.73 $^{+0.07}_{-0.06}$	0.846 $^{+0.029}_{-0.024}$	0.561 $^{+0.019}_{-0.021}$	-3.48 $^{+0.77}_{-0.71}$	520 $^{+108}_{-142}$	-
	grey clouds	2.70 $^{+0.06}_{-0.06}$	0.856 $^{+0.029}_{-0.033}$	0.554 $^{+0.017}_{-0.023}$	-3.27 $^{+0.88}_{-1.09}$	587 $^{+128}_{-140}$	-0.82 $^{+0.77}_{-0.90}$
	non-grey clouds	2.72 $^{+0.06}_{-0.05}$	0.841 $^{+0.023}_{-0.030}$	0.557 $^{+0.015}_{-0.020}$	-3.42 $^{+0.66}_{-0.90}$	525 $^{+107}_{-109}$	-0.91 $^{+1.56}_{-1.12}$
HAT-P-32b	*cloud free	2.87 $^{+0.08}_{-0.06}$	1.222 $^{+0.009}_{-0.010}$	1.784 $^{+0.013}_{-0.016}$	-2.86 $^{+1.28}_{-1.28}$	610 $^{+132}_{-143}$	-
	grey clouds	2.86 $^{+0.07}_{-0.06}$	1.226 $^{+0.009}_{-0.010}$	1.782 $^{+0.012}_{-0.014}$	-2.88 $^{+1.16}_{-1.16}$	642 $^{+144}_{-150}$	-0.55 $^{+1.12}_{-1.05}$
	non-grey clouds	2.86 $^{+0.06}_{-0.06}$	1.221 $^{+0.009}_{-0.009}$	1.782 $^{+0.010}_{-0.015}$	-3.10 $^{+1.09}_{-1.26}$	615 $^{+150}_{-151}$	-0.95 $^{+1.50}_{-1.15}$
HAT-P-38b	cloud free	3.03 $^{+0.07}_{-0.06}$	0.942 $^{+0.051}_{-0.059}$	0.836 $^{+0.046}_{-0.042}$	-5.83 $^{+0.52}_{-0.52}$	1199 $^{+376}_{-739}$	-
	grey clouds	3.02 $^{+0.06}_{-0.06}$	0.952 $^{+0.059}_{-0.041}$	0.808 $^{+0.042}_{-0.042}$	-5.20 $^{+0.52}_{-0.51}$	1539 $^{+713}_{-875}$	-1.42 $^{+1.29}_{-1.01}$
	non-grey clouds	3.03 $^{+0.06}_{-0.06}$	0.939 $^{+0.043}_{-0.041}$	0.831 $^{+0.038}_{-0.037}$	-5.79 $^{+0.52}_{-0.72}$	1204 $^{+367}_{-297}$	-1.00 $^{+1.73}_{-1.21}$
HAT-P-41b	*cloud free	2.87 $^{+0.06}_{-0.05}$	1.720 $^{+0.019}_{-0.019}$	1.658 $^{+0.016}_{-0.026}$	-4.61 $^{+0.80}_{-2.03}$	1064 $^{+293}_{-358}$	-
	grey clouds	2.87 $^{+0.05}_{-0.04}$	1.725 $^{+0.021}_{-0.022}$	1.654 $^{+0.013}_{-0.018}$	-4.20 $^{+0.95}_{-1.72}$	1133 $^{+299}_{-329}$	-0.67 $^{+1.03}_{-1.00}$
	non-grey clouds	2.87 $^{+0.05}_{-0.04}$	1.717 $^{+0.015}_{-0.020}$	1.656 $^{+0.013}_{-0.019}$	-4.73 $^{+0.68}_{-1.27}$	1139 $^{+357}_{-290}$	-0.74 $^{+1.63}_{-1.00}$
HD 97658b	*cloud free	3.21 $^{+0.05}_{-0.05}$	0.736 $^{+0.015}_{-0.012}$	0.205 $^{+0.004}_{-0.005}$	-9.26 $^{+1.83}_{-1.99}$	1625 $^{+217}_{-231}$	-
	grey clouds	3.20 $^{+0.05}_{-0.04}$	0.744 $^{+0.015}_{-0.016}$	0.202 $^{+0.004}_{-0.004}$	-9.08 $^{+1.75}_{-1.75}$	1849 $^{+257}_{-447}$	-0.01 $^{+0.22}_{-0.58}$
	non-grey clouds	3.20 $^{+0.05}_{-0.04}$	0.737 $^{+0.013}_{-0.011}$	0.205 $^{+0.004}_{-0.003}$	-9.09 $^{+1.68}_{-1.61}$	1635 $^{+213}_{-205}$	-0.89 $^{+1.87}_{-1.29}$
HD 149026b	cloud free	3.38 $^{+0.06}_{-0.05}$	1.447 $^{+0.010}_{-0.010}$	0.710 $^{+0.003}_{-0.003}$	-4.09 $^{+2.23}_{-2.44}$	815 $^{+227}_{-546}$	-
	grey clouds	3.35 $^{+0.05}_{-0.05}$	1.518 $^{+0.028}_{-0.029}$	0.701 $^{+0.014}_{-0.016}$	-4.59 $^{+4.62}_{-4.62}$	2424 $^{+531}_{-392}$	-3.09 $^{+0.56}_{-0.76}$
	non-grey clouds	3.35 $^{+0.05}_{-0.05}$	1.515 $^{+0.026}_{-0.026}$	0.699 $^{+0.013}_{-0.009}$	-4.00 $^{+3.37}_{-3.99}$	2412 $^{+482}_{-395}$	-3.05 $^{+0.56}_{-0.80}$
HD 189733b	*cloud free	3.29 $^{+0.02}_{-0.02}$	0.806 $^{+0.010}_{-0.009}$	1.215 $^{+0.015}_{-0.015}$	-2.58 $^{+1.07}_{-0.66}$	504 $^{+90}_{-113}$	-
	grey clouds	3.29 $^{+0.02}_{-0.02}$	0.806 $^{+0.009}_{-0.009}$	1.215 $^{+0.013}_{-0.013}$	-2.43 $^{+1.04}_{-0.73}$	518 $^{+89}_{-127}$	-0.01 $^{+0.83}_{-0.66}$
	non-grey clouds	3.29 $^{+0.02}_{-0.02}$	0.806 $^{+0.008}_{-0.008}$	1.215 $^{+0.013}_{-0.013}$	-2.61 $^{+1.03}_{-0.79}$	507 $^{+81}_{-127}$	-0.90 $^{+1.66}_{-1.20}$
HD 209458b	cloud free	2.93 $^{+0.05}_{-0.05}$	1.226 $^{+0.027}_{-0.024}$	1.424 $^{+0.030}_{-0.030}$	-5.26 $^{+0.31}_{-0.66}$	752 $^{+219}_{-320}$	-
	* grey clouds	2.90 $^{+0.05}_{-0.05}$	1.224 $^{+0.024}_{-0.024}$	1.410 $^{+0.025}_{-0.025}$	-4.25 $^{+2.16}_{-1.36}$	717 $^{+215}_{-315}$	-1.19 $^{+1.06}_{-0.97}$
	non-grey clouds	2.93 $^{+0.05}_{-0.05}$	1.216 $^{+0.018}_{-0.024}$	1.413 $^{+0.022}_{-0.026}$	-5.20 $^{+0.34}_{-0.70}$	743 $^{+189}_{-219}$	-0.74 $^{+0.97}_{-1.10}$
WASP-12b	cloud free	2.99 $^{+0.03}_{-0.03}$	1.570 $^{+0.045}_{-0.036}$	1.814 $^{+0.053}_{-0.042}$	-4.54 $^{+0.53}_{-0.79}$	964 $^{+192}_{-206}$	-
	* grey clouds	2.99 $^{+0.03}_{-0.02}$	1.572 $^{+0.037}_{-0.038}$	1.793 $^{+0.046}_{-0.047}$	-3.85 $^{+0.57}_{-0.66}$	1251 $^{+296}_{-313}$	-1.40 $^{+1.43}_{-0.87}$
	non-grey clouds	2.99 $^{+0.03}_{-0.03}$	1.577 $^{+0.031}_{-0.031}$	1.813 $^{+0.043}_{-0.034}$	-4.39 $^{+0.57}_{-1.07}$	1149 $^{+239}_{-245}$	-1.36 $^{+1.18}_{-0.76}$
WASP-17b	cloud free	2.51 $^{+0.02}_{-0.02}$	1.629 $^{+0.019}_{-0.018}$	1.904 $^{+0.027}_{-0.027}$	-3.51 $^{+1.44}_{-1.75}$	508 $^{+158}_{-299}$	-
	grey clouds	2.50 $^{+0.02}_{-0.02}$	1.628 $^{+0.019}_{-0.019}$	1.899 $^{+0.027}_{-0.027}$	-3.41 $^{+1.37}_{-1.34}$	509 $^{+149}_{-211}$	-0.60 $^{+0.86}_{-0.90}$
	non-grey clouds	2.50 $^{+0.02}_{-0.02}$	1.626 $^{+0.018}_{-0.017}$	1.901 $^{+0.024}_{-0.019}$	-3.57 $^{+1.21}_{-1.39}$	506 $^{+165}_{-255}$	-0.92 $^{+1.53}_{-1.12}$
WASP-19b	*cloud free	3.15 $^{+0.01}_{-0.01}$	1.028 $^{+0.006}_{-0.007}$	1.379 $^{+0.004}_{-0.005}$	-5.04 $^{+0.60}_{-0.66}$	1828 $^{+479}_{-541}$	-
	grey clouds	3.15 $^{+0.01}_{-0.01}$	1.027 $^{+0.006}_{-0.006}$	1.379 $^{+0.005}_{-0.007}$	-4.89 $^{+0.63}_{-0.96}$	1712 $^{+450}_{-422}$	0.21 $^{+0.58}_{-0.47}$
	non-grey clouds	3.15 $^{+0.01}_{-0.01}$	1.027 $^{+0.006}_{-0.006}$	1.379 $^{+0.005}_{-0.007}$	-4.92 $^{+0.59}_{-0.81}$	1724 $^{+446}_{-419}$	-0.86 $^{+1.84}_{-1.20}$
WASP-29b	cloud free	2.95 $^{+0.05}_{-0.05}$	0.827 $^{+0.018}_{-0.025}$	0.781 $^{+0.017}_{-0.023}$	-8.54 $^{+2.25}_{-0.79}$	364 $^{+117}_{-181}$	-
	grey clouds	2.95 $^{+0.04}_{-0.04}$	0.826 $^{+0.016}_{-0.017}</$				

Table A3. Continuation of Table A2.

Planet	Model	$\log g_p$ (cm s $^{-2}$)	R_{\star} (R $_{\odot}$)	R_p (R $_{Jup}$)	$\log X_{H_2O}$	T (K)	$\log P_{\text{cloud-top}}$ (bar)
WASP-31b	cloud free	2.66 $^{+0.03}_{-0.03}$	1.266 $^{+0.017}_{-0.010}$	1.527 $^{+0.019}_{-0.026}$	-4.42 $^{+0.98}_{-1.19}$	454 $^{+175}_{-306}$	-
	grey clouds	2.66 $^{+0.03}_{-0.03}$	1.270 $^{+0.018}_{-0.010}$	1.523 $^{+0.016}_{-0.025}$	-3.87 $^{+1.29}_{-1.47}$	457 $^{+132}_{-261}$	-0.58 $^{+1.18}_{-0.94}$
	non-grey clouds	2.66 $^{+0.03}_{-0.03}$	1.265 $^{+0.014}_{-0.019}$	1.525 $^{+0.016}_{-0.024}$	-4.49 $^{+0.89}_{-1.28}$	484 $^{+183}_{-263}$	-1.00 $^{+1.65}_{-1.25}$
WASP-39b	*cloud free	2.70 $^{+0.05}_{-0.04}$	0.917 $^{+0.020}_{-0.019}$	1.274 $^{+0.028}_{-0.025}$	-1.16 $^{+0.58}_{-0.71}$	408 $^{+89}_{-127}$	-
	grey clouds	2.67 $^{+0.05}_{-0.04}$	0.929 $^{+0.022}_{-0.019}$	1.275 $^{+0.027}_{-0.023}$	-1.42 $^{+1.56}_{-1.19}$	493 $^{+132}_{-194}$	-1.21 $^{+1.32}_{-1.43}$
	non-grey clouds	2.68 $^{+0.04}_{-0.04}$	0.917 $^{+0.019}_{-0.015}$	1.275 $^{+0.026}_{-0.021}$	-1.38 $^{+1.19}_{-0.79}$	361 $^{+60}_{-70}$	-0.97 $^{+1.63}_{-1.18}$
WASP-43b	cloud free	3.68 $^{+0.06}_{-0.06}$	0.675 $^{+0.015}_{-0.022}$	1.045 $^{+0.024}_{-0.034}$	-3.09 $^{+1.27}_{-1.33}$	471 $^{+125}_{-153}$	-
	grey clouds	3.67 $^{+0.06}_{-0.06}$	0.673 $^{+0.017}_{-0.020}$	1.041 $^{+0.018}_{-0.026}$	-3.00 $^{+1.35}_{-1.19}$	518 $^{+137}_{-131}$	-0.28 $^{+1.08}_{-0.80}$
	non-grey clouds	3.68 $^{+0.06}_{-0.05}$	0.673 $^{+0.017}_{-0.017}$	1.041 $^{+0.024}_{-0.026}$	-3.28 $^{+1.22}_{-1.22}$	510 $^{+135}_{-193}$	-1.62 $^{+1.34}_{-1.57}$
WASP-52b	*cloud free	2.80 $^{+0.03}_{-0.03}$	0.796 $^{+0.012}_{-0.012}$	1.259 $^{+0.013}_{-0.020}$	-3.16 $^{+0.92}_{-0.93}$	397 $^{+73}_{-115}$	-
	grey clouds	2.80 $^{+0.03}_{-0.03}$	0.796 $^{+0.008}_{-0.012}$	1.258 $^{+0.014}_{-0.020}$	-3.03 $^{+0.90}_{-0.92}$	396 $^{+66}_{-104}$	0.12 $^{+0.65}_{-0.55}$
	non-grey clouds	2.81 $^{+0.03}_{-0.03}$	0.794 $^{+0.006}_{-0.009}$	1.256 $^{+0.010}_{-0.015}$	-3.19 $^{+0.88}_{-0.87}$	402 $^{+71}_{-109}$	-0.89 $^{+1.76}_{-1.19}$
WASP-63b	cloud free	2.64 $^{+0.05}_{-0.04}$	1.896 $^{+0.045}_{-0.052}$	1.433 $^{+0.036}_{-0.040}$	-4.38 $^{+1.22}_{-1.36}$	301 $^{+72}_{-238}$	-
	grey clouds	2.63 $^{+0.04}_{-0.04}$	1.911 $^{+0.041}_{-0.046}$	1.415 $^{+0.034}_{-0.034}$	-4.40 $^{+2.63}_{-1.62}$	485 $^{+164}_{-236}$	-1.77 $^{+1.08}_{-0.85}$
	non-grey clouds	2.64 $^{+0.04}_{-0.04}$	1.891 $^{+0.039}_{-0.045}$	1.418 $^{+0.027}_{-0.038}$	-4.49 $^{+1.22}_{-1.38}$	378 $^{+131}_{-216}$	-1.38 $^{+1.19}_{-1.16}$
WASP-67b	cloud free	2.70 $^{+0.02}_{-0.01}$	0.873 $^{+0.028}_{-0.029}$	1.367 $^{+0.044}_{-0.047}$	-5.33 $^{+0.67}_{-0.71}$	536 $^{+207}_{-228}$	-
	grey clouds	2.70 $^{+0.02}_{-0.01}$	0.871 $^{+0.026}_{-0.027}$	1.344 $^{+0.054}_{-0.057}$	-5.11 $^{+1.52}_{-1.59}$	612 $^{+249}_{-249}$	-0.93 $^{+1.69}_{-1.22}$
	non-grey clouds	2.70 $^{+0.02}_{-0.01}$	0.872 $^{+0.026}_{-0.024}$	1.363 $^{+0.041}_{-0.039}$	-5.29 $^{+0.70}_{-0.83}$	548 $^{+180}_{-217}$	-1.07 $^{+1.63}_{-1.24}$
WASP-69b	cloud free	2.76 $^{+0.04}_{-0.04}$	0.831 $^{+0.012}_{-0.016}$	1.032 $^{+0.015}_{-0.020}$	-4.41 $^{+0.37}_{-0.39}$	228 $^{+21}_{-48}$	-
	grey clouds	2.75 $^{+0.04}_{-0.04}$	0.829 $^{+0.013}_{-0.013}$	1.026 $^{+0.016}_{-0.015}$	-4.49 $^{+0.41}_{-0.47}$	250 $^{+35}_{-38}$	-0.26 $^{+0.41}_{-0.66}$
	*non-grey clouds	2.74 $^{+0.04}_{-0.04}$	0.830 $^{+0.007}_{-0.008}$	1.022 $^{+0.007}_{-0.008}$	-3.70 $^{+0.84}_{-0.94}$	282 $^{+51}_{-89}$	-1.64 $^{+0.49}_{-0.52}$
WASP-74b	cloud free	2.95 $^{+0.01}_{-0.02}$	1.626 $^{+0.005}_{-0.006}$	1.505 $^{+0.004}_{-0.006}$	-4.59 $^{+0.92}_{-0.92}$	386 $^{+121}_{-182}$	-
	grey clouds	2.96 $^{+0.02}_{-0.01}$	1.625 $^{+0.004}_{-0.005}$	1.505 $^{+0.003}_{-0.005}$	-4.80 $^{+1.47}_{-1.21}$	391 $^{+142}_{-170}$	-0.13 $^{+0.76}_{-0.65}$
	non-grey clouds	2.95 $^{+0.01}_{-0.01}$	1.624 $^{+0.004}_{-0.005}$	1.505 $^{+0.003}_{-0.004}$	-4.76 $^{+1.00}_{-0.94}$	394 $^{+112}_{-156}$	-1.17 $^{+1.45}_{-1.19}$
WASP-76b	cloud free	2.80 $^{+0.02}_{-0.02}$	1.800 $^{+0.023}_{-0.035}$	1.804 $^{+0.023}_{-0.035}$	-2.03 $^{+1.16}_{-1.61}$	416 $^{+57}_{-62}$	-
	grey clouds	2.80 $^{+0.02}_{-0.02}$	1.794 $^{+0.018}_{-0.030}$	1.797 $^{+0.018}_{-0.021}$	-1.94 $^{+0.61}_{-0.64}$	422 $^{+59}_{-89}$	0.22 $^{+0.55}_{-0.49}$
	*non-grey clouds	2.81 $^{+0.02}_{-0.02}$	1.841 $^{+0.016}_{-0.014}$	1.786 $^{+0.010}_{-0.015}$	-4.45 $^{+0.39}_{-0.65}$	1313 $^{+223}_{-223}$	-2.53 $^{+0.55}_{-0.34}$
WASP-80b	cloud free	3.15 $^{+0.01}_{-0.01}$	0.594 $^{+0.007}_{-0.009}$	0.985 $^{+0.011}_{-0.015}$	-5.28 $^{+0.51}_{-0.56}$	425 $^{+163}_{-163}$	-
	grey clouds	3.14 $^{+0.01}_{-0.01}$	0.594 $^{+0.005}_{-0.008}$	0.979 $^{+0.007}_{-0.011}$	-4.46 $^{+0.61}_{-0.67}$	468 $^{+170}_{-196}$	-1.22 $^{+1.07}_{-0.74}$
	non-grey clouds	3.15 $^{+0.01}_{-0.01}$	0.591 $^{+0.004}_{-0.006}$	0.980 $^{+0.007}_{-0.010}$	-5.23 $^{+0.53}_{-0.64}$	429 $^{+146}_{-197}$	-0.72 $^{+1.54}_{-0.92}$
WASP-101b	cloud free	2.76 $^{+0.04}_{-0.04}$	1.315 $^{+0.014}_{-0.021}$	1.382 $^{+0.015}_{-0.022}$	-5.53 $^{+1.70}_{-1.77}$	275 $^{+52}_{-151}$	-
	grey clouds	2.76 $^{+0.04}_{-0.03}$	1.318 $^{+0.015}_{-0.021}$	1.379 $^{+0.012}_{-0.019}$	-5.88 $^{+2.17}_{-1.71}$	324 $^{+89}_{-206}$	-0.32 $^{+1.57}_{-0.80}$
	non-grey clouds	2.77 $^{+0.03}_{-0.03}$	1.323 $^{+0.014}_{-0.016}$	1.378 $^{+0.012}_{-0.016}$	-5.45 $^{+3.05}_{-1.73}$	344 $^{+87}_{-165}$	-2.35 $^{+0.81}_{-0.88}$
WASP-121b	*cloud free	2.98 $^{+0.02}_{-0.01}$	1.569 $^{+0.005}_{-0.007}$	1.826 $^{+0.004}_{-0.008}$	-3.05 $^{+0.78}_{-0.89}$	717 $^{+108}_{-108}$	-
	grey clouds	2.97 $^{+0.02}_{-0.01}$	1.569 $^{+0.004}_{-0.005}$	1.826 $^{+0.003}_{-0.005}$	-3.03 $^{+0.77}_{-0.80}$	700 $^{+121}_{-121}$	-0.00 $^{+0.58}_{-0.51}$
	non-grey clouds	2.97 $^{+0.02}_{-0.01}$	1.568 $^{+0.004}_{-0.004}$	1.825 $^{+0.003}_{-0.004}$	-3.04 $^{+0.69}_{-0.70}$	692 $^{+111}_{-106}$	-0.86 $^{+1.43}_{-1.00}$
XO-1b	*cloud free	3.22 $^{+0.04}_{-0.03}$	0.932 $^{+0.009}_{-0.011}$	1.183 $^{+0.012}_{-0.014}$	-2.34 $^{+1.51}_{-1.49}$	671 $^{+184}_{-171}$	-
	grey clouds	3.22 $^{+0.03}_{-0.03}$	0.932 $^{+0.008}_{-0.009}$	1.181 $^{+0.011}_{-0.012}$	-2.35 $^{+1.54}_{-1.54}$	686 $^{+172}_{-195}$	-0.39 $^{+1.03}_{-0.79}$
	non-grey clouds	3.22 $^{+0.03}_{-0.03}$	0.931 $^{+0.008}_{-0.009}$	1.181 $^{+0.010}_{-0.011}$	-2.93 $^{+1.29}_{-1.18}$	709 $^{+178}_{-234}$	-1.54 $^{+1.45}_{-1.54}$
TRAPPIST-1d (N ₂)	cloud free	2.68 $^{+0.04}_{-0.04}$	0.117 $^{+0.002}_{-0.002}$	0.071 $^{+0.001}_{-0.001}$	-6.87 $^{+3.08}_{-3.60}$	1005 $^{+521}_{-1250}$	-
	grey clouds	2.68 $^{+0.04}_{-0.04}$	0.121 $^{+0.002}_{-0.003}$	0.070 $^{+0.001}_{-0.001}$	-7.37 $^{+3.70}_{-3.31}$	895 $^{+400}_{-912}$	-0.97 $^{+1.68}_{-1.09}$
	non-grey clouds	2.68 $^{+0.04}_{-0.04}$	0.118 $^{+0.002}_{-0.003}$	0.071 $^{+0.001}_{-0.001}$	-6.66 $^{+3.08}_{-3.13}$	865 $^{+395}_{-855}$	-1.34 $^{+1.42}_{-1.32}$
TRAPPIST-1d (H ₂)	cloud free	2.73 $^{+0.04}_{-0.04}$	0.123 $^{+0.002}_{-0.002}$	0.069 $^{+0.001}_{-0.001}$	-6.20 $^{+0.81}_{-0.61}$	219 $^{+14}_{-34}$	-
	grey clouds	2.72 $^{+0.04}_{-0.04}$	0.125 $^{+0.002}_{-0.002}$	0.069 $^{+0.001}_{-0.001}$	-7.89 $^{+2.07}_{-2.00}$	217 $^{+26}_{-26}$	-0.30 $^{+0.29}_{-0.30}$
	non-grey clouds	2.72 $^{+0.04}_{-0.04}$	0.124 $^{+0.002}_{-0.002}$	0.069 $^{+0.001}_{-0.001}$	-6.52 $^{+3.03}_{-0.82}$	218 $^{+13}_{-25}$	-0.45 $^{+1.51}_{-0.53}$
TRAPPIST-1e (N ₂)	cloud free	2.96 $^{+0.03}_{-0.03}$	0.119 $^{+0.002}_{-0.001}$	0.082 $^{+0.001}_{-0.001}$	-6.49 $^{+3.74}_{-3.58}$	1640 $^{+943}_{-977}$	-
	grey clouds	2.96 $^{+0.03}_{-0.03}$	0.121 $^{+0.002}_{-0.002}$	0.081 $^{+0.001}_{-0.001}$	-6.63 $^{+3.50}_{-3.50}$	1105 $^{+577}_{-1033}$	-0.56 $^{+1.72}_{-1.00}$
	non-grey clouds	2.96 $^{+0.03}_{-0.03}$	0.120 $^{+0.002}_{-0.001}$	0.082 $^{+0.001}_{-0.001}$	-6.68 $^{+3.27}_{-3.61}$	1553 $^{+863}_{-933}$	-1.20 $^{+1.67}_{-1.40}$
TRAPPIST-1e (H ₂)	cloud free	2.97 $^{+0.03}_{-0.03}$	0.121 $^{+0.002}_{-0.002}$	0.081 $^{+0.001}_{-0.001}$	-6.00 $^{+4.10}_{-4.00}$	343 $^{+106}_{-167}$	-
	grey clouds	2.96 $^{+0.03}_{-0.03}$	0.123 $^{+0.002}_{-0.002}$	0.080 $^{+0.001}_{-0.001}$	-6.15 $^{+3.85}_{-3.85}$	316 $^{+86}_{-86}$	0.43 $^{+0.54}_{-0.36}$
	non-grey clouds	2.97 $^{+0.03}_{-0.03}$	0.121 $^{+0.002}_{-0.002}$	0.081 $^{+0.001}_{-0.001}$	-6.27 $^{+3.66}_{-3.68}$	328 $^{+149}_{-157}$	-1.25 $^{+0.95}_{-1.67}$
TRAPPIST-1f (N ₂)	cloud free	2.92 $^{+0.02}_{-0.02}$	0.121 $^{+0.002}_{-0.002}$	0.093 $^{+0.002}_{-0.002}$	-6.54 $^{+3.54}_{-3.84}$	1625 $^{+942}_{-963}$	-
	grey clouds	2.92 $^{+0.02}_{-0.02}$	0.122 $^{+0.002}_{-0.002}$	0.092 $^{+0.002}_{-0.002}$	-6.71 $^{+3.85}_{-3.85}$	981 $^{+519}_{-1093}$	-0.49 $^{+1.62}_{-0.98}$
	non-grey clouds	2.92 $^{+0.02}_{-0.02}$	0.121 $^{+0.002}_{-0.002}$	0.093 $^{+0.002}_{-0.002}$	-6.48 $^{+3.36}_{-3.43}$	1570 $^{+902}_{-967}$	-1.12 $^{+1.79}_{-1.38}$
TRAPPIST-1f (H ₂)	cloud free	2.93 $^{+0.02}_{-0.02}$	0.122 $^{+0.002}_{-0.002}$	0.092 $^{+0.002}_{-0.002}$	-5.86 $^{+4.22}_{-3.97}$	329 $^{+93}_{-159}$	-
	grey clouds	2.93 $^{+0.02}_{-0.02}$	0.123 $^{+0.002}_{-0.002}$	0.093 $^{+0.001}_{-0.002}$	-6.12 $^{+3.87}_{-3.87}$	304 $^{+138}_{-138}$	0.39 $^{+0.51}_{-0.38}$
	non-grey clouds	2.93 $^{+0.02}_{-0.02}$	0.122 $^{+0.002}_{-0.002}$	0.092 $^{+0.001}_{-0.002}$	-6.00 $^{+3.92}_{-3.70}$	316 $^{+84}_{-155}$	-1.16 $^{+1.82}_{-1.56}$
TRAPPIST-1g (N ₂)	cloud free	2.93 $^{+0.02}_{-0.02}$	0.120 $^{+0.002}_{-0.002}$	0.103 $^{+0.002}_{-0.001}$	-6.76 $^{+3.42}_{-3.58}$	1689 $^{+1010}_{-953}$	-
	grey clouds	2.93 $^{+0.02}_{-0.02}$	0.122 $^{+0.002}_{-0.002}$	0.102 $^{+0.002}_{-0.002}$	-6.47 $^{+3.29}_{-3.29}$	1125 $^{+380}_{-907}$	-0.55 $^{+1.84}_{-0.99}$
	non-grey clouds	2.93 $^{+0.02}_{-0.02}$	0.120 $^{+0.002}_{-0.002}$	0.103 $^{+0.002}_{-0.001}$	-6.27 $^{+3.45}_{-3.40}$	1550 $^{+866}_{-932}$	-1.13 $^{+1.72}_{-1.34}$
TRAPPIST-1g (H ₂)	cloud free	2.93 $^{+0.02}_{-0.02}$	0.122 $^{+0.002}_{-0.002}$	0.102 $^{+0.002}_{-0.002}$	-6.85 $^{+3.45}_{-3.75}$	422 $^{+152}_{-210}$	-
	grey clouds	2.93 $^{+0.02}_{-0.02}$	0.123 $^{+0.002}_{-0.002}$	0.101 $^{+0.001$			

Table A4. Free parameters retrieved using cloud-free and grey-cloud non-isothermal models (see Section 5.4.1).

Planet	Model	$\log g_p$ (cm s $^{-2}$)	R_{star} (R $_{\odot}$)	R_p (R $_{\text{Jup}}$)	$\log X_{\text{H}_2\text{O}}$	$\log P_{\text{cloud-top}}$ (bar)
HD 209458b	cloud free	$2.92^{+0.05}_{-0.05}$	$1.22^{+0.02}_{-0.02}$	$1.41^{+0.03}_{-0.02}$	$-5.07^{+0.50}_{-0.36}$	–
	grey clouds	$2.9^{+0.05}_{-0.05}$	$1.23^{+0.02}_{-0.02}$	$1.41^{+0.03}_{-0.02}$	$-4.44^{+1.15}_{-0.77}$	$-0.93^{+0.62}_{-0.65}$
WASP-12b	cloud free	$2.99^{+0.03}_{-0.03}$	$1.57^{+0.03}_{-0.04}$	$1.81^{+0.04}_{-0.04}$	$-4.5^{+0.77}_{-0.49}$	–
	grey clouds	$2.99^{+0.02}_{-0.02}$	$1.58^{+0.03}_{-0.04}$	$1.79^{+0.04}_{-0.05}$	$-4.25^{+0.95}_{-0.65}$	$-1.07^{+0.57}_{-0.70}$
WASP-39b	cloud free	$2.67^{+0.04}_{-0.04}$	$0.94^{+0.02}_{-0.02}$	$1.27^{+0.02}_{-0.03}$	$-1.46^{+0.37}_{-1.14}$	–
	grey clouds	$2.67^{+0.04}_{-0.04}$	$0.94^{+0.02}_{-0.02}$	$1.28^{+0.02}_{-0.03}$	$-1.51^{+0.40}_{-1.27}$	$-0.79^{+0.98}_{-1.01}$

Table A5. Free parameters retrieved using cloud-free and grey-cloud models with tight temperature priors (see Section 5.4.2).

Planet	Model	$\log g_p$ (cm s $^{-2}$)	R_{star} (R $_{\odot}$)	R_p (R $_{\text{Jup}}$)	$\log X_{\text{H}_2\text{O}}$	T (K)	$\log P_{\text{cloud-top}}$ (bar)
HD 209458b	cloud free	$3.04^{+0.03}_{-0.04}$	$1.25^{+0.03}_{-0.04}$	$1.45^{+0.03}_{-0.04}$	$-5.94^{+0.14}_{-0.14}$	1360^{+25}_{-13}	–
	grey clouds	$2.9^{+0.05}_{-0.05}$	$1.25^{+0.02}_{-0.02}$	$1.4^{+0.03}_{-0.01}$	$-5.11^{+1.09}_{-0.62}$	1417^{+75}_{-31}	$-1.59^{+0.63}_{-1.10}$
WASP-12b	cloud free	$3.03^{+0.02}_{-0.02}$	$1.6^{+0.03}_{-0.04}$	$1.83^{+0.03}_{-0.05}$	$-6.16^{+0.17}_{-0.19}$	2176^{+51}_{-24}	–
	grey clouds	$2.99^{+0.02}_{-0.02}$	$1.59^{+0.03}_{-0.03}$	$1.75^{+0.04}_{-0.03}$	$-4.41^{+1.02}_{-0.84}$	2276^{+168}_{-91}	$-2.13^{+0.85}_{-1.03}$
WASP-39b	cloud free	$2.73^{+0.03}_{-0.04}$	$0.94^{+0.01}_{-0.02}$	$1.28^{+0.02}_{-0.03}$	$-5.43^{+0.17}_{-0.17}$	966^{+27}_{-13}	–
	grey clouds	$2.65^{+0.04}_{-0.05}$	$0.98^{+0.02}_{-0.03}$	$1.27^{+0.02}_{-0.03}$	$-3.94^{+0.92}_{-0.86}$	1045^{+107}_{-66}	$-2.35^{+0.85}_{-0.91}$

Table A6. Free parameters retrieved using cloud-free and grey-cloud models from Sing et al. (2016) data (see Section 6.2).

Planet	Data	Model	$\log g_p$ (cm s $^{-2}$)	R_{star} (R $_{\odot}$)	R_p (R $_{\text{Jup}}$)	$\log X_{\text{H}_2\text{O}}$	T (K)	$\log P_{\text{cloud-top}}$ (bar)
HD 209458b	WFC3 only	cloud free	$3.0^{+0.04}_{-0.05}$	$1.24^{+0.04}_{-0.03}$	$1.44^{+0.04}_{-0.04}$	$-5.49^{+0.21}_{-0.19}$	1081^{+97}_{-94}	–
		grey clouds	$2.92^{+0.04}_{-0.05}$	$1.25^{+0.02}_{-0.02}$	$1.4^{+0.02}_{-0.01}$	$-5.38^{+0.66}_{-0.41}$	1783^{+236}_{-304}	$-1.35^{+0.47}_{-0.64}$
HD 209458b	STIS+WFC3+IRAC	cloud free	$3.04^{+0.02}_{-0.03}$	$1.23^{+0.02}_{-0.03}$	$1.44^{+0.03}_{-0.04}$	$-5.25^{+0.13}_{-0.11}$	907^{+49}_{-36}	–
		grey clouds	$3.03^{+0.03}_{-0.06}$	$1.23^{+0.02}_{-0.02}$	$1.43^{+0.03}_{-0.03}$	$-5.23^{+0.15}_{-0.12}$	899^{+57}_{-95}	$-0.11^{+0.64}_{-0.37}$