Retrieval analysis of 38 WFC3 transmission spectra and resolution of the normalization degeneracy

Chloe Fisher[*][†] and Kevin Heng
Center for Space and Habitability, University of Bern, Gesellschaftsstrasse 6, CH-3012 Bern, Switzerland

ABSTRACT

A comprehensive analysis of 38 previously published Wide Field Camera 3 (WFC3) transmission spectra is performed using a hierarchy of nested-sampling retrievals: with versus without clouds, grey versus non-grey clouds, isothermal versus non-isothermal transit chords, and with water, hydrogen cyanide, and/or ammonia. We revisit the 'normalization degeneracy': the relative abundances of molecules are degenerate at the order-of-magnitude level with the absolute normalization of the transmission spectrum. Using a suite of mock retrievals, we demonstrate that the normalization degeneracy may be partially broken using WFC3 data alone, even in the absence of optical/visible data and without appealing to the presence of patchy clouds, although lower limits to the mixing ratios may be prior-dominated depending on the measurement uncertainties. With *James Webb Space Telescope*-like spectral resolutions, the normalization degeneracy may be completely broken from infrared spectra alone. We find no trend in the retrieved water abundances across nearly two orders of magnitude in exoplanet mass and a factor of 5 in retrieved temperature (about 500–2500 K). We further show that there is a general lack of strong Bayesian evidence to support interpretations of non-grey over grey clouds (only for WASP-69b and WASP-76b) and non-isothermal over isothermal atmospheres (no objects). 35 out of 38 WFC3 transmission spectra are well fitted by an isothermal transit chord with grey clouds and water only, while 8 are adequately explained by flat lines. Generally, the cloud composition is unconstrained.

Key words: planets and satellites: atmospheres.

1 INTRODUCTION

At the time of writing, we are in the transitional period between the *Hubble* and *James Webb Space Telescopes* (*HST* and *JWST*). In the foreseeable future, Wide Field Camera 3 (WFC3) transmission spectra spanning 0.8–1.7 μm will be superceded by NIRSpec data ranging from 0.6 to 5 μm and at enhanced spectral resolution. It is therefore timely to perform a uniform theoretical analysis of a consolidated data set of WFC3 transmission spectra, which is the overarching motivation behind the current study.

1.1 Observational motivation: a statistical study of cloudy atmospheres

Following the work of Iyer et al. (2016), Fu et al. (2017) recently conducted a statistical study of the transmission spectra of 34 exoplanets (mostly hot Jupiters) measured using WFC3 onboard *HST*, which were mostly gathered from Tsiaras et al. (2018). In order to isolate the spectral feature due to water,[1] they quantified the strength of absorption between 1.3 and 1.65 μm, relative to the continuum, in terms of the number of pressure scale heights, which they represented by A_H. Based on the finding that both A_H and the equilibrium temperature (T_{eq}) follow log-normal distributions, Fu et al. (2017) concluded that their sample of A_H is affected by observational bias. Tsiaras et al. (2018) defined an Atmospheric Detectability Index (ADI) to quantify the strength of detection of the water feature, but do not explicitly link the ADI to any trends in cloud properties. They concluded that all of their WFC3 transmission spectra, except for WASP-69b, are consistent with the presence of a grey cloud deck.

Our intention is to build upon the Fu et al. (2017) and Tsiaras et al. (2018) studies by subjecting their WFC3 sample to a detailed atmospheric retrieval study and elucidating the presence of assumptions, limitations, and trends. It follows the principle that the same data sets should be analysed by different groups (using different codes and techniques) within the community, so as to check for the

---

[*] E-mail: chloe.fisher@csh.unibe.ch
[†] University of Bern International 2021 Ph.D. Fellowship.

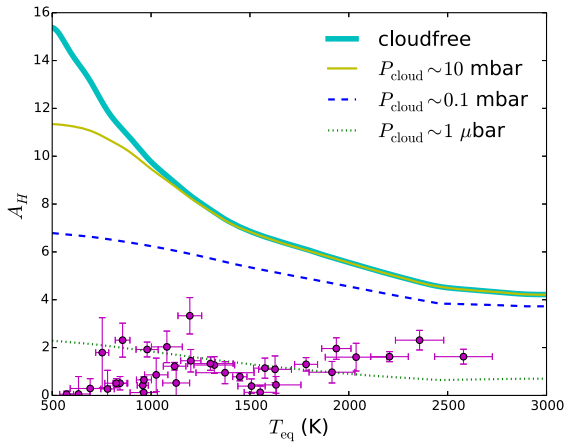[1]Technically, it is due to a collection of unresolved water lines.

Figure 1. Strength of WFC3 water feature, A_H, in terms of pressure scale heights as a function of the equilibrium temperature. Also shown are the theoretical predictions of A_H for cloud-free and cloudy atmospheres. For the latter, the curves correspond to transit chords probing $P_{cloud} \sim 1~\mu$bar, ~ 0.1 mbar and ~ 10 mbar if the opacity was solely due to grey clouds and the gravity is $\sim 10^3$ cm s^{-2}. It is apparent that all of the 34 atmospheres are cloudy if only water is assumed to be present.

Figure 2. Opacities of water, hydrogen cyanide, and ammonia as functions of wavelength. The `ExoMol` spectroscopic line list was used as input for computing these opacities. For water and hydrogen cyanide, we show a sequence of opacities from 900 to 2100 K. Darker shades of the same colour correspond to higher temperatures. For ammonia, the temperature sequence is terminated at 1500 K, because `ExoMol` does not provide any data for higher temperatures. The spectral resolution is 5 cm^{-1} and the pressure is fixed at $P = 10$ mbar, because these values are what we use in our retrievals (see the text for details).

consistency and robustness of theoretical interpretations (Fortney et al. 2016).

From a theoretical standpoint, A_H is an elegant quantity to examine, because the difference in transit radii between the peak of the water feature and the continuum is simply

$$A_H = \ln\left(\frac{\kappa_{max}}{\kappa_{min}}\right), \qquad (1)$$

where κ_{max} and κ_{min} are the maximum and minimum values of the water opacity in the WFC3 range of wavelengths. The preceding equation naturally derives from equation (2), if the volume mixing ratio of water (X_{H_2O}) is assumed to be uniform across altitude, and is free of the normalization degeneracy (see next subsection). Its simplicity allows us to do a first check on if the 34 objects in the sample gathered by Fu et al. (2017) have cloudy atmospheres.

In Fig. 1, we show curves of A_H for completely cloud-free atmospheres by assuming that the temperature (sampled by the water opacity) is the equilibrium temperature. Also shown are curves of A_H corresponding to cloudy atmospheres with constant opacities. For example, an opacity of 1 cm^2 g^{-1} corresponds to a transit chord probing a pressure ~ 0.1 mbar if only clouds (and not molecules) are present in the atmosphere. By comparing these theoretical curves to the measured data points of Fu et al. (2017), we tentatively conclude that all of the 34 transiting exoplanets in their sample have cloudy atmospheres. It is one of the goals of this study to examine if this conclusion is robust. Assuming that the temperature is some fraction of the equilibrium temperature merely translates the theoretical curves along the horizontal axis (not shown).

There is an additional supporting argument for the atmospheres being cloudy. By visual inspection of measured WFC3 transmission spectra, we noticed that the continuum bluewards of the 1.4 μm water feature tends to be somewhat flat, in contrast to the opacity of water which tends to be rather structured at these wavelengths (Fig. 2). This suggests that most, if not all, of the WFC3 transmission spectra measured so far are probing cloudy atmospheres – at least at the atmospheric limbs. However, this argument becomes less clear if ammonia and hydrogen cyanide are present, as they may mimic these effects on the spectra.
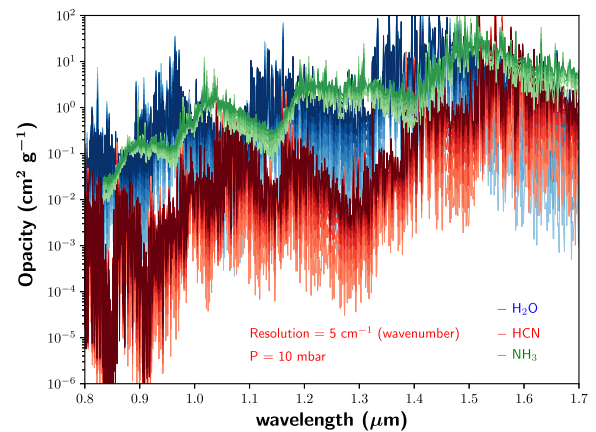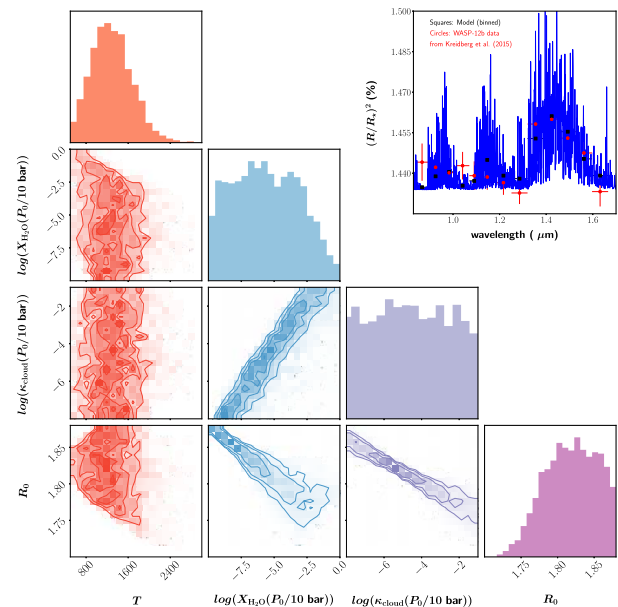

Figure 3. Posterior distributions of the water volume mixing ratio (X_{H_2O}), temperature (T), grey/constant cloud opacity (κ_{cloud}), and reference transit radius (R_0; uniform prior from 1.7 to 1.88 R_J) from a retrieval analysis of the WASP-12b transmission spectrum. The degeneracies between R_0 and the other quantities is apparent; R_0 is bounded from below because X_{H_2O} is bounded from above by unity. In this test, we have set $P_0 = 10$ bar but in our subsequent retrieval of the WASP-12b WFC3 transmission spectrum we will fit for P_0 (see the text and Fig. 20). The measured data and best-fitting model are shown in the top right panel. The physical units of T and κ_{cloud} are K and cm^2 g^{-1}, respectively, while R_0 is given in units of Jupiter radii (R_J). This retrieval assumes a constant mean molecular mass and ignores the effect of collisional induced absorption, which we will explore later in the current study.

Table 1. Assumed prior distributions for retrievals and values of physical constants used.

Quantity	Symbol	Range	Assumption	Units
Temperature	T	$(100, 2900)$	Uniform	K
Water mixing ratio	X_{H_2O}	$(10^{-13}, 1)$	Log-uniform	–
Hydrogen cyanide mixing ratio	X_{HCN}	$(10^{-13}, 1)$	Log-uniform	–
Ammonia mixing ratio	X_{NH_3}	$(10^{-13}, 1)$	Log-uniform	–
Grey cloud opacity	κ_{cloud}	$(10^{-12}, 10^2)$	Log-uniform	cm^2 g^{-1}
Opacity normalization for non-grey cloud model	κ_0	$(10^{-10}, 10^{-1})$	Log-uniform	cm^2 g^{-1}
Composition parameter in non-grey cloud model	Q_0	$(1, 100)$	Uniform	–
Index in non-grey cloud model	a	$(3, 6)$	Uniform	–
Monodisperse, spherical cloud particle radius	r_c	$(10^{-7}, 10^{-1})$	Log-uniform	cm
Non-isothermal temperature profile parameter	b	$(-30, -1), (1, 30)$	Uniform	–
Reference transit radius^a	R_0	$(1.619, 1.799)$	Uniform	R_J
Reference pressure	P_0	$(10^{-1}, 10^3)$	Log-uniform	bar
Equatorial radius of Jupiter	R_J	7.1492×10^9	–	cm
Mass of hydrogen atom	m_H		–	cm
Atomic mass unit	m_{amu}	$1.66053904 \times 10^{-24}$	–	g
Boltzmann constant	k_B	$1.38064852 \times 10^{-16}$	–	erg K^{-1}

Notes. ^{a}Only used in the test retrievals of WASP-17b (Section 3.2.2).

In the current study, one of our goals is to formalise this finding by performing atmospheric retrieval, within a nested-sampling framework (e.g. Skilling 2006; Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009; Benneke & Seager 2013; Feroz et al. 2013; Waldmann et al. 2015; Lavie et al. 2017; Tsiaras et al. 2018), on each of the 34 objects in the Fu et al. (2017) sample. We construct a hierarchy of models with increasing levels of sophistication: cloud-free model (two parameters), cloudy model with constant/grey cloud opacity (three parameters), cloudy model with non-grey opacity (six parameters). It is assumed that the main molecular absorber is water. If hydrogen cyanide (HCN) and ammonia (NH$_3$) are added to the analysis (MacDonald & Madhusudhan 2017), then it adds two more free parameters for a maximum of eight parameters for the isothermal model. Our non-isothermal model adds another parameter. For comparison, MacDonald & Madhusudhan (2017) employ a 16-parameter model based partly on the heritage from Madhusudhan & Seager (2009).

We use the computed Bayesian evidence (Trotta 2008) from the retrievals to select the best model given the quality of the data, and hence determine if the atmospheres are cloudy, if cloud properties may be meaningfully constrained, and if NH$_3$ and/or HCN are detected in a given WFC3 spectrum. Unlike the approach adopted by MacDonald & Madhusudhan (2017), we do not test for patchy clouds. Rather, we test essentially for whether the cloud particles are small or large (compared to the wavelengths probed).

1.2 Theoretical motivation: the normalization degeneracy

Atmospheric retrievals of transmission spectrum typically specify a plane-parallel model atmosphere, assume azimuthal symmetry and then trace a transit chord through a set of atmospheric columns (each approximated by a plane-parallel atmosphere) (Madhusudhan & Seager 2009; Benneke & Seager 2012, 2013; Line et al. 2013; Waldmann et al. 2015). This brute-force procedure for calculating the transmission spectrum was previously described by Brown (2001) and Hubbard et al. (2001). In the current study, our intention is to build a nested-sampling retrieval framework around a validated analytical model for computing the transmission spectrum that bypasses the need for a brute-force calculation. Complementary

to previous retrieval studies, we make a different set of investments, approximations, and simplifications.

Building on the work of Lecavelier des Etangs et al. (2008), de Wit & Seager (2013), and Bétrémieux & Swain (2017), Heng & Kitzmann (2017) demonstrated that an analytical expression for the isothermal transit chord of an atmosphere,

$$R = R_0 + H(\gamma + E_1 + \ln \tau), \qquad (2)$$

is accurate enough^2 to model WFC3 transmission spectra for atmospheres with temperatures $\sim$1000 K or hotter, where we have

$$\tau \equiv \frac{\kappa P_0}{g} \sqrt{\frac{2\pi R_0}{H}}. \qquad (3)$$

The pressure scale height is given by H, the Euler–Mascheroni constant by $\gamma \approx 0.57721$, and the surface gravity by g. The exponential integral of the first order is given by $E_1(\tau)$ (Abramowitz & Stegun 1970; Arfken & Weber 1995), which has the argument τ. For a WFC3 spectrum dominated by water, the opacity is $\kappa \propto X_{H_2O}$, where X_{H_2O} is the volume mixing ratio of water. Equation (2) assumes that $R_0 < R$; if the layer of the atmosphere located at R_0 is opaque in the WFC3 bandpass ($\tau \gg 1$), then the E_1 term may be dropped.

Equation (2) straightforwardly shows that there exists a three-way degeneracy between the reference transit radius (R_0), reference pressure (P_0), and X_{H_2O}, which was first noticed numerically^3 by Benneke & Seager (2012) and Griffith (2014). The values of R_0 and P_0, as well as the relationship between them, are a priori unknown, because it is akin to having prior knowledge of the structure of the exoplanet. It is apparent that a small change in R_0 causes a large variation in X_{H_2O}. Furthermore, it is $X_{H_2O}P_0$, and not X_{H_2O} alone, that is being retrieved from the data. It is worth emphasising that these obstacles do not exist in the forward problem, where one makes a specific set of assumptions (e.g. solar metallicity, chemical equilibrium) and computes the transmission spectrum, but they are

---

^{2}Meaning that the errors incurred are smaller than the noise floor (about 50 parts per million) of *HST* and the expected noise floor of *JWST*.
^{3}Our stand is that a numerical demonstration of an effect alone does not qualify as attaining full understanding of it, until its theoretical (analytical) formalism has been elucidated.

Table 2. Summary of retrieval outcomes (38 objects, 42 sets of retrievals).

Name	T_{eq} (K)	T (K)	$\log(X_{H_2O})$	$\log(X_{HCN})$	$\log(X_{NH_3})$	Cloudy?	Non-grey clouds?	$\log(\kappa_{cloud})$ (cm^2 g^{-1})
GJ436b	633	a	a	a	a	a	a	a
GJ3470b	692	629^{+911}_{-239}	$-4.75^{+2.47}_{-2.93}$	–	–	Yes	No	$-2.17^{+2.67}_{-1.29}$
HAT-P-1b	1320	1223^{+435}_{-492}	$-0.75^{+0.23}_{-1.0}$	–	$-4.44^{+1.59}_{-5.65}$	Maybe	–	–
HAT-P-3b	1127	1145^{+1011}_{-604}	$-7.88^{+5.99}_{-3.47}$	–	–	Maybe	–	$-2.42^{+2.63}_{-2.35}$
HAT-P-11b	856	1002^{+524}_{-255}	$-2.56^{+2.13}_{-2.2}$	–	–	Maybe	–	$-2.52^{+1.45}_{-1.68}$
HAT-P-12b	958	a	a	a	a	a	a	a
HAT-P-17b	780	1114^{+984}_{-1255}	$-8.25^{+3.4}_{-2.91}$	$-4.37^{+2.08}_{-5.17}$	$-7.39^{+4.02}_{-3.58}$	Yes	No	$-4.82^{+3.74}_{-4.56}$
HAT-P-18b	843	347^{+146}_{-137}	$-1.83^{+0.8}_{-1.48}$	–	–	Maybe	–	$-7.25^{+2.57}_{-3.07}$
HAT-P-26b	980	647^{+118}_{-82}	$-2.37^{+0.86}_{-1.24}$	–	–	Yes	No	$-2.66^{+0.65}_{-0.94}$
HAT-P-32b	1784	1109^{+251}_{-190}	$-2.39^{+1.09}_{-1.57}$	–	–	Yes	No	$-1.9^{+0.88}_{-1.22}$
HAT-P-38b	1080	1876^{+637}_{-1074}	$-0.41^{+0.22}_{-1.8}$	–	$-5.25^{+2.92}_{-5.31}$	Maybe	–	–
HAT-P-41b	1937	1561^{+624}_{-507}	$-0.9^{+0.28}_{-1.2}$	–	$-5.32^{+2.27}_{-4.94}$	Maybe	–	–
HD149026b	1627	1672^{+679}_{-687}	$-4.69^{+4.11}_{-5.02}$	–	–	Maybe	–	$-2.33^{+2.14}_{-2.61}$
HD189733b	1201	782^{+172}_{-107}	$-2.3^{+0.87}_{-1.26}$	–	–	Yes	No	$-2.62^{+0.84}_{-0.86}$
HD209458b	1449	777^{+193}_{-95}	$-2.65^{+0.81}_{-1.24}$	–	–	Yes	No	$-1.75^{+0.69}_{-0.94}$
WASP-29b	963	a	a	a	a	a	a	a
WASP-31b	1576	a	a	a	a	a	a	a
WASP-39b	1119	600^{+86}_{-72}	$-2.3^{+0.4}_{-0.67}$	–	–	Yes	No	$-5.69^{+1.94}_{-4.54}$
WASP-43b	1374	835^{+340}_{-121}	$-2.89^{+1.13}_{-3.07}$	–	–	Yes	No	$-2.03^{+1.04}_{-1.01}$
WASP-52b	1300	776^{+278}_{-149}	$-2.65^{+0.84}_{-1.03}$	–	–	Yes	No	$-2.48^{+0.74}_{-0.83}$
WASP-63b	1508	1068^{+700}_{-352}	$-5.83^{+2.9}_{-4.23}$	–	–	Yes	No	$-2.02^{+2.03}_{-1.56}$
WASP-67b	1026	a	a	a	a	a	a	a
WASP-69b	964	658^{+148}_{-107}	$-4.24^{+1.03}_{-1.09}$	–	–	Yes	Yes	$-2.44^{+0.81}_{-0.89}$
WASP-74b	1915	1152^{+798}_{-354}	$-7.94^{+3.72}_{-3.64}$	–	–	Yes	No	$-1.23^{+1.9}_{-1.64}$
WASP-76b	2206	1647^{+185}_{-178}	$-5.3^{+0.61}_{-0.61}$	–	–	Yes	Yes	$-1.78^{+0.47}_{-0.65}$
WASP-80b	824	a	a	a	a	a	a	a
WASP-101b	1552	1616^{+256}_{-288}	$-9.03^{+3.13}_{-2.63}$	$-2.56^{+0.5}_{-0.62}$	$-8.3^{+3.84}_{-3.06}$	Yes	No	$-6.79^{+3.45}_{-3.37}$
WASP-121b	2358	1523^{+468}_{-290}	$-3.09^{+1.01}_{-1.26}$	–	–	Yes	No	$-2.32^{+0.86}_{-1.05}$
XO-1b	1196	977^{+254}_{-174}	$-1.06^{+0.29}_{-0.47}$	–	–	Maybe	–	–
GJ1214b	573	a	a	a	a	a	a	a
HD97658b	753	1323^{+224}_{-286}	$-7.48^{+4.88}_{-3.7}$	–	$-0.48^{+0.19}_{-0.23}$	Maybe	–	–
WASP-17b	1632	1678^{+610}_{-448}	$-0.98^{+0.46}_{-4.94}$	–	–	Maybe	–	$-3.61^{+2.4}_{-4.31}$
WASP-19b	2037	2039^{+381}_{-338}	$-2.86^{+1.2}_{-1.49}$	–	–	Maybe	–	–
WASP-12b	2580	1540^{+358}_{-242}	$-3.02^{+1.09}_{-1.36}$	–	–	Yes	No	$-1.9^{+0.97}_{-1.11}$
TRAPPIST-1d	288	a	a	a	a	a	a	a
		a,b	a,b	a,b	a,b	a,b	a,b	a,b
TRAPPIST-1e	251	a	a	a	a	a	a	a
		$1173^{+1108}_{-729}{}^{b}$	$-10.02^{+2.02}_{-2.02}{}^{b}$	–	–	Maybe^{b}	–	–
TRAPPIST-1f	219	a	a	a	a	a	a	a
		$1214^{+1089}_{-815}{}^{b}$	$-10.09^{+1.98}_{-1.94}{}^{b}$	–	–	Maybe^{b}	–	–
TRAPPIST-1g	199	a	a	a	a	a	a	a
		$896^{+1238}_{-590}{}^{b}$	$-9.61^{+2.5}_{-2.24}{}^{b}$	–	–	Maybe^{b}	–	–

Notes.

For 'Cloudy?': 'Yes' refers to cases where all of the cloud-free models have Bayes factors of unity or more. 'No' means only cloud-free models have Bayes factor of less than unity. 'Maybe' means a mixture of cloud-free and cloudy models have Bayes factor of less than unity. **For 'Non-grey clouds?':** 'Yes' refers to cases where only non-grey-cloud models have Bayes factors of less than unity. 'No' means a mixture of non-grey cloud and grey-cloud models have Bayes factors of less than unity.

[a]Flat-line fit has Bayes factor of less than unity and no atmospheric properties may be retrieved.

[b]For the TRAPPIST-1 exoplanets, we also examine Earth-like atmospheres ($m = 29m_H$).

front and centre in the inverse problem. Heng & Kitzmann (2017) pointed out these issues, but they did not examine them further within a Bayesian retrieval framework, which partially motivates the current study.

Fig. 3 shows a retrieval calculation performed using a new code (HELIOS-T) presented as part of the current study, which we constructed specifically to perform fast retrievals on transmission spec-tra at low spectral resolution.[4] (A detailed description of methodology will come later in Section 2.) It demonstrates that while the

[4]At high spectral resolution, the fully resolved spectral lines may span many orders of magnitude in pressure between the line peaks and wings, thereby violating the isobaric assumption.

Table 3. Summary of input parameters (38 objects, 42 sets of retrievals).

Name	$R_\star$ ($R_\odot$)	M (M_J)	R_0 (R_J)	g (cm s^{-2})	References	$\bar{R}_{\rm WFC3}$ (R_J)
GJ436b	0.455	$0.078^{+0.007}_{-0.008}$	0.3532	1318	von Braun et al. (2012)	0.3693
GJ3470b	0.48	0.043 ± 0.005	0.3287	676	Biddle et al. (2014)	0.3630
HAT-P-1b	1.115	0.524 ± 0.031	1.213	750	Johnson et al. (2008), Sing et al. (2016)	1.272
HAT-P-3b	0.799	0.591 ± 0.018	0.8383	2138	Chan et al. (2011)	0.8559
HAT-P-11b	0.75	0.081 ± 0.009	0.4077	1122	Bakos et al. (2010)	0.4332
HAT-P-12b	0.701	0.211 ± 0.012	0.8770	562	Hartman et al. (2009)	0.9341
HAT-P-17b	0.838	0.534 ± 0.018	0.9677	1288	Howard et al. (2012)	0.9880
HAT-P-18b	0.717	0.196 ± 0.008	0.9349	542	Esposito et al. (2014)	0.9552
HAT-P-26b	0.788	0.059 ± 0.007	0.4741	447	Hartman et al. (2011a)	0.5475
HAT-P-32b	1.219	0.860 ± 0.164	1.714	661	Hartman et al. (2011b)	1.804
HAT-P-38b	0.923	0.267 ± 0.020	0.8010	977	Sato et al. (2012)	0.8380
HAT-P-41b	1.683	0.800 ± 0.102	1.568	692	Hartman et al. (2012)	1.662
HD149026b	1.368	$0.359^{+0.022}_{-0.021}$	0.6536	2291	Torres, Winn & Holman (2008)	0.6774
HD189733b	0.805	1.162 ± 0.058	1.200	1950	Boyajian et al. (2015)	1.221
HD209458b	1.203	0.64 ± 0.09	1.350	759	Boyajian et al. (2015)	1.414
WASP-29b	0.808	0.244 ± 0.020	0.7330	891	Hellier et al. (2010)	0.7692
WASP-31b	1.252	0.478 ± 0.029	1.379	456	Anderson et al. (2011)	1.535
WASP-39b	0.918	0.283 ± 0.041	1.207	414	Maciejewski et al. (2016)	1.297
WASP-43b	0.67	1.78 ± 0.10	1.029	4699	Hellier et al. (2011)	1.039
WASP-52b	0.79	0.46 ± 0.02	1.199	646	Hébrard et al. (2013)	1.266
WASP-63b	1.88	0.38 ± 0.03	1.316	417	Hellier et al. (2012)	1.437
WASP-67b	0.87	0.42 ± 0.04	1.314	501	Hellier et al. (2012)	1.383
WASP-69b	0.813	0.260 ± 0.017	0.9563	532	Anderson et al. (2014)	1.017
WASP-74b	1.64	0.95 ± 0.06	1.456	891	Hellier et al. (2015)	1.528
WASP-76b	1.73	0.92 ± 0.03	1.635	631	West et al. (2016)	1.752
WASP-80b	0.586	$0.538^{+0.035}_{-0.036}$	0.9562	1396	Triaud et al. (2015)	0.9760
WASP-101b	1.29	0.50 ± 0.04	1.274	575	Hellier et al. (2014)	1.364
WASP-121b	1.458	$1.183^{+0.064}_{-0.062}$	1.633	940	Delrez et al. (2016)	1.717
XO-1b	0.934	$0.918^{+0.081}_{-0.078}$	1.172	1626	Torres et al. (2008)	1.197
GJ1214b	0.211	0.019 ± 0.003	0.2135	768	Anglada-Escudé et al. (2013)	0.2385
HD97658b	0.741	$0.024^{+0.003}_{-0.002}$	0.2036	1466	van Grootel et al. (2014)	0.2208
WASP-17b	1.583	0.477 ± 0.033	1.709	316	Southworth et al. (2012)	1.897
WASP-19b	1.004	1.114 ± 0.036	1.311	1419	Tregloan-Reed, Southworth & Tappert (2013)	1.378
WASP-12b	1.57	1.41 ± 0.10	1.748	977	Hebb et al. (2009), Kreidberg et al. (2015)	1.836
TRAPPIST-1d	0.121	$9.34^{+1.10}_{-1.23} \times 10^{-4}$	0.05402 0.07268^a	474	Grimm et al. (2018), van Grootel et al. (2018)	0.07436
TRAPPIST-1e	0.121	$2.43^{+0.24}_{-0.25} \times 10^{-3}$	0.07329 0.08174^a	912	Grimm et al. (2018), van Grootel et al. (2018)	0.08250
TRAPPIST-1f	0.121	$2.94 \pm 0.25 \times 10^{-3}$	0.08490 0.09294^a	837	Grimm et al. (2018), van Grootel et al. (2018)	0.09366
TRAPPIST-1g	0.121	$3.61^{+0.30}_{-0.31} \times 10^{-3}$	0.09580 0.1030^a	854	Grimm et al. (2018), van Grootel et al. (2018)	0.1036

Note. ^{a}For the TRAPPIST-1 exoplanets, we also examine Earth-like atmospheres ($m = 29m_{\rm H}$).

temperature may be robustly retrieved, there are order-of-magnitude degeneracies associated with the water mixing ratio and cloud opacity that arise from small variations of R_0 (in the third significant figure), as previously elucidated by Heng & Kitzmann (2017). In this study, we wish to examine if R_0 or P_0 may be used as a fitting parameter to break the normalization degeneracy. We further examine if the normalization degeneracy may be broken with WFC3 transmission spectra alone, or if *JWST*-like spectra is needed.

1.3 Layout of the current study

In Section 2, we describe our theoretical methodology, including how we compute transit radii and opacities. In Section 3, we perform suites of tests, a detailed analysis of the 38 WFC3 transmission spectra in the Tsiaras et al. (2018) and de Wit et al. (2018) samples and elucidate trends among the retrieved quantities. The implications of our results are discussed in Section 4. Table 1 lists our

assumptions for the prior distributions of parameters. Table 2 summarizes our retrieval results. Table 3 summarizes some of the input parameters for the retrievals.

Our new nested-sampling retrieval code for transmission spectra, HELIOS-T, is part of our open-source suite of tools for analysing exoplanetary atmospheres known as the Exoclimes Simulation Platform (www.exoclime.org or https://github.com/exoclime).

2 METHODOLOGY

2.1 Transmission spectra

As explained in Section 1, equation (2) describes our forward model for transforming a given temperature, surface gravity, opacity function, reference transit radius, and reference pressure into a transmission spectrum. The accuracy of equation (2) has previously been demonstrated by Heng & Kitzmann (2017) and we will not repeat
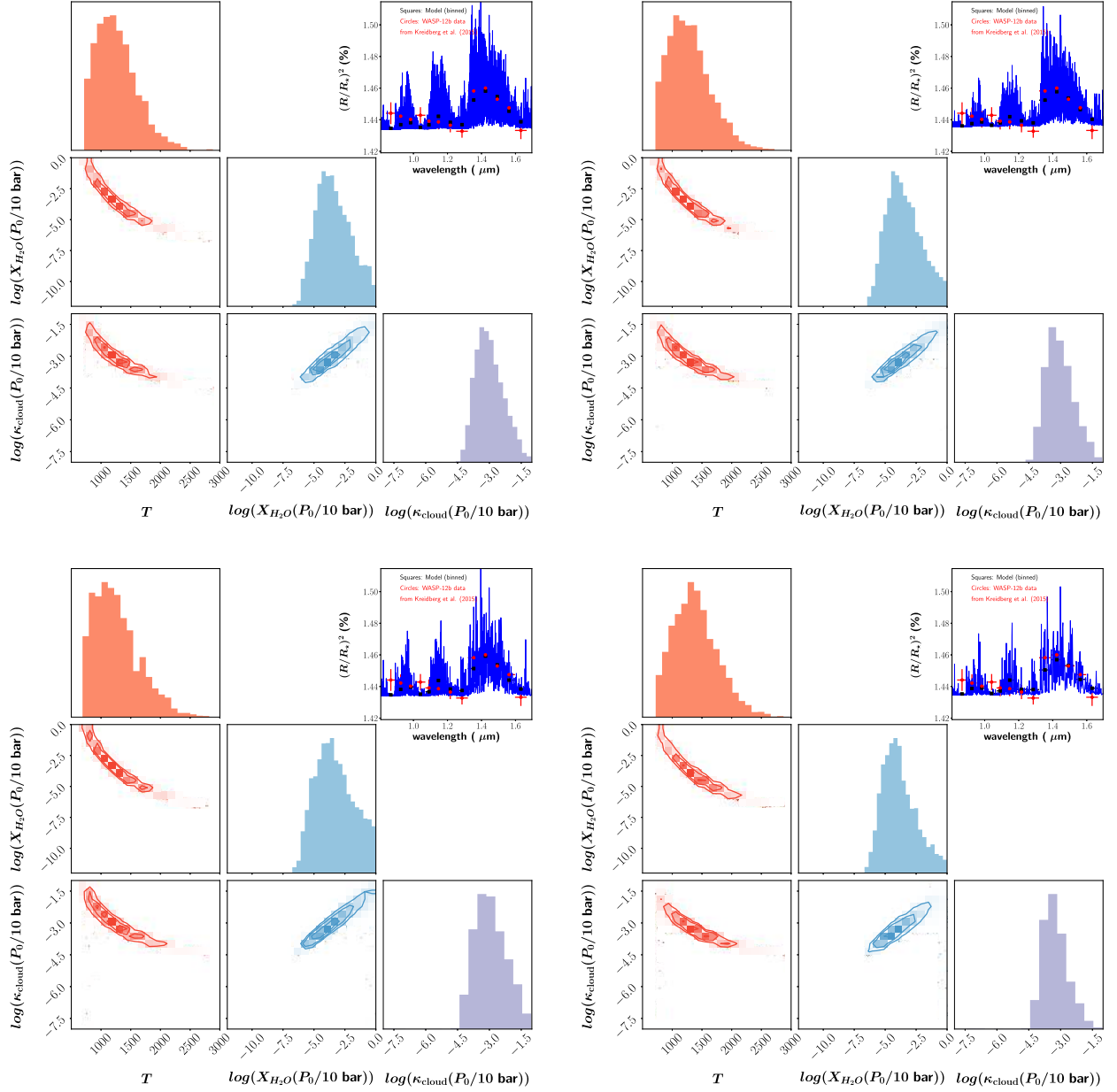
Figure 4. Same as Fig. 3, but using spectral resolutions of 1 cm^{-1} (top left panel), 2 cm^{-1} (top right panel), 5 cm^{-1} (bottom left panel), and 10 cm^{-1} (bottom right panel) for the ExoMol water opacity.

the analysis and explanation here. To test for non-isothermality, we use another formula derived by Heng & Kitzmann (2017),

$$R = R_0 + H\tau^{1/b}\left(\gamma + E_1 + \ln\tau\right), \qquad (4)$$

where the reference optical depth is now given by

$$\tau = \frac{\pi P_0 \kappa}{2g}\sqrt{\frac{2R_0\,|b|}{H}}. \qquad (5)$$

We again have $E_1 = E_1(\tau)$. The dimensionless index b is the ratio of the non-isothermal to the isothermal scale height. If the values of $|b|$ are much larger than unity, then it means that the behaviour is close to being isothermal. Essentially, our simplified temperature–pressure profile is described by two parameters.

Our approach is complementary to other approaches in the literature that use more complicated prescriptions for temperature–pressure profiles. For example, Madhusudhan & Seager (2009) and MacDonald & Madhusudhan (2017) use nine- and seven-parameter fitting functions, respectively. Again, we make a different investment: we choose to simplify the temperature profile prescription in order to isolate the effects of the other parameters. It allows us to more cleanly study degeneracies.

2.2 Opacities

Our three H$_2$O, HCN, and NH$_3$ opacities are taken from the ExoMol spectroscopic database (Barber et al. 2006; Yurchenko, Barber & Tennyson 2011; Yurchenko et al. 2013; Barber et al. 2014; Yurchenko & Tennyson 2014). In a single set of tests (see Section 3.1.4), we also use the HITRAN (Rothman et al. 1987, 1992, 1998, 2003, 2005, 2009, 2013) and HITEMP (Rothman et al. 2010)
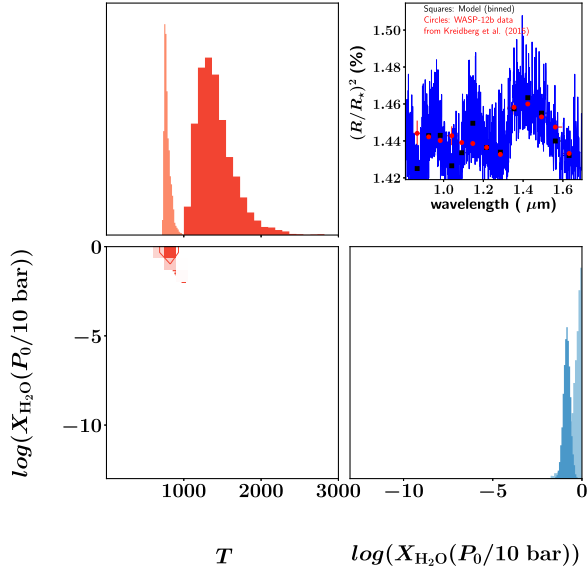
Figure 5. Same as Fig. 3, but for cloud-free models in which we fix $m = 2.4 m_H$ (log X_{H_2O} posterior bumps up against 0) versus a variable m (posteriors distributions are in a darker shade) that takes into account water-rich atmospheres.

databases for water. For a review of the spectroscopic databases, please see Tennyson & Yurchenko (2017). For the procedure on how to use the ExoMol inputs to compute opacities, we refer the reader to Grimm & Heng (2015), chapter 5 of Heng (2017), and Yurchenko, Al-Refaie & Tennyson (2018). Examples of opacities for all three molecules are given in Fig. 2.

The opacity function used in equation (2) is given by

$$\kappa = \frac{X_{H_2O} m_{H_2O} \kappa_{H_2O}}{m} + \frac{X_{HCN} m_{HCN} \kappa_{HCN}}{m} + \frac{X_{NH_3} m_{NH_3} \kappa_{NH_3}}{m} + \kappa_{cloud}, \tag{6}$$

where m is the mean molecular mass, m_{H_2O} is the mass of the water molecule, κ_{H_2O} is the water opacity, X_{HCN} is the volume mixing ratio of hydrogen cyanide, m_{HCN} is the mass of the hydrogen cyanide molecule, κ_{HCN} is the hydrogen cyanide opacity, X_{NH_3} is the volume mixing ratio of ammonia, m_{NH_3} is the mass of the ammonia molecule, κ_{NH_3} is the ammonia opacity, and κ_{cloud} is the cloud opacity.

Denoting the atomic mass unit by m_{amu}, the mean molecular weight ($\mu = m/m_{amu}$) is given by

$$\mu = 2.4 X_{H_2} + \frac{X_{H_2O} m_{H_2O}}{m_{amu}} + \frac{X_{HCN} m_{HCN}}{m_{amu}} + \frac{X_{NH_3} m_{NH_3}}{m_{amu}}. \tag{7}$$

The mixing ratio of molecular hydrogen is determined by demanding that all mixing ratios sum to unity,

$$1.1 X_{H_2} + X_{H_2O} + X_{HCN} + X_{NH_3} = 1, \tag{8}$$

where we have assumed that the helium mixing ratio follows cosmic abundance ($X_{He} = 0.1 X_{H_2}$).

The molecular opacities are sampled at 1 mbar for the first suite of tests (Section 3.1; to ensure continuity with Heng & Kitzmann 2017) and 10 mbar for our second suite of tests (Section 3.2.2) and actual results (see Section 3.3). The cloud mixing ratio is subsumed into κ_{cloud}. The opacity associated with collision-induced absorption (both H_2–H_2 and H_2–He) are taken from Rothman et al. (2013).

An unresolved physics problem inherent in the computation of opacities concerns the effects of pressure broadening. The spectral lines of *isolated* atoms and molecules are described rather well by a Voigt profile. As a population, collisions between them become important at high enough pressures, which modify the shape of the far line wings of the profile. It remains unknown exactly what 'far' means and how to compute these modified profiles. In practice, various workers in the field have resorted to truncating the Voigt profiles at some ad hoc distance from line centre (see Grimm & Heng 2015 and references therein). For this study, we use a line-wing cut-off of 100 cm^{-1}. Fortunately, since transmission spectra probe pressures that are tenuous enough such that pressure broadening has a negligible effect for $\sim$1000 K atmospheres, this is not a limiting issue.

Another limitation is that, at the time of writing, the NH$_3$ opacities do not exist for temperatures above 1600 K (Yurchenko et al. 2011). In the absence of these data, we set the opacity for NH$_3$ to be zero for temperatures above 1600 K.

2.3 Cloud models

We consider both grey and non-grey clouds. For our grey cloud model, we assume a constant cloud opacity, which is physically equivalent to assuming that the cloud particles are much larger than the WFC3 wavelengths being probed. Our non-grey cloud model uses the opacity of (Kitzmann & Heng 2018),

$$\kappa_{cloud} = \frac{\kappa_0}{Q_0 x^{-a} + x^{0.2}}, \tag{9}$$

where $x = 2\pi r_c/\lambda$ is the dimensionless size parameter, r_c is the particle radius, and λ is the wavelength. In their study of 32 condensate species, Kitzmann & Heng (2018) showed that $Q_0 \approx$ 0.1–65 is a proxy for cloud composition with larger values corresponding to more volatile species. For example, water ice has $Q_0 = 64.98$ and olivine has $Q_0 \approx 10$. The index a ranges from 3 to 7; $a = 4$ corresponds to Rayleigh scattering. Our non-grey cloud model has four free parameters: κ_0, Q_0, r_c, and a. The immediate implication of the preceding equation is that if the cloud is grey ($a \approx 0$), then the composition cannot be decisively constrained.

Conceptually, the treatment of Lee, Heng & Irwin (2013) and Kitzmann & Heng (2018) are identical in that they both allow smooth transitions between the Rayleigh and large-particle regimes. However, Lee et al. (2013) assumed $a = 4$, whereas Kitzmann & Heng (2018) calibrated Q_0 and a against a larger library of species and composition.

2.4 Data

For 30 out of 38 objects, the WFC3 transmission spectra were obtained from Tsiaras et al. (2018) and provided in electronic form by the first author (Tsiaras, private communication). For WASP-17b, WASP-19b, GJ 1214b, and HD97658b, the WFC3 transmission spectra were obtained from Mandell et al. (2013), Huitson et al. (2013), Kreidberg et al. (2014a), and Knutson et al. (2014b), respectively. The WFC3 transmission spectra of TRAPPIST-1d, e, f, and g were taken from de Wit et al. (2018). The stellar radii and surface gravities for each object are listed in Table 3. Uncertainties in the stellar radii manifest themselves as uncertainties in the normalization of the transmission spectra.
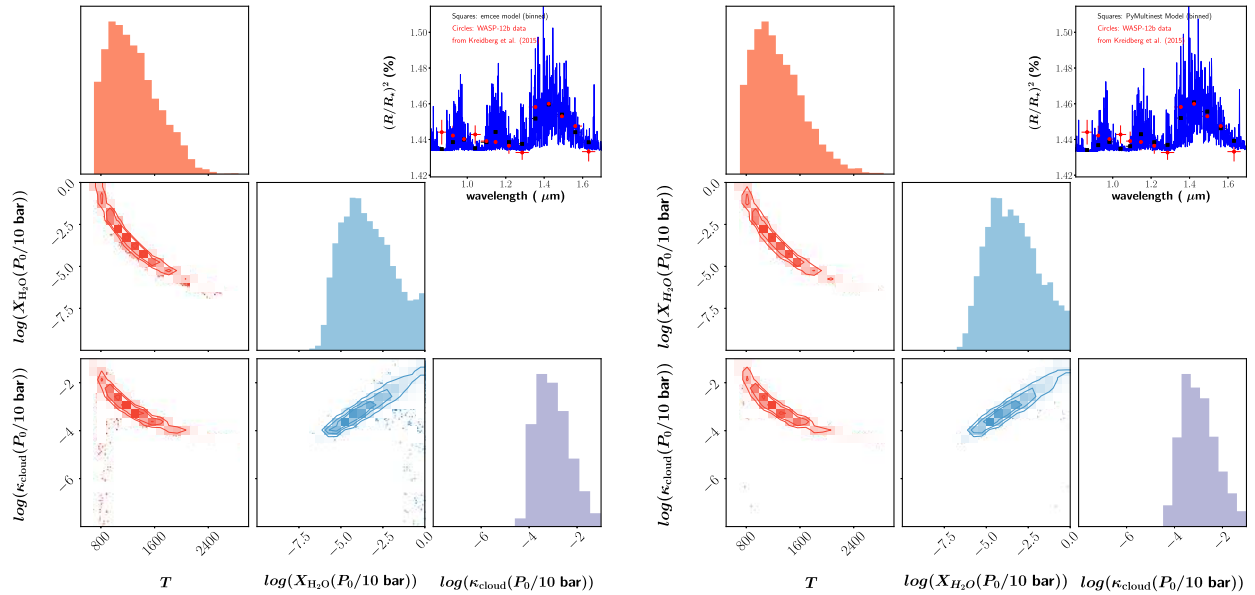
Figure 6. Same as Fig. 3, but comparing a MCMC (left-hand panel) versus nested-sampling (right-hand panel) retrieval approach.

3 RESULTS

3.1 Suite of tests on WASP-12b transmission spectrum

To provide continuity between Heng & Kitzmann (2017) and this study, we use the WFC3 transmission spectrum of WASP-12b (13 data points), measured by Kreidberg et al. (2015), as our starting point for tests. To cleanly isolate the effects studied, we begin with using a constant/grey cloud opacity. Note that for these WASP-12b tests only (Figs 4–9), the molecular opacities are sampled at 1 mbar, CIA is not included and m is fixed at $2.4m_H$, where m_H is the mass of the hydrogen atom (which we take to be one atomic mass unit, m_{amu}), unless otherwise stated. In these tests only, we set $R_0 = 1.79 R_J$ and $P_0 = 10$ bar. These restrictions are lifted for the rest of the study.

3.1.1 Spectral resolution of opacities

In Fig. 4, we perform resolution tests associated with the sampling of the water opacity across wavenumber. We show retrieval outcomes for spectral resolutions of 1, 2, 5, and 10 cm^{-1}. For all of these values, the posterior distributions of T, X_{H_2O}, and κ_{cloud} are somewhat similar. Specifically, the retrieved temperatures are 1218^{+388}_{-297}, 1252^{+393}_{-307}, 1203^{+468}_{-323}, and 1363^{+403}_{-343}, respectively. The retrieved water volume mixing ratios are $-3.51^{+1.66}_{-1.26}$, $-3.61^{+1.68}_{-1.27}$, $-3.46^{+1.93}_{-1.49}$, and $-4.04^{+1.62}_{-1.14}$, respectively. For the rest of the study, we will adopt a sampling resolution of 5 cm^{-1}.

3.1.2 Cloudy versus cloud-free

Another necessary check is to determine that cloudy models are necessary in the first place for WASP-12b. In Fig. 5, we subject the WASP-12b WFC3 transmission spectrum to two cloud-free retrievals: the first has a fixed $m = 2.4m_H$, while the second has a variable m. For the retrieval with a fixed m, the outcome is implausible as the water volume mixing ratio is ~10–100 per cent. The

retrieval with a variable m produces more plausible posteriors, but even by visual inspection it is apparent that the cloud-free model struggles to match the somewhat flat spectral continuum bluewards of the 1.4 µm water feature. For the rest of the WFC3 transmission spectra, we will not show the posterior distributions associated with the cloud-free retrieval (unless it has the highest Bayesian evidence in the model hierarchy), but we will still include them in the overall analysis.

3.1.3 MCMC versus nested sampling

The next logical test is to compare cloudy retrievals obtained using a Markov Chain Monte Carlo (MCMC) versus nested sampling approach. For the former, we use the open-source EMCEE package (Foreman-Mackey et al. 2013). For the latter, we use the open-source PYMULTINEST package (Buchner et al. 2014). Fig. 6 compares the outcome from this pair of retrievals. It is reassuring that the posterior distributions of T, X_{H_2O}, and κ_{cloud} are somewhat similar, although we note that the retrieval performed with MCMC produces higher values of the water volume mixing ratio in the tail of the distribution (towards $X_{H_2O} = 1$). The reason to select the nested-sampling approach over MCMC is because it allows us to straightforwardly compute the Bayesian evidence associated with each model, which in turn allows us to formally apply Occam's Razor.

3.1.4 Choice of spectroscopic databases: HITRAN versus HITEMP versus ExoMol

Perhaps the most surprising outcome of our series of WASP-12b tests is shown in Fig. 7, where we examine the retrieval outcomes using the HITRAN, HITEMP, and ExoMol spectroscopic databases to construct the water opacity. The main shortcoming with HITRAN is that it omits the weak lines of water that contribute prominently to the spectral continuum when $T \sim 1000$ K or hotter. HITEMP
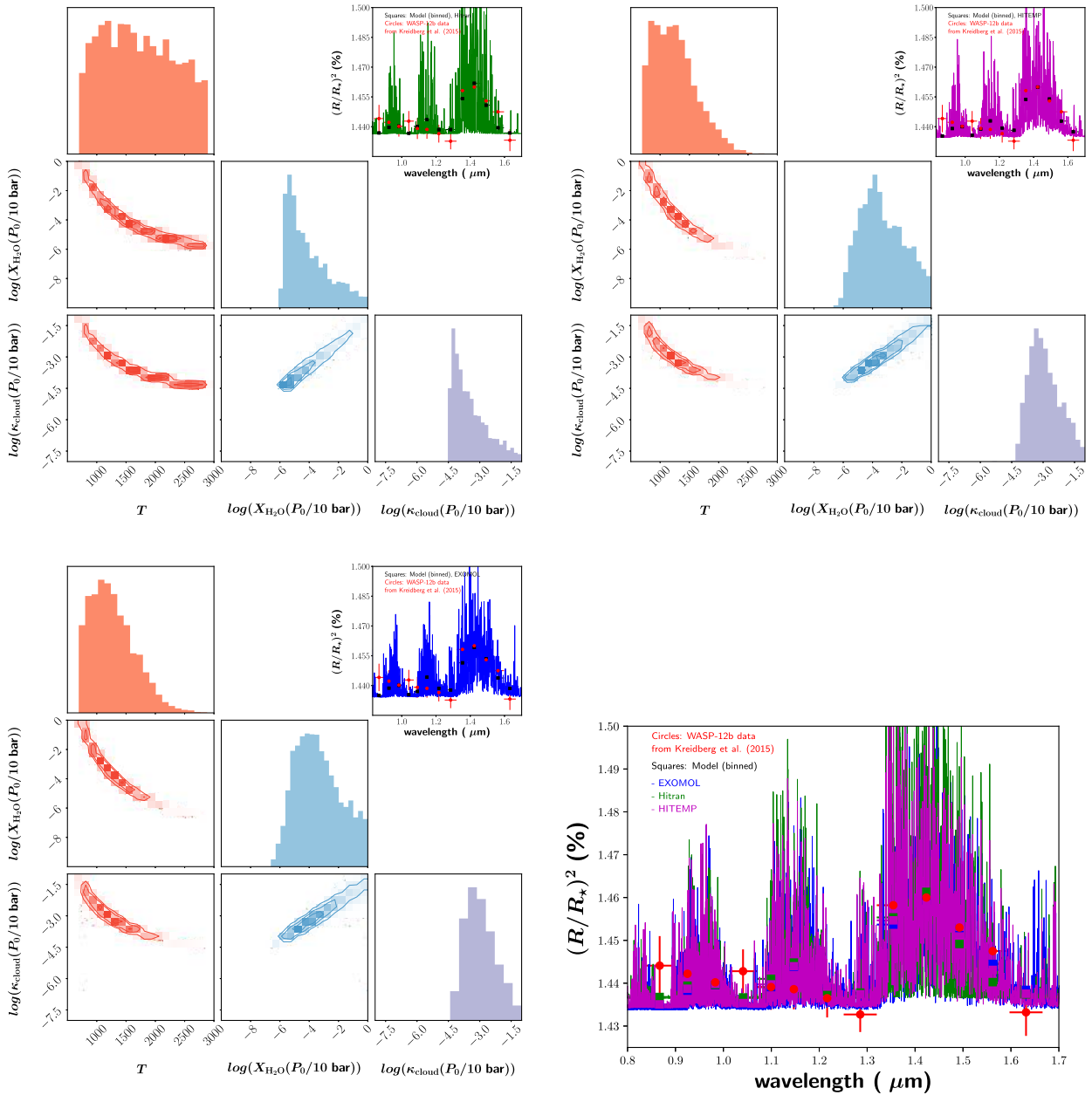
Figure 7. Same as Fig. 3, but comparing the use of `HITRAN` (top left panel) versus `HITEMP` (top right panel) versus `ExoMol` (bottom left panel) spectroscopic line lists for water. Additionally, the best-fitting spectra are compared in the bottom right panel.

addresses this issue somewhat, but it is widely accepted by the exoplanet community that `ExoMol` addresses this issue most completely to date (see Tennyson & Yurchenko 2017 for a review). With an equilibrium temperature in excess of 2500 K, WASP-12b is an ideal target for testing if discrepancies from retrievals arise from the use of different line lists. Yet, Fig. 7 shows us that the choice of line list for the water opacity is irrelevant at the present spectral resolution and signal-to-noise attainable for the WFC transmission spectrum of WASP-12b. It suggests that the retrievals performed on the other WFC3 transmission spectra are robust to the choice of spectroscopic line list. Despite this finding, we persist in using the `ExoMol` line list in order to dispel any notion that our results lack robustness.

3.1.5 Insensitivity to pressure broadening

Pressure broadening is an ill-defined source of uncertainty, because there is no first-principles theory to describe it (see discussion in Section 2.2). Nevertheless, we quantify its effect as the final test in this WASP-12b suite. Fig. 8 shows two retrievals with $P = 1$ mbar versus 1 bar, which span the conceivable range of pressures probed by the WFC3 transmission spectrum. The effects on the posterior distributions of the temperature, water mixing ratio and grey cloud opacity are minimal, even with a factor of 1000 difference in pressure between the pair of retrievals.

For the rest of the study, we will fix the pressure associated with pressure broadening at 10 mbar. The reasoning is that departures
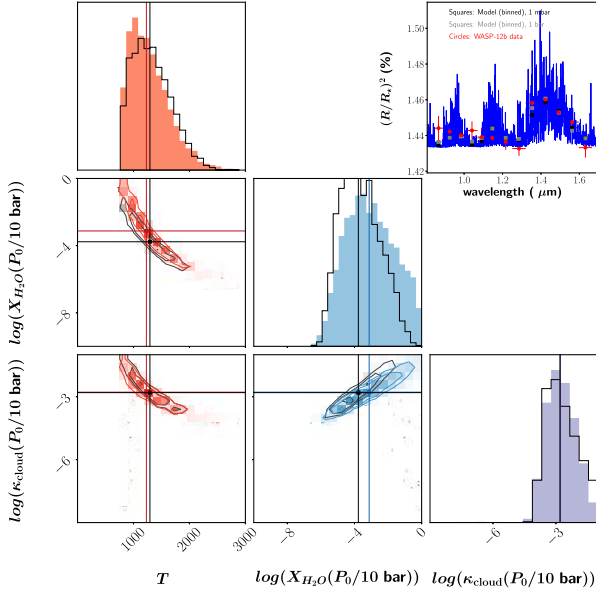
Figure 8. Same as Fig. 3, but elucidating the effect of pressure broadening. The posteriors are for $P = 1$ mbar, while the posteriors associated with $P = 1$ bar are overplotted as the solid curves. The vertical and horizontal lines represent the median values.
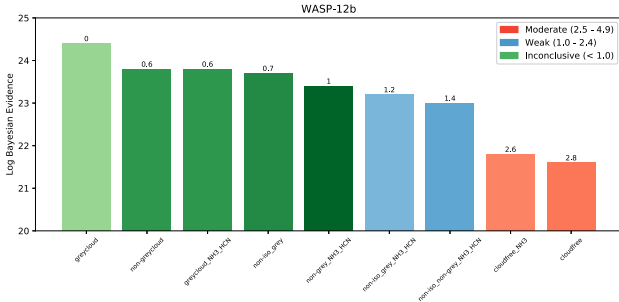
Figure 9. Logarithm of the Bayesian evidence and corresponding Bayes factor between each model compared to the best model (as indicated by the number on top of each bar). The entry marked by '0' is the best model, i.e. the model with the highest Bayesian evidence. The legend lists the correspondence between the Bayes factor and the strength or weakness of the evidence in favour of a given model (compared to the best model). For this set of Bayesian evidences only, we sample the opacities at 1 mbar, ignore CIA and use a fixed $m = 2.4m_H$ (see the text).

from this value will result in minor errors to the retrieved posterior distributions, which are subsumed as errors in the grey cloud opacity. Given that the exact functional form of pressure broadening cannot be specified from first principles, this is a reasonable approach, because it allows us to include pressure broadening in a more controlled way.

3.1.6 Comparison of Bayesian evidence

Following these tests, we analyse the WFC3 transmission spectrum of WASP-12b using a hierarchy of models: with and without clouds, grey versus non-grey clouds, isothermal versus non-isothermal, and with various permutations of the three molecules being present. Fig. 9 shows the Bayes factor for each model, which is the logarithm of the ratio of the Bayesian evidence of a given model compared to the best model. The value of the Bayes factor may be interpreted

as being weak, moderate, or strong evidence for the best model in favour of a given model (Trotta 2008). It may also be used to infer that the comparison is inconclusive, i.e. there is no evidence to favour one model over the other, if the Bayes factor between them is less than unity.

A few conclusions may be drawn from Fig. 9. First, cloud-free models are disfavoured. Secondly, there is weak evidence for non-isothermal behaviour, non-grey clouds, and the presence of HCN and/or NH_3, but overall the isothermal model with only water present and grey clouds is sufficient to fit the WFC3 transmission spectrum. In other words, there is no evidence for more complicated models to be favoured.

Again, note that the molecular opacities are sampled at 1 mbar, CIA is not included, m is fixed at $2.4m_H$ and we have fixed $R_0 = 1.79 R_J$ and $P_0 = 10$ bar. These assumptions will be lifted for WASP-12b later in the study.

3.2 Breaking the normalization degeneracy for cloud-free objects

3.2.1 Deriving $R_0(P_0)$: case study of WASP-17b

Heng (2016) previously concluded that the atmospheres of WASP-17b and WASP-31b are cloud-free based on optical transmission spectra recorded by STIS (Sing et al. 2016). This conclusion was based on the reasoning that the sodium and potassium lines may serve as diagnostics for cloudiness. The peaks of these resonant lines are hardly affected by clouds, but the line wings are, which makes the distance between the line peak and wing highly sensitive to the degree of cloudiness. If the optical transit chord is cloud-free, then we may associate the measured optical spectral slope with Rayleigh scattering by hydrogen molecules (H_2), which yields a direct measurement of the pressure scale height (Lecavelier des Etangs et al. 2008; Heng 2016),

$$H = -\frac{1}{4}\frac{\partial R}{\partial(\ln \lambda)}, \tag{10}$$

where λ is the wavelength. Such an approach is possible only because we have $\kappa = X_{H_2}m_{H_2}\kappa_{H_2}/m$, $X_{H_2} \approx 1$, and κ_{H_2} is known from first principles. If the optical transit chord is cloudy, then $\kappa = X_{cloud}m_{cloud}\kappa_{cloud}/m$. The cloud volume mixing ratio (X_{cloud}), composition of the cloud particles (and hence their mass, m_{cloud}), and opacity (κ_{cloud}) are now generally unknown and cannot be uniquely retrieved from either the optical or WFC3 transmission spectra.

We use WASP-17b as a working example, for which Heng (2016) previously estimated $H = 1896$ km using two data points from Sing et al. (2016) and $R_\star = 1.583 R_\odot$ (Southworth et al. 2012). In the current study, we fit a line to the optical spectral slope (comprising 15 data points) and derive $H = 1950$ km (not shown).

In a hydrogen-dominated atmosphere, the opacity associated with Rayleigh scattering alone is $\kappa = \sigma_{H_2}/m$. The cross-section for Rayleigh scattering by hydrogen molecules is (Sneep & Ubachs 2005),

$$\sigma_{H_2} = \frac{24\pi^3}{n_{ref}^2\lambda^4}\left(\frac{n_r^2-1}{n_r^2+2}\right)^2, \tag{11}$$

where $n_{ref} = 2.68678 \times 10^{19}$ cm^{-3} and the real part of the index of refraction is (Cox 2000)

$$n_r = 1.358 \times 10^{-4}\left[1 + 7.52 \times 10^{-3}\left(\frac{\lambda}{1\,\mu m}\right)^{-2}\right] + 1. \tag{12}$$
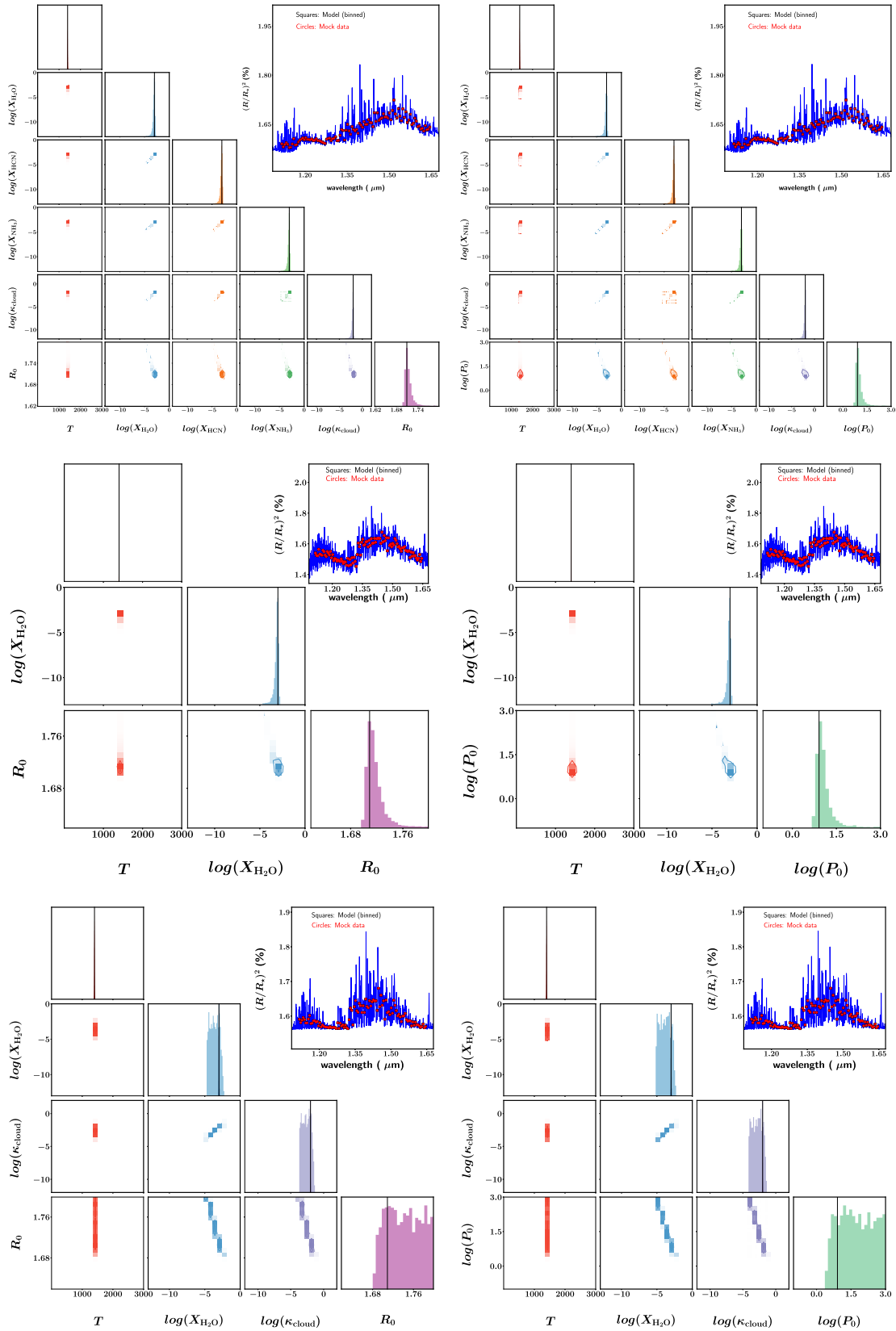
Figure 10. High-resolution (*JWST*-like) mock retrievals for WASP-17b using $R_0 = 1.709\,R_J$ and $P_0 = 8$ bar. The left column of retrievals hold P_0 fixed at 8 bar and fit for R_0, while the right column holds R_0 fixed at $1.709\,R_J$ and fit for P_0. The top, middle, and bottom rows are for three molecules with grey clouds, water only (cloud-free), and water only with grey clouds, respectively. All mock retrievals assume isothermal atmospheres and uncertainties of 10 ppm. Vertical lines indicate the true (input) values of the parameters.
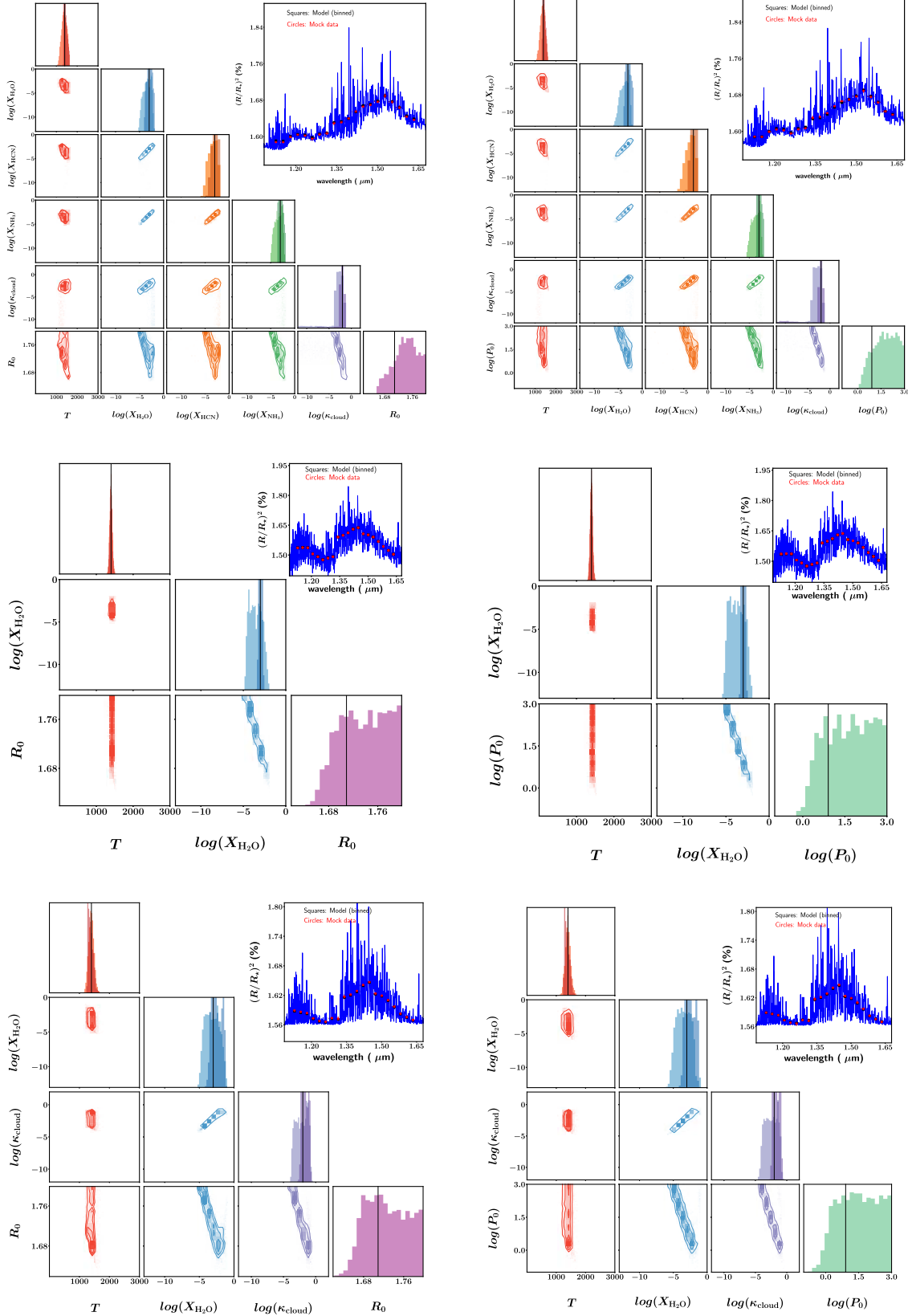
Figure 11. Same as Fig. 10, but for low-resolution (WFC3-like) spectra. Additionally, the darker posterior distribution in each panel corresponds to an additional retrieval in which R_0 (left column) or P_0 (right column) is held fixed at its true value (1.709 R_J or 8 bar). The uncertainties on each mock data point are assumed to be 50 ppm.
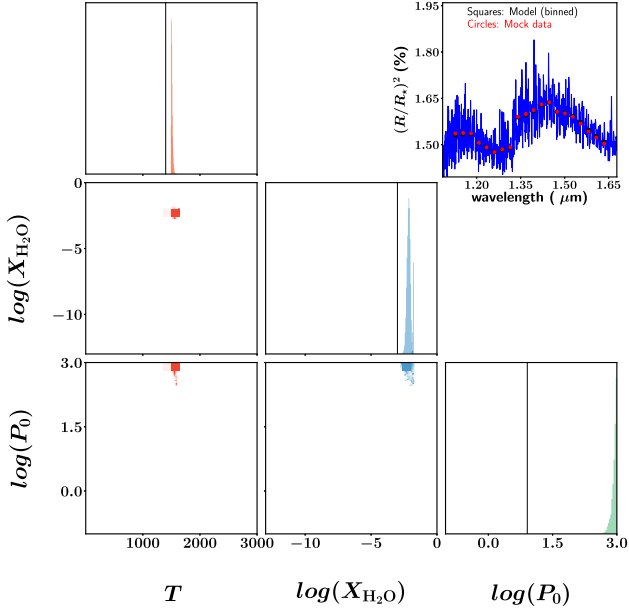
Figure 12. Low-resolution (WFC3-like) mock retrieval of WASP-17b for an isothermal, cloud-free atmosphere with water only. The mock spectrum is created using $R_0 = 1.709\,R_J$ and $P_0 = 8$ bar. The retrieval is performed with a fixed value of R_0 that is reduced by 10 per cent to $1.5381\,R_J$. The corresponding value of P_0 now lies outside of its assumed prior range (0.1–1000 bar). It is an illustration of how a bad assumption on R_0 can lead to an erroneous retrieval outcome. We emphasize that there is no unique value of R_0 one can assume, but it is related to P_0 via hydrostatic equilibrium. Retrievals with different R_0–P_0 pairs should yield the same outcome as long as the prior range of values of P_0 is set correctly.

Figure 14. Full atmospheric retrieval analysis of WASP-17b. Top panel: Posterior distributions of parameters for the isothermal model with water only and grey clouds, which has the highest Bayes factor among the model hierarchy. The vertical solid line is the median value of each posterior, while the vertical dotted lines are the 1σ uncertainties. The vertical dashed line is the best-fitting value of each posterior. Also shown is a second retrieval where the R_0–P_0 relationship is determined by the values derived using optical data (see the text). Bottom panel: Logarithm of the Bayesian evidence and corresponding Bayes factor between each model compared to the best model.

If the optical spectral slope is associated with H_2 Rayleigh scattering alone, then hydrostatic equilibrium allows us to derive a unique solution for P_0,

$$P_0 = \frac{0.56\,mg}{\sigma_{H_2}}\sqrt{\frac{H}{2\pi R_0}}\,\exp\left(\frac{R - R_0}{H}\right), \tag{13}$$

based on equation (2) and assuming that R_0 is associated with the part of the atmosphere that is opaque to both optical and infrared radiation.

For WASP-17b, we take $R = 1.890\,R_J$ at $\lambda = 0.405$ μm from the measurements of Sing et al. (2016). We then select a reference radius that is three orders of magnitude in pressure greater than that probed by WFC3,

$$R_0 = \bar{R}_{WFC3} - 6.908H, \tag{14}$$

where $\bar{R}_{WFC3}$ is the average value of the transit radius in the measured WFC3 bandpass. The preceding expression assumes hydrostatic equilibrium. For WASP-17b, we have $\bar{R}_{WFC3} = 1.897\,R_J$ and $R_0 = 1.709\,R_J$. Using the measured value of R and equation (13),

Figure 13. Low-resolution (WFC3-like) mock retrieval of WASP-17b for an isothermal atmosphere with grey clouds and all three molecules present. The pressure associated with CIA is a fitting parameter of the retrieval; its true value is $P_{CIA} = 0.1$ bar. The darker posterior distribution in each panel corresponds to an additional retrieval in which P_{CIA} is held fixed at its true value. Vertical lines indicate the true (input) values of the parameters.
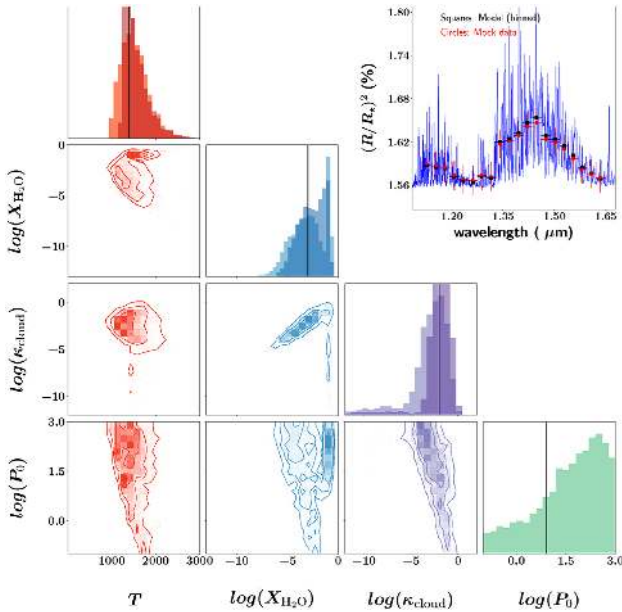
Figure 15. Additional low-resolution (WFC3-like) mock retrieval analysis of WASP-17b, but with larger uncertainties on each data point (200 ppm instead of 50 ppm). The darker posterior distribution in each panel corresponds to a second retrieval in which R_0 is held fixed at its true value (1.709 R_J). This pair of retrievals should be compared to the lower right panel of Fig. 11.
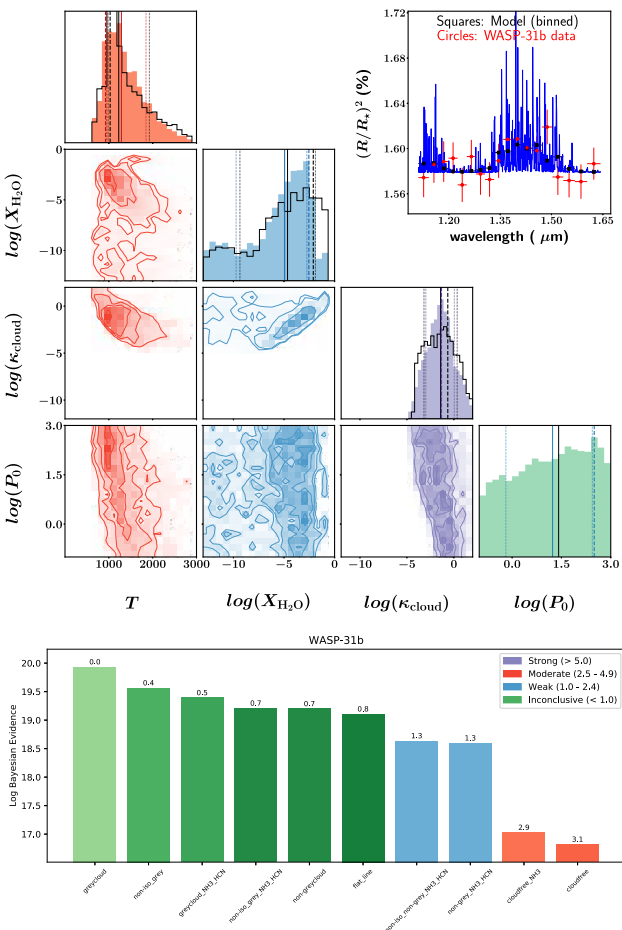

Figure 16. Same as Fig. 14, but for WASP-31b.

we estimate that $P_0 = 8$ bar. This means that the pressure probed in the WFC3 bandpass is, on average, about 8 mbar. We note that the pressure probed at $\lambda = 0.405 \, \mu m$ is about 10 mbar.

We do the same analysis for WASP-31b. We use $R_\star = 1.252 \, R_\odot$ (Anderson et al. 2011) and derived $H = 1619$ km. Heng (2016) previously derived $H = 1390$ km based on using $R_\star = 1.12 \, R_\odot$ (Anderson et al. 2011). We estimate $R_0 = 1.379 \, R_J$ and $P_0 = 26$ bar, based on $R = 1.547 \, R_J$ at $\lambda = 0.4032 \, \mu m$. This means that the WFC3 bandpass and the optical data point correspond to about 26 and 15 mbar, respectively.

These estimates are broadly consistent with our approach of assuming 10 mbar for the molecular opacities.

3.2.2 Mock retrievals of WASP-17b: breaking the normalization degeneracy

Using the derived $R_0 = 1.709 \, R_J$ and $P_0 = 8$ bar, we perform suites of mock retrievals to study if the normalization degeneracy may be broken. A uniform prior distribution of 1.619–1.799 R_J is set for R_0, while a log-uniform prior distribution of 0.1–1000 bar is set for P_0.

First, we create high-resolution mock transmission spectra with 100 data points that are representative of what will be possible with *JWST*. The uncertainty on each data point is assumed to be 10 parts per million (ppm). We explore pairs of retrievals in which R_0 is held fixed and P_0 is a fitting parameter, and vice versa. Secondly, we create a hierarchy of mock spectra to gain understanding into the retrieval outcomes: three molecules with grey clouds, water only with grey clouds and water only (cloud-free). All volume mixing ratios are set to 10^{-3} for illustration, with a grey-cloud opacity of $10^{-2} \, cm^2 \, g^{-1}$.

Fig. 10 shows the outcomes of six retrievals on high-resolution mock spectra. Unexpectedly, the peaks of the narrow posterior distributions of all six parameters, including R_0 or P_0, land exactly on the true values. The pair of cloud-free retrievals with water only also manages to locate the correct solution. In fact, the posterior distribution on the temperature is essentially a narrow spike with no width. This is straightforward to understand, because the temperature controls the 'stretch factor' in the spectrum and a unique solution is obtained by correctly fitting for the difference between the peaks and troughs of the spectrum. By contrast, R_0 or P_0 serves as a 'translation factor', which shifts the spectrum up or down in transit radius or depth without altering its shape. Further insight is obtained by examining a pair of retrievals with water only but with grey clouds present. The presence of grey clouds provides an extra degree of freedom in the system in the form of a constant spectral continuum. Grey clouds mute spectral features, which may be compensated by an increase in the volume mixing ratio of water, which is clearly seen in Fig. 10. Note that the normalization degeneracy is simultaneously present, as increases in X_{H_2O} and κ_{cloud} are negated by decreases in P_0 or R_0. The lower bound on the water mixing ratio in this pair of retrievals is artificial and is set by the chosen upper limit of our prior on R_0 or P_0. This pair of cloudy retrievals with water only allows us to understand that the degeneracy may be broken, even in the presence of clouds, if multiple molecules are present to provide additional information on the shape of the spectrum.

Fig. 11 shows the same suite of retrievals but for a low-resolution, WFC3-like spectrum with 20 data points. The uncertainty on each data point is assumed to be 50 ppm. For each of the six retrievals, we perform an additional retrieval in which R_0 or P_0 is held fixed at its true value (1.709 R_J or 8 bar). The lessons learnt and insights gleaned
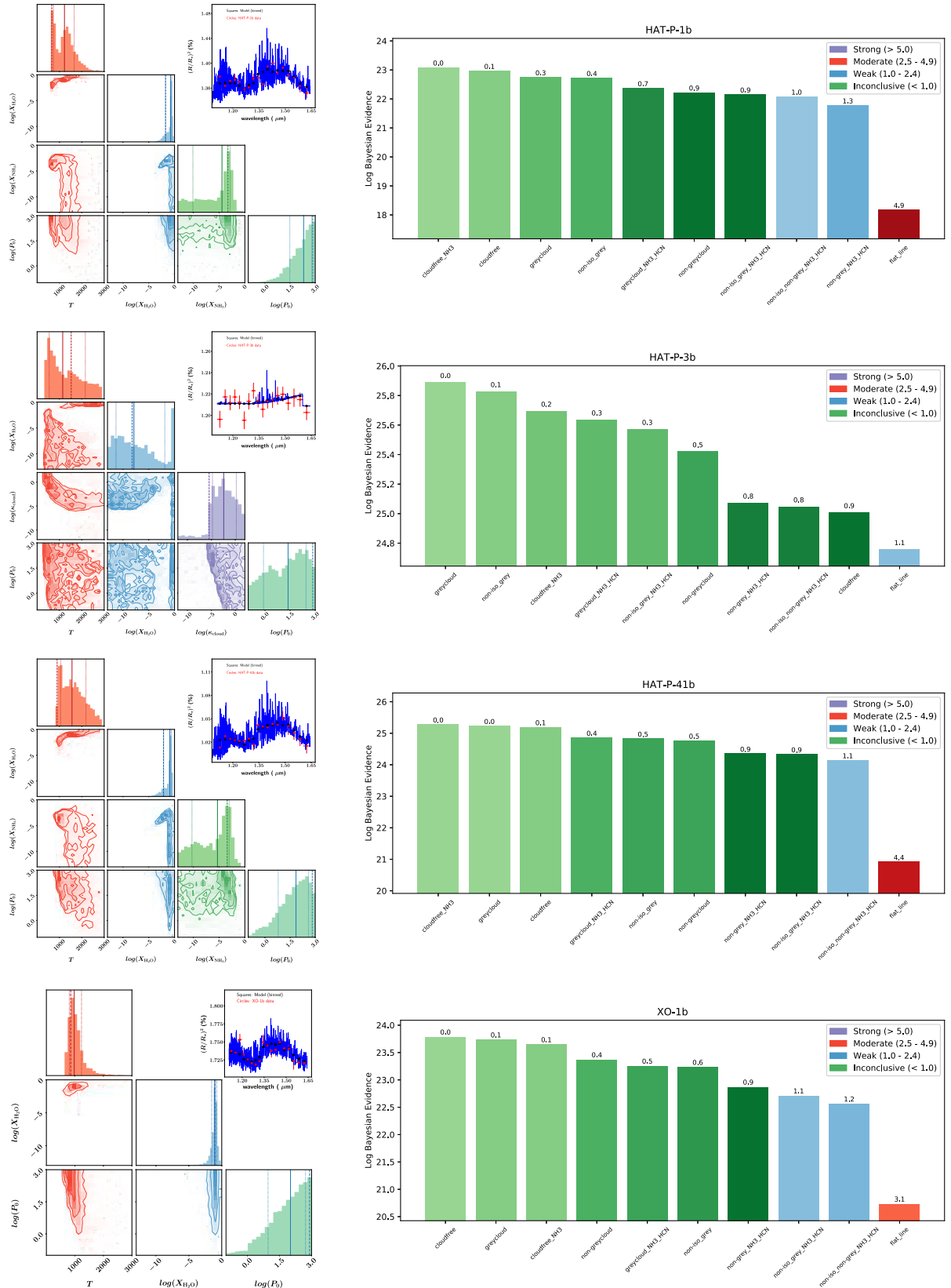
Figure 17. Left column: Posterior distributions and synthetic spectrum for the best model (as selected by the Bayesian evidence). Right column: Comparison of Bayesian evidence for objects for which it is not possible to distinguish between cloudy atmospheres containing water only versus cloud-free atmospheres with both water and ammonia present. The solid, dotted, and dashed vertical lines represent the median value, the 1σ uncertainties associated with the median and the best-fitting value of each posterior distribution, respectively. XO-1b is one of two objects with the highest Bayesian evidence for the cloud-free, isothermal model with water only (excluding the TRAPPIST-1 exoplanets).
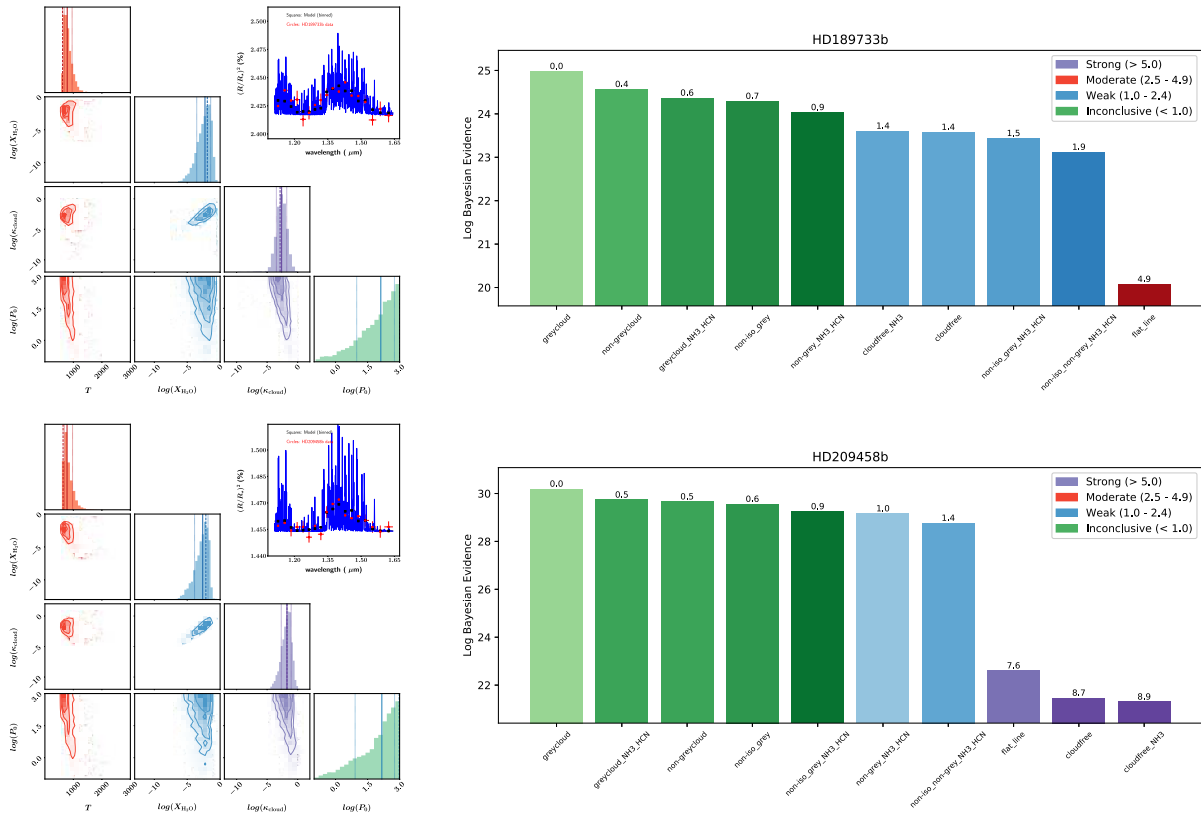
Figure 18. Same as Fig. 17, but for the prototypical hot Jupiters HD 189733b and HD 209458b.

from the high-resolution retrievals carry over to the low-resolution ones. Tight constraints are obtained on the temperature. For the volume mixing ratios of the molecules, constraints are obtained at the order-of-magnitude level that encompass the true values, but it is important to note that the lower bounds are artefacts of assuming an upper limit for the prior of R_0 or P_0. Unlike in the high-resolution regime, the low-resolution retrievals do not provide tight constraints on either R_0 or P_0.

The key lesson learnt is that, for meaningful retrieval outcomes to be obtained, we have to assume a reasonable range of prior values for R_0 or P_0. Since we find it easier to have an intuition about P_0, we will set the range of 0.1–1000 bar as the prior on P_0. It then becomes important to set a value of R_0 that corresponds to this range of P_0 values (see Section 3.2.3).

To illustrate this point, we perform an additional mock retrieval in which the value of R_0 is reduced and the corresponding P_0 value falls outside of the 0.1–1000 bar range. Fig. 12 shows that the posterior distribution for P_0 bumps up against the upper boundary of the prior distribution, which results in errors in the retrieved values of temperature and water mixing ratio.

3.2.3 Catalogue of R_0 values for other objects

For the other 36 objects in our sample, we first assume the WFC3 bandpass to probe a pressure of 10 mbar. We then use equation (14) to estimate the value of R_0 that corresponds to 10 bar (Table 3). The pressure scale height is estimated using $H = k_B T_{eq}/mg$, where k_B is the Boltzmann constant and T_{eq} is the equilibrium temperature (as

was done by Heng 2016). These R_0 values are then used as input in our retrievals.

We emphasise that while the value of R_0 is fixed to the tabulated value for each object, our retrievals ultimately use P_0 as a fitting parameter as justified by our tests in Section 3.2.2. The reason to use these values is to have R_0 be in approximately the range of values corresponding to 0.1–1000 bar, such that the retrieval will converge meaningfully.

3.2.4 Collision-induced absorption

As a final test on mock WASP-17b spectra, we consider an isothermal model atmosphere with all three molecules present, grey clouds, and CIA. We set the pressure associated with CIA at 0.1 bar, but allow the retrieval to treat this pressure as a fitting parameter (P_{CIA}). Fig. 13 shows that the retrieval outcome is insensitive to the retrieved value of P_{CIA}. Similar to our treatment of pressure broadening, we set the pressure associated with CIA to be 0.1 bar for the rest of the study with the reasoning that any deviations from this value may be visualized as errors that are subsumed into the grey cloud opacity. Figure 10 of Tsiaras et al. (2018) shows that CIA contributes a roughly flat continuum to the WFC3 spectrum.

3.2.5 Retrieval analysis of WASP-17b WFC3 transmission spectrum

Following our suite of tests, we now perform a full retrieval analysis on the WFC3 transmission spectrum of WASP-17b using a hierarchy of models. Additionally, we attempt to fit the spectrum with a flat line (one parameter only) and compute its corresponding
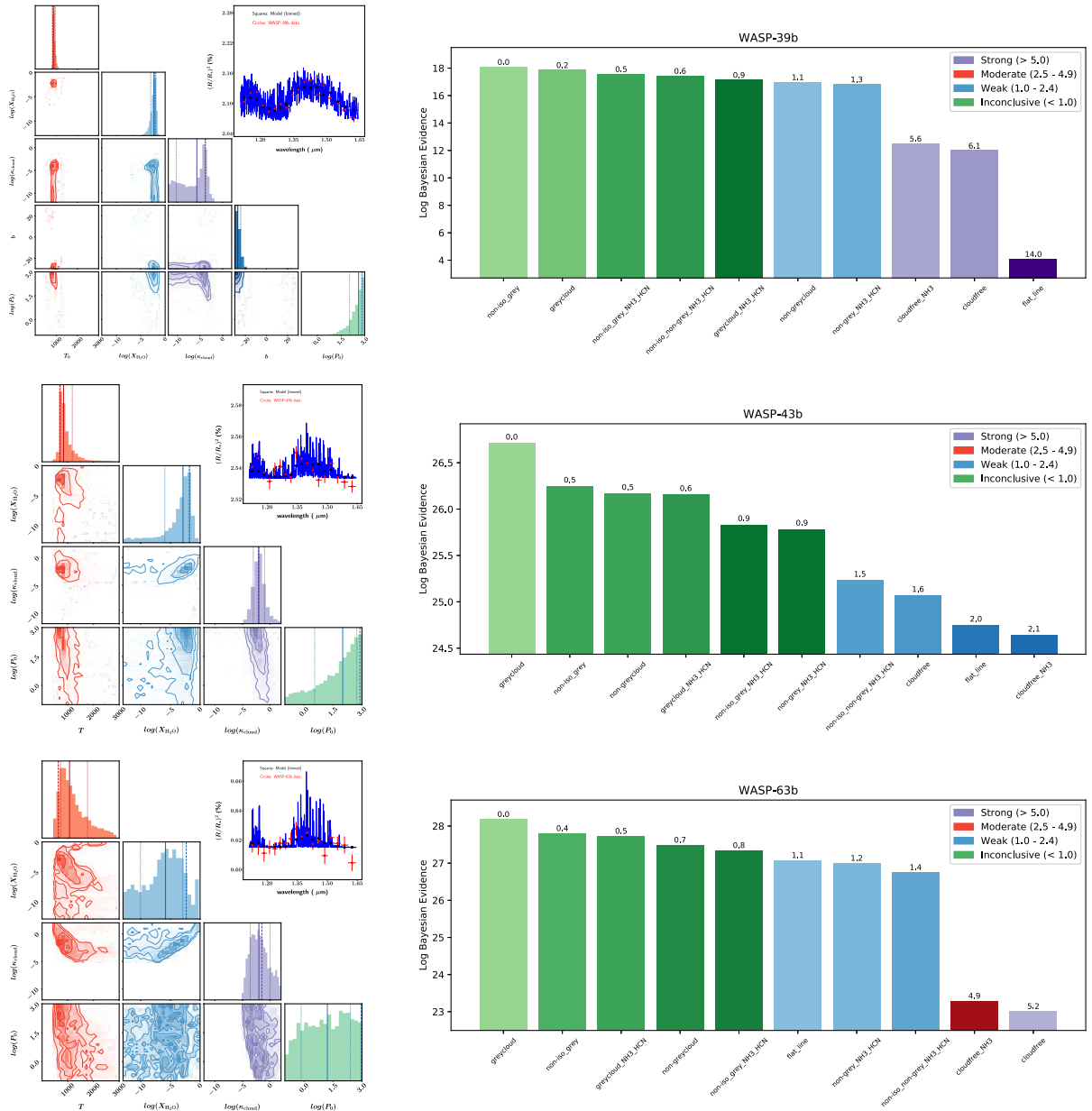
Figure 19. Same as Fig. 17, but for Early Release Science (ERS) objects: WASP-39b and WASP-43b (for *JWST*) and WASP-63b (for *HST*).

Bayesian evidence. We see that there is weak evidence against the flat-line fit, but several models are consistent with the data (Fig. 14). The isothermal model atmosphere with water only and grey clouds has the highest Bayesian evidence, which motivates us to display the posterior distributions of parameters associated with it in Fig. 14. Alongside this retrieval, we perform a second retrieval where $R_0 = 1.709\,R_J$ and $P_0 = 8$ bar as derived using the optical spectral slope. The posterior distribution for P_0 is only loosely constrained. The median value of P_0 is a factor of about 6 larger than its true value (8 bar); its best-fitting value hits the upper boundary of the prior distribution at 1000 bar.

Yet, despite this inaccuracy in retrieving P_0, the posterior distributions of the pair of retrievals agree well. This is somewhat surprising, because in our mock, low-resolution retrievals of WASP-

17b we discovered that the volume mixing ratio of water is prior-dominated on its lower bound (and corresponds to the upper limit set on the prior of P_0). To investigate this issue further, we ran an additional mock retrieval where the uncertainty on each data point is 200 ppm, instead of 50 ppm. Fig. 15 shows that the pair of retrievals now have posterior distributions that are more similar to each other, which implies that the retrieval with variable P_0 is no longer as prior-dominated because there is now a larger parameter space of possibilities available to fit the mock spectrum. However, the retrieval outcomes are still better (the posterior distributions are narrower) when the uncertainties are smaller. The lesson learnt is that the lower bounds to volume mixing ratios retrieved from WFC3 transmission spectra may (or may not) be prior-dominated, depending on the measurement uncertainties.
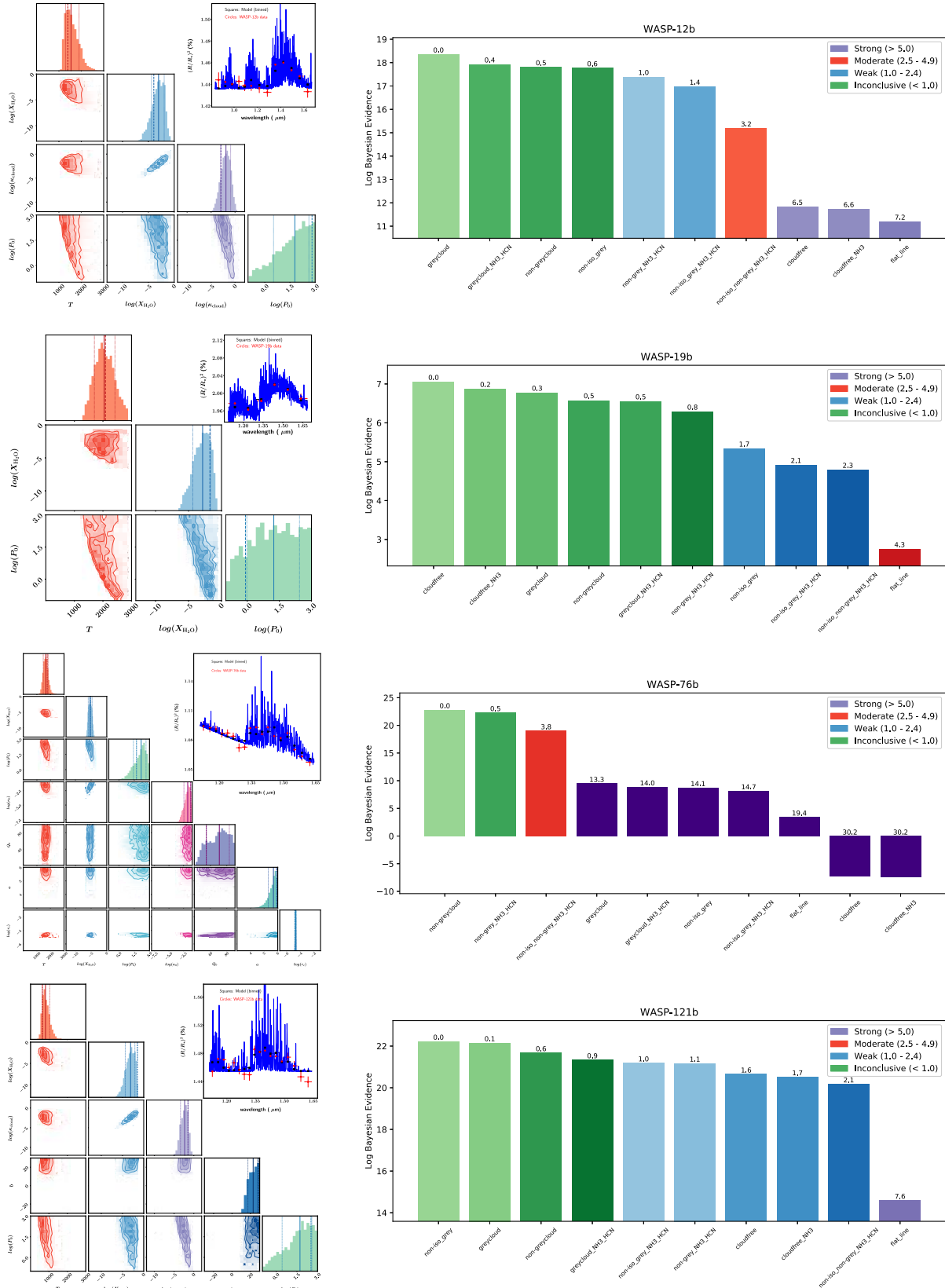
Figure 20. Same as Fig. 17, but for very hot Jupiters ($T_{eq} > 2000$ K): WASP-12b, WASP-19b, WASP-76b, and WASP-121b. WASP-19b is one of two objects with the highest Bayesian evidence for the cloud-free, isothermal model with water only (excluding the TRAPPIST-1 exoplanets). WASP-76b is one of two objects where non-grey clouds are needed to fit the data.
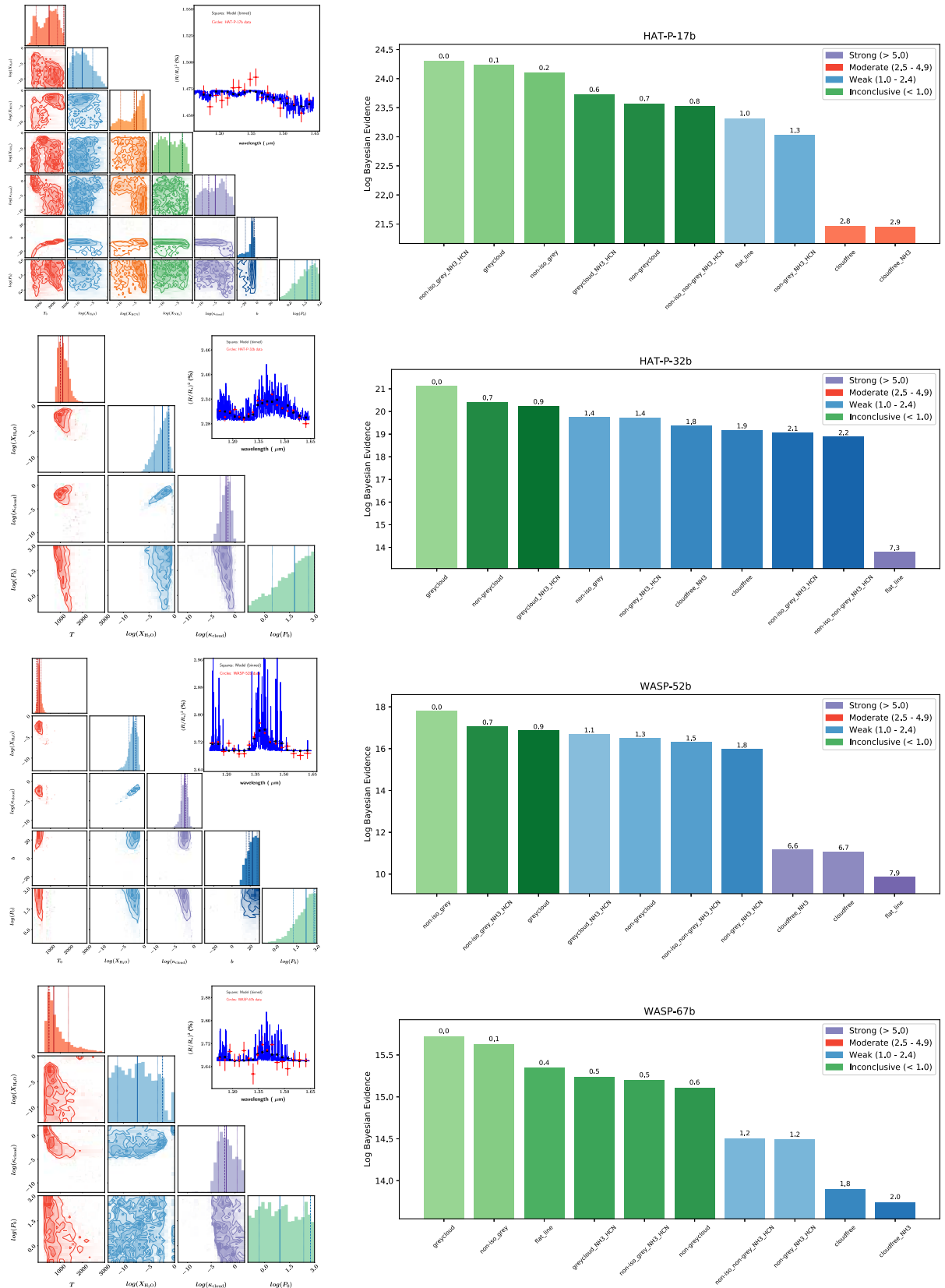
Figure 21. Same as Fig. 17, but for the rest of the hot Jupiters: HAT-P-17b, HAT-P-32b, WASP-52b, and WASP-67b.
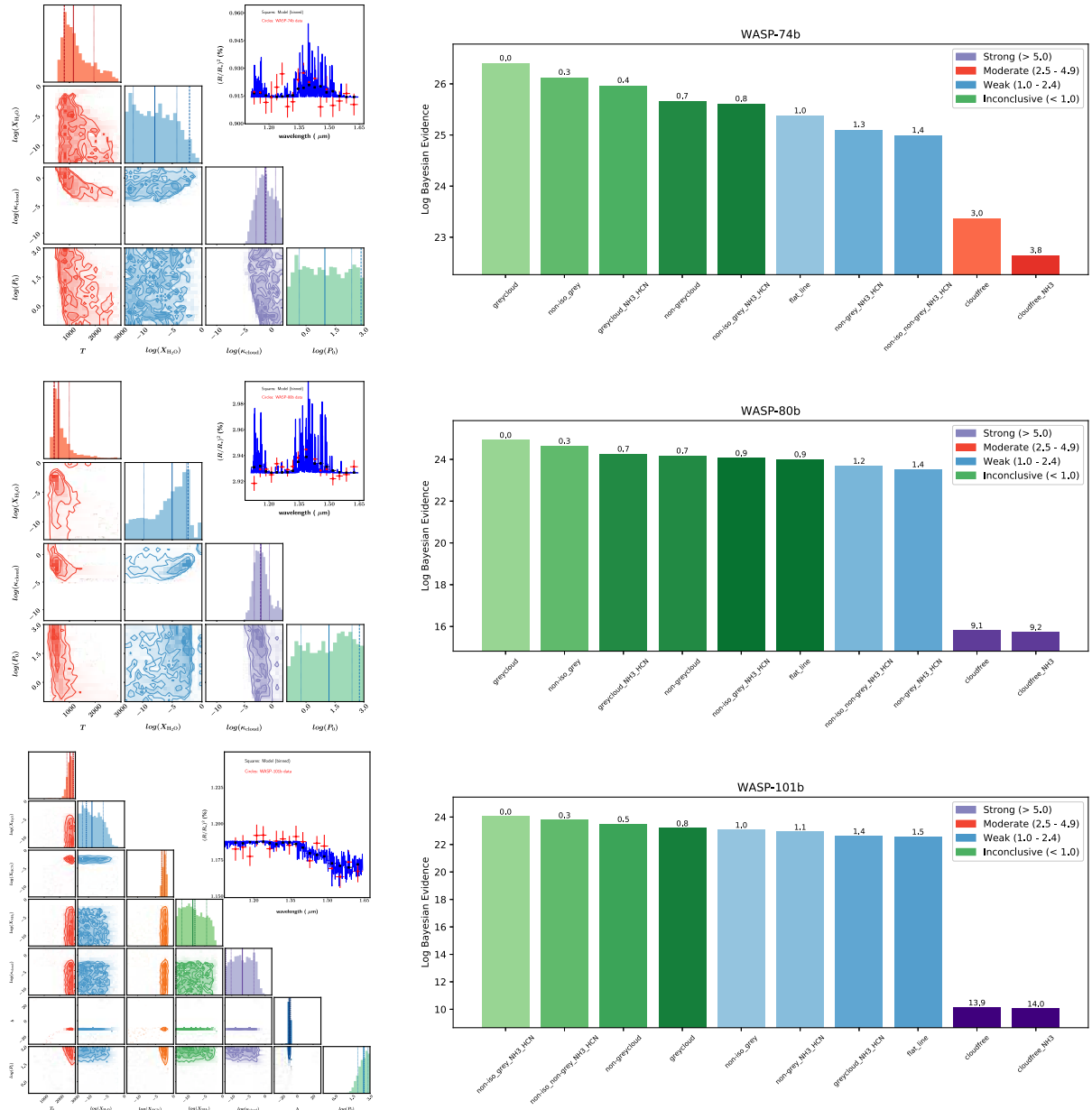
Figure 22. Continuation of Fig. 21 for the rest of the hot Jupiters: WASP-74b, WASP-80b, and WASP-101b. WASP-101b stands out as the only object for which HCN is significantly detected over water and ammonia.

3.2.6 Retrieval analysis of WASP-31b WFC3 transmission spectrum

Since WASP-31b is the other object in our sample where we can robustly derive R_0 and P_0 from the optical spectral slope, we subject it to the same retrieval analysis we performed for WASP-17b. In Fig. 16, we again subject the WFC3 transmission spectrum to a hierarchy of retrievals. Again, the isothermal model with water only and grey clouds has the highest Bayesian evidence. Two key differences are that the flat-line fit is not ruled out and that there is moderate evidence against cloud-free models. As before, we perform a second retrieval with $R_0 = 1.379\,R_J$ and $P_0 = 26$ bar. The median value of P_0 is about 16 bar and the best-fitting value of P_0 almost hits the prior boundary at 594 bar, but despite these outcomes the posterior distributions of parameters from the pair of retrievals agree surprisingly well.

Our general conclusions from studying WASP-17b and WASP-31b are that P_0 can be robustly used as a fitting parameter as long as one's guess for R_0 corresponds to the range of prior values set on P_0. Even if P_0 is not tightly constrained, the posterior distributions of the other parameters are, despite the low spectral resolution of the WFC3 transmission spectra.

3.3 Comparison of retrieval models for the other 36 WFC3 transmission spectra

Following our suites of tests in Sections 3.1 and 3.2.2, as well as our retrieval analyses of WASP-17b (Section 3.2.5) and WASP-31b (Section 3.2.6), we now apply our retrieval technique to the other 36 WFC3 transmission spectra in our sample. For each object, we use the value of R_0 listed in Table 3 and allow P_0 to be a fitting parameter (with a log-uniform prior between 0.1 and 1000 bar).
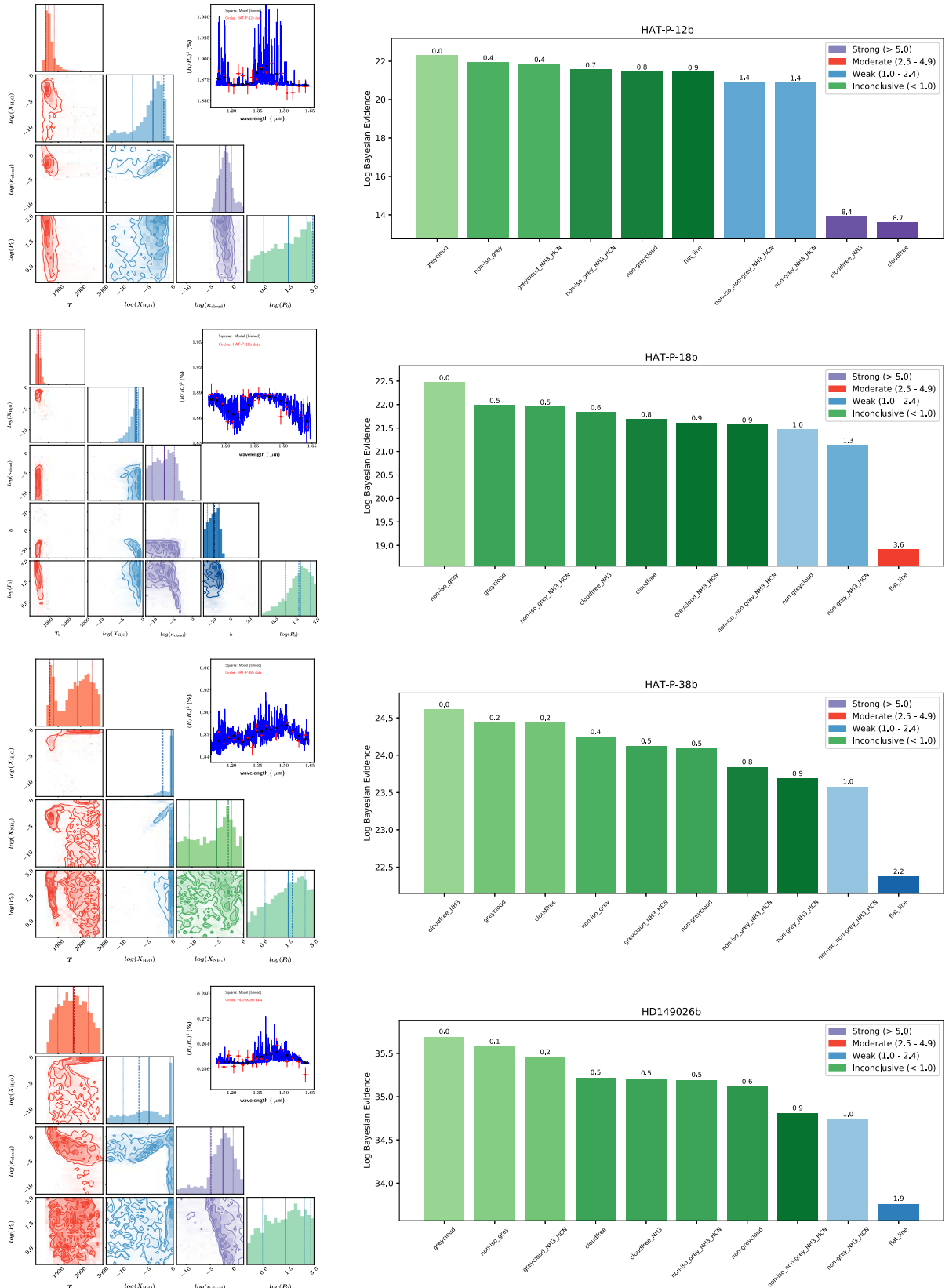
Figure 23. Same as Fig. 17, but for exo-Saturns (0.2–0.4M_J): HAT-P-12b, HAT-P-18b, HAT-P-38b, and HD 149026b.
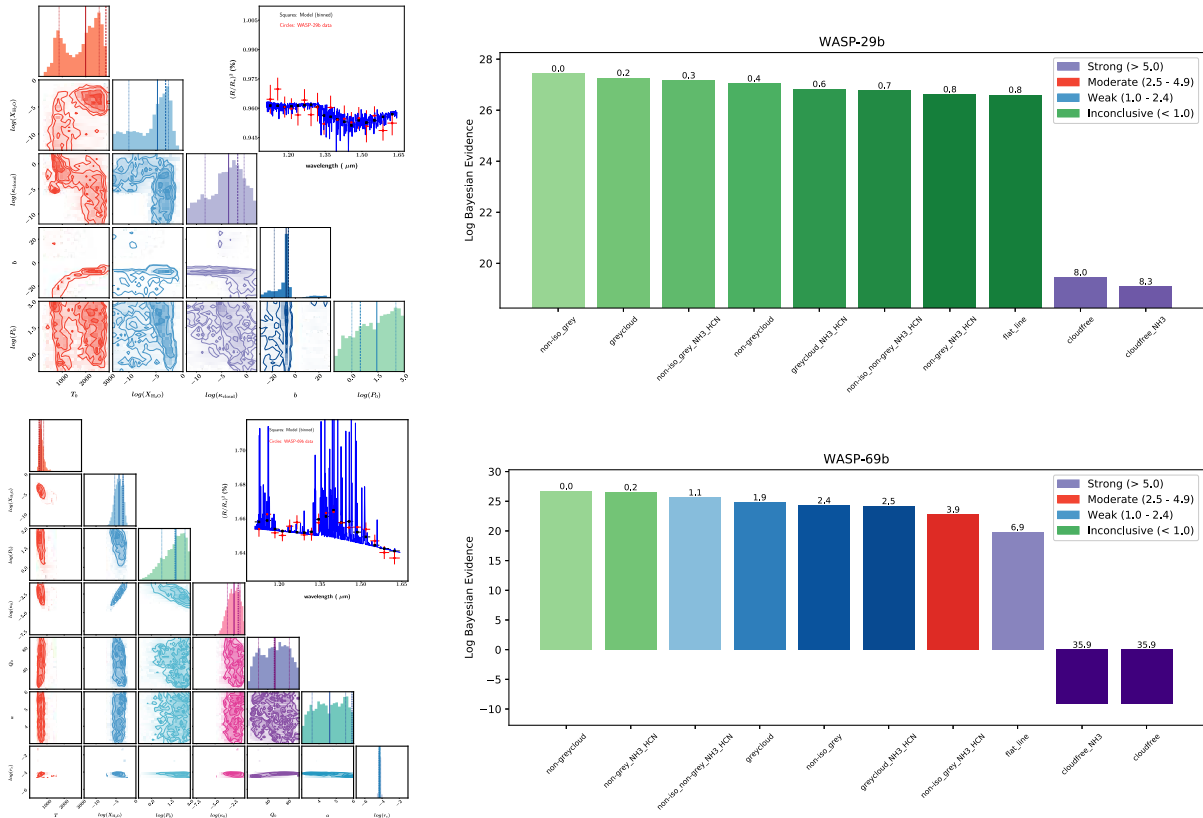
Figure 24. Continuation of Fig. 23 for exo-Saturns (0.2–0.4M_J): WASP-29b and WASP-69b. Additionally, WASP-69b is one of two objects where non-grey clouds are needed to fit the data.

The results are shown in Table 2, where the parameter values shown are the median and 1σ uncertainties from the best model (highest Bayesian evidence). Additionally, we ask several questions of the outcome. An atmosphere is deemed to be cloudy if all of the cloud-free models have Bayes factors of unity or more. Cloud-free atmospheres have *only* cloud-free models with Bayes factors of less than unity. If the models with Bayes factors of less than unity are a mixture of cloudy and cloud-free, then we tag the object with 'Maybe'. For non-grey clouds, our criterion is stricter: it refers to objects where *only* non-grey cloud models have Bayes factors of less than unity.

If the flat-line fit has a Bayes factor of less than unity, then we deem the retrieval to be inconclusive. In these cases, we do not report any retrieved properties of the WFC3 transit chord.

3.3.1 Ammonia may mimic cloudiness

By visual inspection of the atmospheric opacities (Fig. 2), we had suspected that it would be possible for ammonia to mimic the flattening of the spectral continuum bluewards of the 1.4 μm water feature. Note that the one-parameter flat-line fits are disfavoured. Fig. 17 shows four examples of objects (HAT-P-1b, HAT-P-3b, HAT-P-41b, and XO-1b) where the Bayes factor between the model with grey clouds and water only versus the cloud-free model with water and ammonia is below unity, indicating that there is a lack of Bayesian evidence to favour one model over the other (Trotta 2008). This interpretation holds also for WASP-17b (Fig. 14), WASP-19b (Fig. 20), HAT-P-38b and HD149026b (Fig. 23), and HAT-P-11b (Fig. 25).

With WFC3 transmission spectra, the cautionary tale is that cloudiness may be mimicked by the presence of ammonia and this occurs for 9 out of 38 objects in our sample.

3.3.2 Prototypical hot Jupiters

HD 189733b and HD 209458b are among the most studied hot Jupiters so far. In Fig. 18, we see that the WFC3 data definitively rule out cloud-free WFC3 transit chords for HD 209458b, and weakly rule cm out for HD 189733b. The simplest cloudy model, which is that of an isothermal atmosphere with grey clouds and water only, explains the WFC3 data well for both prototypical hot Jupiters. For HD 209458b, our retrieved temperature of ≈800 K is roughly consistent with MacDonald & Madhusudhan (2017), but our retrieved water abundance of $\log(X_{H_2O}) = -2.65^{+0.81}_{-1.24}$ is more than two orders of magnitude higher than their retrieved value of $-5.24^{+0.36}_{-0.27}$. It is unclear how to compare these values, because it is unclear how MacDonald & Madhusudhan (2017) have broken the normalization degeneracy. Unlike MacDonald & Madhusudhan (2017), we find a lack of evidence for the detection of either NH_3 or HCN. For example, the isothermal model with grey clouds and water only versus that with all three molecules have a Bayes factor of 0.5, indicating that one cannot favour one model over the other (Trotta 2008). For HD 189733b, we compare our results with those of Madhusudhan et al. (2014) in Section 4.5.

3.3.3 Early Release Science objects

WASP-39b and WASP-43b are among the Early Release Science (ERS) objects proposed for *JWST* (Batalha et al. 2017). WASP-63b
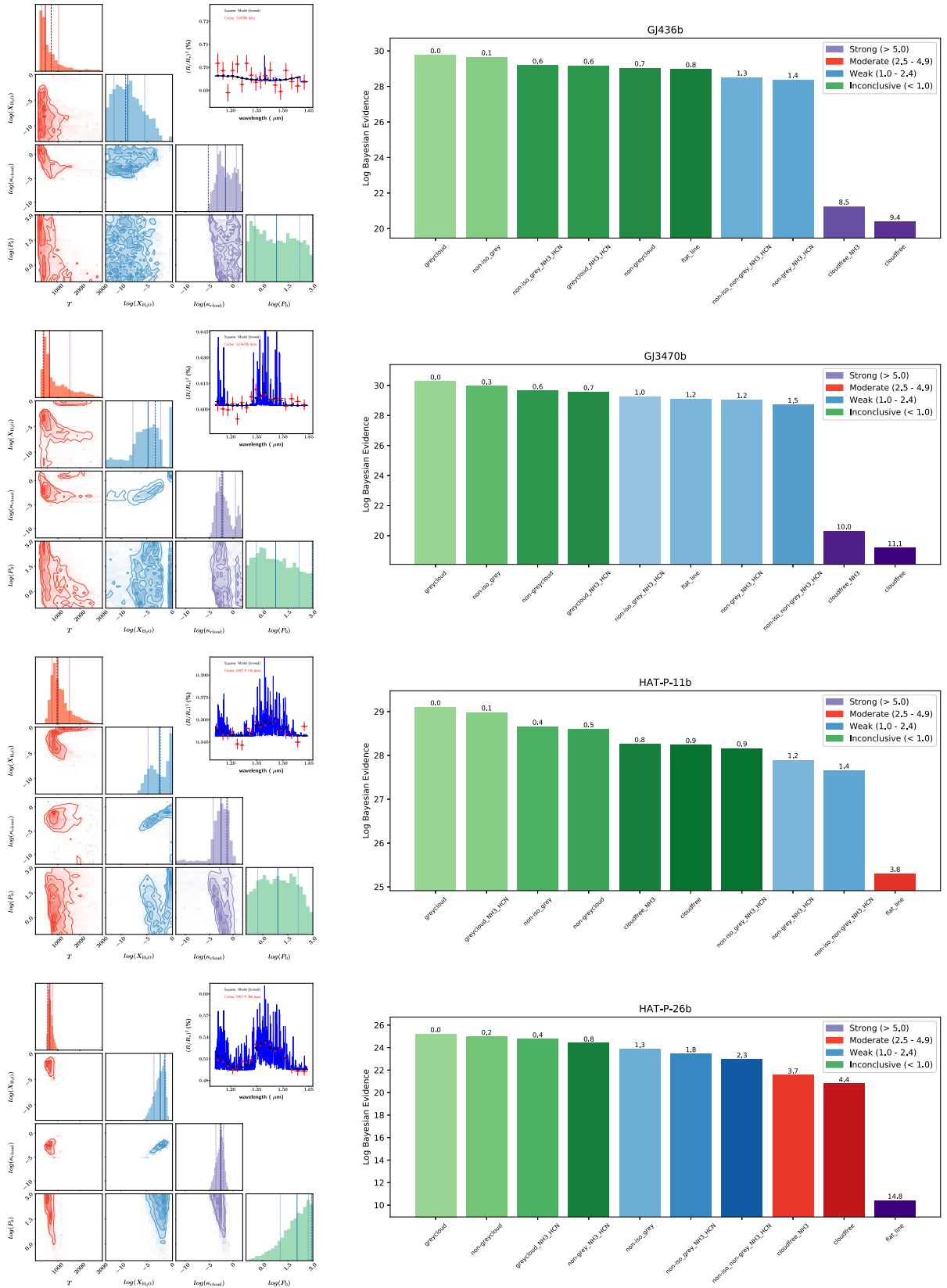
Figure 25. Same as Fig. 17, but for exo-Neptunes: GJ 436b, GJ 3470b, HAT-P-11b, and HAT-P-26b.
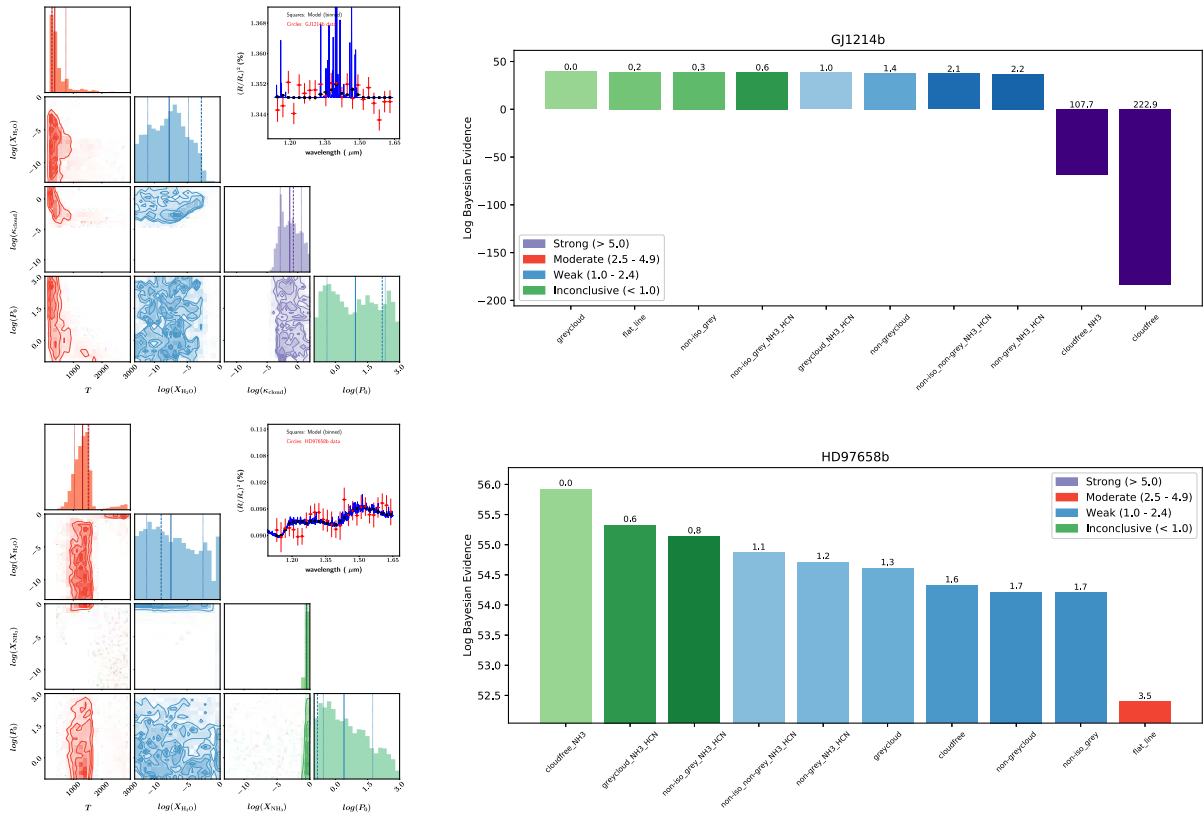
Figure 26. Same as Fig. 17, but for super Earths: GJ 1214b and HD 97658b. HD 97658b stands out as an object where ammonia is significantly detected but the abundance of water is essentially unconstrained.

is an ERS object for *HST* (Kilpatrick et al. 2017). Additionally, WASP-43b is one of the few hot Jupiters to have multiwavelength phase curves from *HST*, due to its sub-day orbit that circumvents the thermal breathing obstacle with *HST* (Stevenson et al. 2014). In Fig. 19 we see that none of the three objects are cloud-free in the WFC3 bandpass, and the simplest cloudy model fits the WFC3 data well. There is no definitive evidence for the detection of either HCN or NH_3. For WASP-63b, this is consistent with the analysis of Kilpatrick et al. (2017). For WASP-43b, our retrieved $\log(X_{H_2O}) = -2.89^{+1.13}_{-3.07}$ is broadly consistent with the $-3.6^{+0.8}_{-0.9}$ value reported by Kreidberg et al. (2014b), although it should be noted that Kreidberg et al. (2014b) included carbon dioxide, carbon monoxide, and methane in their analysis, while we excluded these molecules and included ammonia and hydrogen cyanide instead. Interestingly, Kreidberg et al. (2014b) reported a logarithm of the 'reference pressure' of $-2.4^{+0.6}_{-0.4}$ (pressure in bar), which is broadly consistent with the pressure of 10 mbar that we assume the WFC3 bandpass to probe. It is unclear how to compare the reference pressures between the two studies.

estimate on the carbon-to-oxygen ratio, unlike in Kreidberg et al. (2015). Our retrieved $\log(X_{H_2O}) = -3.02^{+1.09}_{-1.36}$ is broadly consistent with the $\sim 10^{-4}$–10^{-2} value reported by Kreidberg et al. (2015). In the case of WASP-19b, a cloud-free model with water only is a viable explanation – a rare occurrence in our sample. WASP-76b is an interesting object in that several scenarios are strongly ruled out: cloud-free with either water only or water and ammonia, the simplest cloudy model, etc. In fact, it seems to show strong evidence for any model featuring a non-grey cloud.

3.3.5 Other hot Jupiters

Figs 21 and 22 show the retrieval outcomes for seven other hot Jupiters. In all cases, cloud-free models are either unlikely or ruled out. All of these seven objects have WFC3 transmission spectra that may be explained by model atmospheres with grey clouds, meaning that non-grey clouds are not necessary to explain the data. WASP-101b is the only object where HCN is detected at significant levels, while only upper limits are obtained on the abundances of H_2O and NH_3.

3.3.6 Saturns

Figs 23 and 24 show the retrieval outcomes for six Saturn-mass (0.2–$0.4 M_J$) exoplanets. WASP-39b, an ERS object, also belongs to this category. With the exception of WASP-69b, the WFC3 transmission spectra are explained by the simplest cloudy model. WASP-69b requires non-grey clouds along its transit chord to explain the WFC3
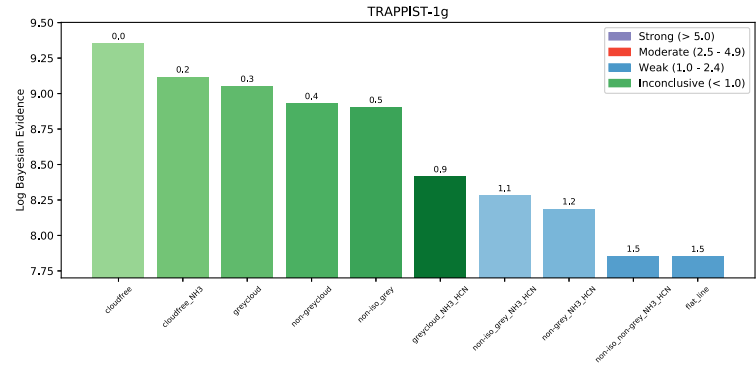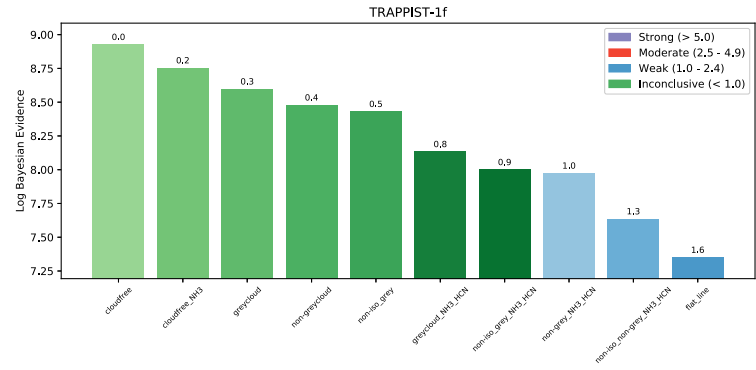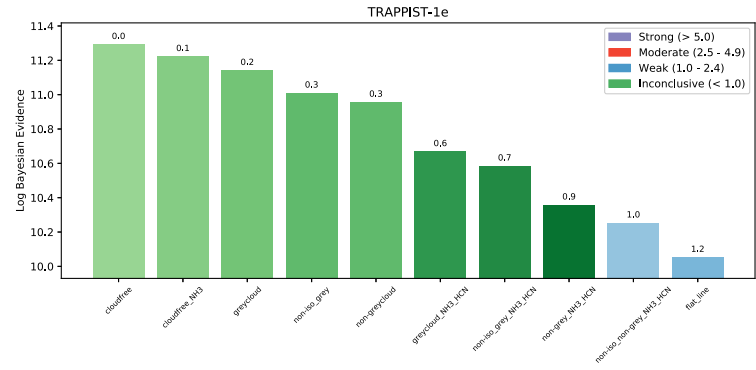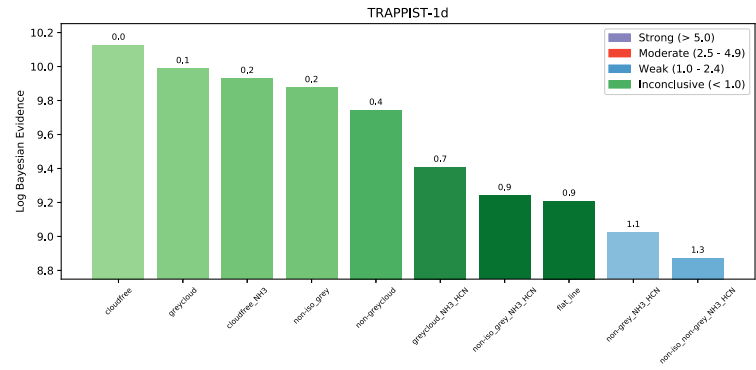
3.3.4 Very hot Jupiters

In our sample, four objects have equilibrium temperatures exceeding 2000 K: WASP-12b, WASP-19b, WASP-76b, and WASP-121b shown inn Fig. 20. For WASP-12b, the WFC3 transmission spectrum may be explained by models with HCN and NH_3 and also models with only water (i.e. these models all fall within Bayes factors of less than unity), which implies that we are unable to offer any
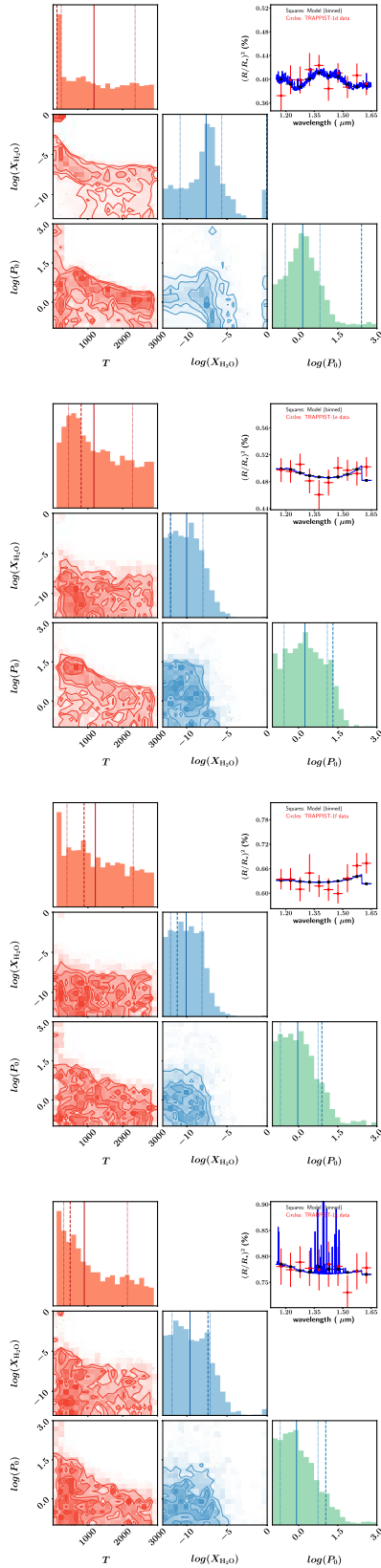
Figure 27. Same as Fig. 17, but for the TRAPPIST-1 exoplanets assuming Earth-like atmospheres ($m = 29\,m_{\rm H}$).
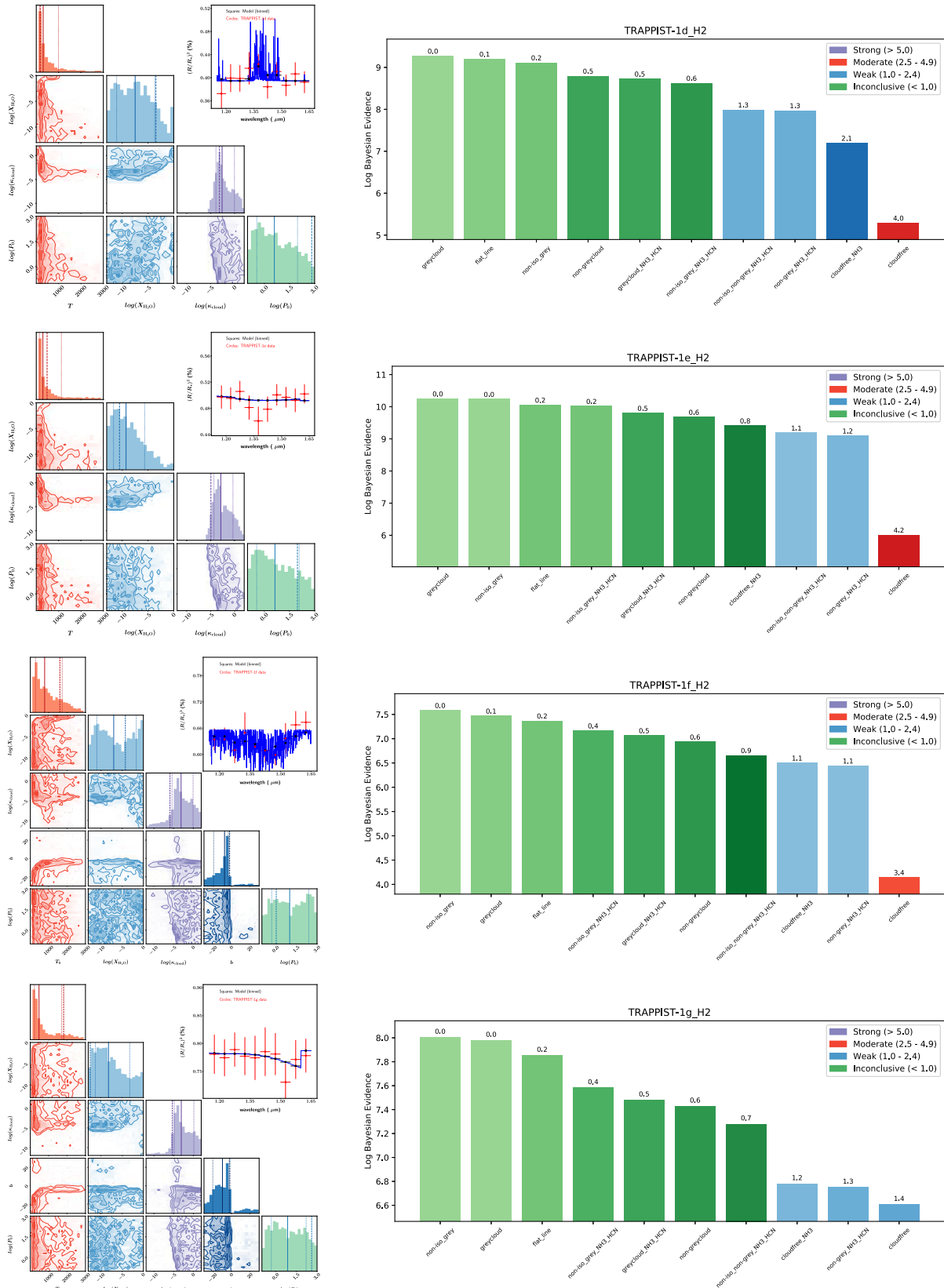
Figure 28. Same as Fig. 27, but assuming atmospheres dominated by molecular hydrogen (variable *m*), where the pressure scale height is larger by about an order of magnitude.
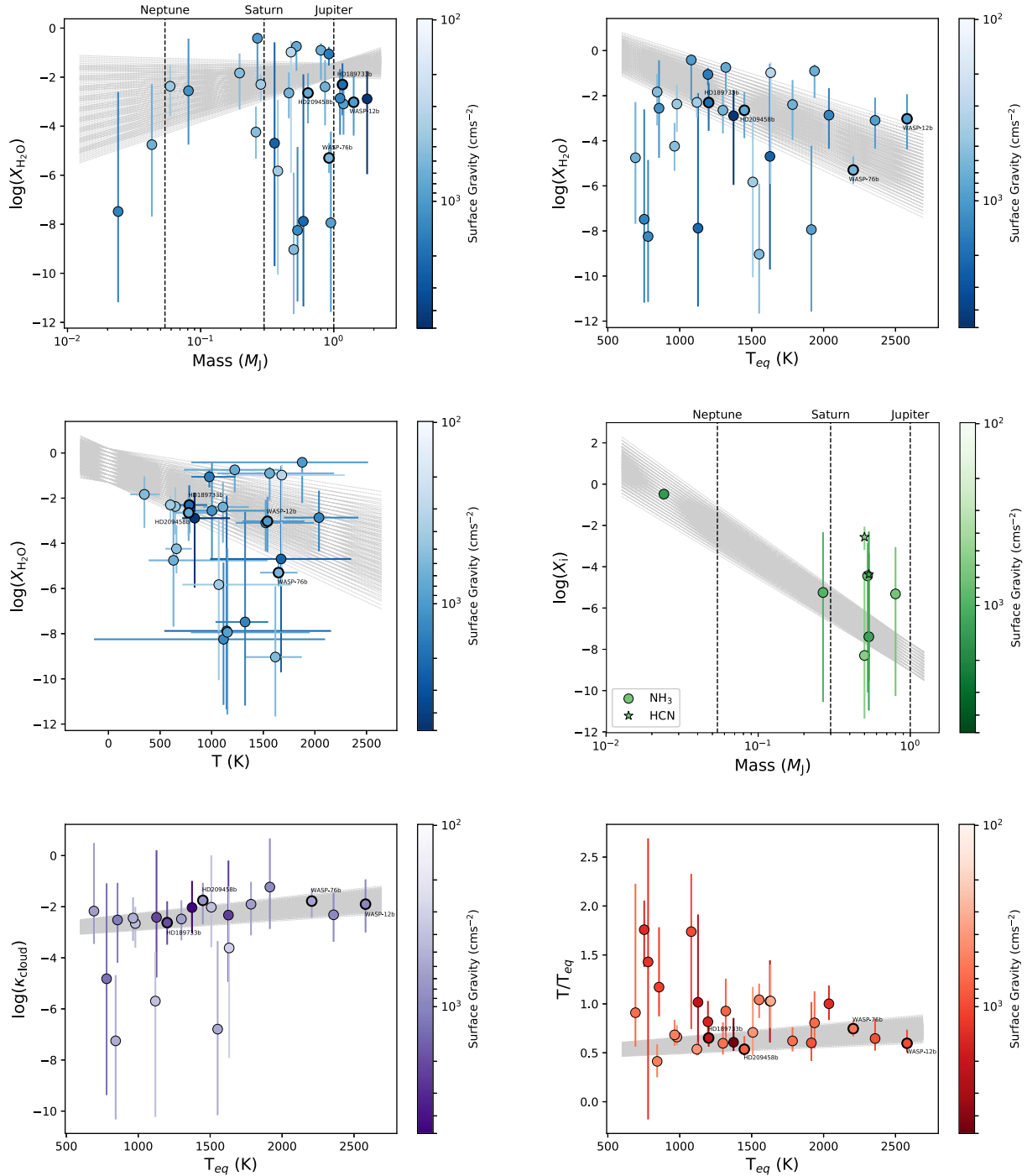
Figure 29. A search for trends between the retrieved atmospheric properties based on the best model (highest Bayesian evidence) of each object. We have excluded seven objects that can be adequately fitted by a one-parameter flat line. We have also excluded the four TRAPPIST-1 exoplanets. The family of lines in each panel shows the Monte Carlo fits of a two-parameter straight line (slope and constant offset). Top row: Water volume mixing ratio versus exoplanet mass (top left panel; slope of 2.56 ± 0.35) and equilibrium temperature (top right panel; slope of -0.00245 ± 0.00054 K^{-1}). Middle left panel: Water volume mixing ratio versus retrieved atmospheric temperature; slope of -0.00134 ± 0.00078 K^{-1}. Middle right panel: Ammonia and hydrogen cyanide volume mixing ratios versus exoplanet mass; slope of -5.28 ± 0.42. Bottom left panel: Grey cloud opacity versus equilibrium temperature; slope of 0.000517 ± 0.000135 K^{-1}. Bottom right panel: Ratio of retrieved to equilibrium temperatures versus equilibrium temperature; slope of 0.000103 ± 0.000038 K^{-1}.

data. For HAT-P-18b, HAT-P-38b, and HD 149026b, the isothermal cloud-free model with water only provides a viable explanation for the data; several other models also have Bayes factor of less than unity.

3.3.7 Neptunes

There is strong evidence against a cloud-free interpretation of the somewhat flat WFC3 transmission spectra of the exo-Neptunes GJ 436b and GJ 3470b (Fig. 25). For GJ 436b, this is consistent with the

findings of Knutson et al. (2014a). In fact, the WFC3 transmission spectrum of GJ 436b can simply be fit by a one-parameter flat line, rendering it impossible to report atmospheric properties in a meaningful sense. HAT-P-26b does not have a flat transmission spectrum and cloud-free interpretations are strongly ruled out (Bayes factor exceeding 5.0). Wakeford et al. (2017) previously analysed the transmission spectrum of HAT-P-26b, which includes STIS, WFC3, and *Spitzer* data spanning 0.5–5 μm, using a suite of models incorporating carbon monoxide, carbon dioxide, methane and water. Using the Bayesian information criterion, they disfavoured cloud-free models. Our WFC3-only analysis is consistent with the conclusion of Wakeford et al. (2017). The best model, in terms of the Bayesian evidence, is the simplest cloudy one: an isothermal atmosphere with grey clouds and water only, but a variety of cloud models have Bayes factors below unity compared to this best model.

3.3.8 Super Earths

Besides being a super Earth, GJ 1214b is the prototypical example of a flat transmission spectrum (Kreidberg et al. 2014a). The retrieval outcome in Fig. 26 corroborates this view and it is unsurprising that a one-parameter flat-line fit suffices. In our analysis, HD 97658b is inconclusively favoured by a cloud-free model with water and NH_3, though the quantities of ammonia required to match the data may be implausibly high. More data are needed to corroborate or refute this finding.

3.3.9 TRAPPIST-1 exoplanets

de Wit et al. (2018) previously measured somewhat flat WFC3 transmission spectra for TRAPPIST-1d, e, f, and g. We note an ongoing debate concerning the robustness of these measured WFC3 transmission spectra, as it has been argued that the shapes of the spectral bandheads may have been contaminated by starspots and faculae from TRAPPIST-1 (Ducrot et al. 2018; Morris et al. 2018; Rackham, Apai & Giampapa 2018). Nevertheless, we will analyse these spectra as given. de Wit et al. (2018) ruled out cloud-free, H_2-dominated atmospheres for TRAPPIST-1d, e, and f, but not for g. We wish to corroborate or refute this conclusion and also to go slightly further, by considering both Earth-like ($m = 29\,m_H$) or H_2-dominated (variable m as defined in equation 7) atmospheres in two separate suites of retrievals, shown in Fig. 27 and 28, respectively.

For Earth-like atmospheres, the WFC3 spectra are explained by the majority of the models in our hierarchy. With the exception of TRAPPIST-1d, there is weak evidence against the WFC3 transmission spectra being explained by a flat line. This is unsurprising (compared to the retrievals with H_2-dominated atmospheres), because for a nitrogen-dominated atmosphere the scale height is an order of magnitude smaller than for the H_2-dominated atmosphere, which implies that even small departures from a flat line require spectral features spanning several scale heights to explain the data. Overall, when Earth-like atmospheres are assumed, the retrieval analyses are inconclusive.

When H_2-dominated atmospheres are assumed, we rule out cloud-free atmospheres with water only for TRAPPIST-1d, e, and f. For all four exoplanets, the WFC3 transmission spectrum is adequately explained by a one-parameter flat-line fit, which implies that atmospheric properties cannot be meaningfully retrieved.

We do not consider arguments based on the evolution of the exoplanet or atmospheric escape, as they are out of the scope of the present study. Our inclusion of the TRAPPIST-1 exoplanets is for completeness and they will not be included in our analysis of the

trends associated with the water volume mixing ratios in Section 3.4. However, when compiling population statistics, we will include the outcomes only from the retrievals of the TRAPPIST-1 exoplanets assuming Earth-like atmospheres.

3.4 Trends

All of the techniques developed and tests performed in this study culminate in a singular result: to examine if there are trends in the retrieved atmospheric properties. In particular, we wish to examine if X_{H_2O} correlates with the equilibrium temperature (T_{eq}), retrieved temperature (T) or mass of the exoplanet (M). The equilibrium temperature is a proxy for the strength of insolation or stellar irradiation. Previous studies have plotted the 'metallicity' versus the exoplanet mass and claimed a correlation between the two quantities (Kreidberg et al. 2014b; Wakeford et al. 2017, 2018; Arcangeli et al. 2018; Mansfield et al. 2018; Nikolov et al. 2018).

In Fig. 29, we find little to no evidence for a correlation between X_{H_2O} and M, T_{eq} or T. If the abundance of water is assumed to be a direct proxy for the elemental abundance of oxygen (see Section 4.6), then this outcome runs contrary to previous claims. There is a lack of correlation between κ_{cloud} and T_{eq}, which has two implications. First, it suggests that our inferred X_{H_2O} values are not biased by the degree of cloudiness (or haziness) in these atmospheres. Secondly, the majority of atmospheric transit chords probed by WFC3 appear to have $\kappa_{cloud} \sim 10^{-2}$ cm^2 g^{-1} (corresponding to ~10 mbar), regardless of the surface gravity or strength of insolation. The lack of evidence for non-grey clouds implies that the particle radii are $r_c \gtrsim 0.1$ μm. Overall, these outcomes may be interpreted as the transit chords being affected by haze.[5] The ratio of the retrieved to the equilibrium temperatures (T/T_{eq}) appears to have a lower limit of about 0.5.

It is unclear how to relate our results to claimed correlations between the bulk metallicity of exoplanets and their masses based on the analysis of mass–radius relations (Miller & Fortney 2011; Thorngren et al. 2016).

4 DISCUSSION

4.1 Comparison to a previous retrieval study

It is natural to compare our study to Tsiaras et al. (2018), since the WFC3 transmission spectra of 30 objects from our sample are taken from it. Furthermore, some of the modelling choices made by Tsiaras et al. (2018) are the same as ours: isothermal transit chord, nested sampling. Our cloud models differ, because Tsiaras et al. (2018) use the formulation of Lee et al. (2013), which also allows for a smooth transition between the Rayleigh and large-particle regimes but predates Kitzmann & Heng (2018), and also assume a cloud-top boundary (which we do not). Furthermore, Tsiaras et al. (2018) include methane, carbon monoxide, carbon dioxide, titanium oxide (TiO), and vanadium oxide (VO) in their retrievals in addition to water and ammonia; they do not include hydrogen cyanide. By contrast, we only include water, ammonia, and hydrogen cyanide in our model hierarchy. Inevitably, these choices lead to differences in some of the retrieval outcomes.

Table 3 of Tsiaras et al. (2018) lists their retrieved water volume mixing ratios. For GJ 436b, HAT-P-12b, WASP-29b, WASP-31b, WASP-67b, and WASP-80b, we do not report any retrieved

---

[5]We adopt the planetary science definition of 'cloud' versus 'haze': the former is formed thermochemically, while the latter is formed photochemically.

atmospheric properties, unlike for Tsiaras et al. (2018), as the one-parameter flat-line fit is among the models with Bayes factors of less than unity. For GJ 3470b, HAT-P-1b, HAT-P-3b, HAT-P-11b, HAT-P-17b, HAT-P-18b, HAT-P-26b, HAT-P-32b, HAT-P-38b, HD 149026b, HD 189733b, HD 209458b, WASP-12b, HAT-P-41b, WASP-43b, WASP-52b, WASP-63b, WASP-69b, WASP-74b, WASP-101b, WASP-121b, and XO-1b, our retrieved water mixing ratios are broadly consistent with those of Tsiaras et al. (2018). For WASP-39b and WASP-76b, our retrieved water mixing ratios differ at the order-of-magnitude level compared to Tsiaras et al. (2018). Interestingly, these two objects also have the highest values of the ADI in the Tsiaras et al. (2018) sample of 30 objects.

Of particular interest is WASP-76b, which is one of two objects in our sample that requires a non-grey cloud to fit the data. Tsiaras et al. (2018) reported that their retrieval favours a cloud-free interpretation, because the non-flat spectral continuum blue-wards of the $1.4\,\mu m$ water feature may be fitted by the spectral features of TiO and VO. Tsiaras et al. (2018) remark that their retrieved $\log X_{TiO} \sim -2.5$ is 'likely unphysical'. Our retrieval yields $\log X_{H_2O} = -5.3 \pm 0.61$, which is inconsistent with the $\log X_{H_2O} = -2.7 \pm 1.07$ reported by Tsiaras et al. (2018). The WFC3 transmission spectrum of WASP-76b demonstrates that a wider wavelength range is required to resolve the degeneracy associated with these modelling choices, which will be provided by *JWST* spectra.

It is unclear why our retrieval outcome for WASP-39b differs from that of Tsiaras et al. (2018), because they did not publish the full set of posterior distributions for this object, unlike for WASP-76b in their fig. 11. For example, it is unclear if the high value of the ADI for WASP-39b translates into a cloud-free interpretation (which is the case for WASP-76b).

4.2 Is there evidence for non-grey clouds? Is cloud composition constrained?

Cloud models of varying sophistication have been employed in retrieval models. Our approach is somewhat different in that we include in our hierarchy of retrievals both grey and non-grey cloud models, as well as a one-parameter flat line. For 8 out of 38 objects, the WFC3 transmission spectrum is explained by a flat line. For 35 out of 38 objects, an isothermal grey cloud model with water only is sufficient to explain the data. Only WASP-69b and WASP-76b have WFC3 transmission spectra that require an explanation by model atmospheres with non-grey clouds along the transit chord. Otherwise, there is generally no evidence for non-grey clouds being present in the sample of 38 objects.

Since the cloud composition may only be inferred for non-grey clouds, this implies that the composition is generally unconstrained, which is consistent with the conclusion drawn by Tsiaras et al. (2018). Even for WASP-69b and WASP-76b, the parameter Q_0 is largely unconstrained because it spans the entire range of values set by the prior.

Given the retrieval outcomes, our approach to not consider patchy clouds (Line & Parmentier 2016) is justified. We have also shown that the normalization degeneracy may be broken without appealing to the more complicated patchy cloud model, which was invoked by MacDonald & Madhusudhan (2017) to break the degeneracy.

4.3 Is there evidence for non-isothermal transit chords?

For all 38 objects in our sample, we find a lack of strong Bayesian evidence to support non-isothermal transit chords probed by WFC3.

4.4 How prevalent is HCN or NH$_3$?

Based on the best model selected by the Bayesian evidence, we find that only six objects have tentative evidence for the detection of ammonia: HAT-P-1b, HAT-P-17b, HAT-P-38b, HAT-P-41b, WASP-101b, and HD 97658b. However, the retrieved value for HD 97658b is $\log(X_{NH_3}) = -0.48^{+0.19}_{-0.23}$, which may be unphysically high. This is unsurprising as our model contains no chemistry, so there is nothing to prevent unphysical values being retrieved. HAT-P-17b and WASP-101b also have tentative detections of hydrogen cyanide.

4.5 Subsolar water abundances in hot Jupiters?

Madhusudhan et al. (2014) previously analysed the WFC3 transmission spectra of HD 189733b, HD 209458b, and WASP-12b using cloud-free retrieval models and found $\log(X_{H_2O}) = -5.20^{+1.68}_{-0.18}$, $-5.27^{+0.65}_{-0.16}$, and $-5.35^{+1.85}_{-1.99}$, respectively. They concluded that the water abundances from these hot Jupiters are subsolar by about one to two orders of magnitude. By contrast, our retrievals find $\log(X_{H_2O})$ values that are several orders of magnitude higher: $-2.3^{+0.87}_{-1.26}$, $-2.65^{+0.81}_{-1.24}$, and $-3.02^{+1.09}_{-1.36}$, respectively. We estimate that $\log(X_{H_2O}) \approx -3.2$ assuming chemical equilibrium, solar abundance, and a pressure of 10 mbar, which suggests that our retrieved water abundances are super- rather than subsolar as claimed by Madhusudhan et al. (2014). The discrepancy arises from the retrievals of Madhusudhan et al. (2014) being cloud-free, while we have included a cloud model that smoothly transitions between the Rayleigh and large-particle regimes. It is consistent with the fact that cloud opacity diminishes the strength of the water feature, which may be negated by increasing X_{H_2O}.

4.6 What does the 'metallicity' mean when interpreting spectra of exoplanetary atmospheres?

Several published studies have plotted the 'metallicity' (in 'solar' units) versus the mass of the exoplanet with entries from the Solar system gas and ice giants overplotted (Kreidberg et al. 2014b; Wakeford et al. 2017, 2018; Arcangeli et al. 2018; Mansfield et al. 2018; Nikolov et al. 2018). As already elucidated by Heng (2018), there are several caveats to these plots. First, the 'metallicity' is predominantly O/H in these studies. Secondly, the 'mixing ratio of water at solar abundance' is a temperature- and pressure-dependent statement. Given a fixed value of O/H, the mixing ratio of water still depends on temperature and pressure. In other words, it is a *function* and not a number. Thirdly, the conversion factor between the water mixing ratio and O/H is not always unity and depends on the elemental abundances (O/H, C/H, etc.), carbon-to-oxygen ratio, temperature, pressure, photochemistry, atmospheric mixing, condensation, etc. For all of these reasons, we have chosen to present our retrieved water abundances as they are in Fig. 29, rather than convert them to O/H.

REFERENCES

Abramowitz M., Stegun I. A., 1970, Handbook of Mathematical Functions, 9th printing. Dover Publications, New York
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
Arcangeli J. et al., 2018, ApJ, 855, L30
Arfken G. B., Weber H. J., 1995, Mathematical Methods for Physicists, 4th edn. Academic Press, San Diego
Bakos G. A. et al., 2010, ApJ, 710, 1724
Barber R. J., Tennyson J., Harris G. J., Tolchenov R. N., 2006, MNRAS, 368, 1087
Barber R. J., Strange J. K., Hill C., Polyansky O. L., Mellau G. C., Yurchenko S. N., Tennyson J., 2014, MNRAS, 437, 1828
Batalha N. et al., 2017, JWST Proposal ID 1366, Cycle 0 Early Release Science
Benneke B., Seager S., 2012, ApJ, 753, 100
Benneke B., Seager S., 2013, ApJ, 778, 153
Bétrémieux Y., Swain M. R., 2017, MNRAS, 467, 2834
Biddle L. I. et al., 2014, MNRAS, 443, 1810
Boyajian T. et al., 2015, MNRAS, 447, 846
Brown T. M., 2001, ApJ, 553, 1006
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, 4th edn. Springer-Verlag, New York
de Wit J., Seager S., 2013, Science, 342, 1473
de Wit J. et al., 2018, Nat. Astron., 2, 214
Delrez L. et al., 2016, MNRAS, 458, 4025
Ducrot E. et al., 2018, preprint (arXiv:1807.01402)
Esposito M. et al., 2014, A&A, 564, L13
Feroz F., Hobson M. P., 2008, MNRAS, 384, 449
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, preprint (arXiv:1306.2144)
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Fortney J. J. et al., 2016, preprint (arXiv:1602.06305)
Fu G., Deming D., Knutson H., Madhusudhan N., Mandell A., Fraine J., 2017, ApJ, 847, L22
Griffith C. A., 2014, Phil. Trans. R. Soc. A, 372, 86
Grimm S. L., Heng K., 2015, ApJ, 808, 182
Grimm S. L. et al., 2018, A&A, 613, A68
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., 2016, ApJ, 826, L16
Heng K., 2017, Exoplanetary Atmospheres: Theoretical Concepts and Foundations. Princeton Univ. Press, Oxford
Heng K., 2018, RNAAS, 2, 128
Heng K., Kitzmann D., 2017, MNRAS, 470, 2972
Howard A. W. et al., 2012, ApJ, 749, 134
Hubbard W. B., Fortney J. J., Lunine J. I., Burrows A., Sudarsky D., Pinto P., 2001, ApJ, 560, 413
Huitson C. M. et al., 2013, MNRAS, 434, 3252
Iyer A. R., Swain M. R., Zellem R. T., Line M. R., Roudier G., Rocha G., Livingston J. H., 2016, ApJ, 823, 109

Johnson J. A. et al., 2008, ApJ, 686, 649
Kilpatrick B. M. et al., 2018, AJ, 156, 103
Kitzmann D., Heng K., 2018, MNRAS, 475, 94
Knutson H. A., Benneke B., Deming D., Homeier D., 2014a, Nature, 505, 66
Knutson H. A. et al., 2014b, ApJ, 794, 155
Kreidberg L. et al., 2014a, Nature, 505, 69
Kreidberg L. et al., 2014b, ApJ, 793, L27
Kreidberg L. et al., 2015, ApJ, 814, 66
Lavie B. et al., 2017, AJ, 154, 91
Lecavelier Des Etangs A., Pont F., Vidal-Madjar A., Sing D., 2008, A&A, 481, L83
Lee J.-M., Heng K., Irwin P. G. J., 2013, ApJ, 778, 97
Line M. R., Parmentier V., 2016, ApJ, 820, 78
Line M. R., Knutson H., Deming D., Wilkins A., Desert J.-M., 2013, ApJ, 778, 183
MacDonald R. J., Madhusudhan N., 2017, MNRAS, 469, 1979
Maciejewski G. et al., 2016, Acta Astron., 66, 55
Madhusudhan N., Seager S., 2009, ApJ, 707, 24
Madhusudhan N., Crouzet N., McCullough P. R., Deming D., Hedges C., 2014, ApJ, 791, L9
Mandell A. M., Haynes K., Sinukoff E., Madhusudhan N., Burrows A., Deming D., 2013, ApJ, 779, 128
Mansfield M. et al., 2018, AJ, 156, 10
Miller N., Fortney J. J., 2011, ApJ, 736, L29
Morris B. M., Agol E., Davenport J. R.A., Hawley S. L., 2018, ApJ, 857, 39
Nikolov N. et al., 2018, Nature, 557, 526
Rackham B. V., Apai D., Giampapa M. S., 2018, ApJ, 853, 122
Rothman L. S. et al., 1987, Appl. Opt., 26, 4058
Rothman L. S. et al., 1992, J. Quant. Spectrosc. Radiat. Transfer, 48, 469
Rothman L. S. et al., 1996, J. Quant. Spectrosc. Radiat. Transfer, 60, 665
Rothman L. S. et al., 2003, J. Quant. Spectrosc. Radiat. Transfer, 82, 5
Rothman L. S. et al., 2005, J. Quant. Spectrosc. Radiat. Transfer, 96, 139
Rothman L. S. et al., 2009, J. Quant. Spectrosc. Radiat. Transfer, 110, 533
Rothman L. S. et al., 2010, J. Quant. Spectrosc. Radiat. Transfer, 111, 2139
Rothman L. S. et al., 2013, J. Quant. Spectrosc. Radiat. Transfer, 130, 4
Sato B. et al., 2012, PASJ, 64, 97
Sing D. K. et al., 2016, Nature, 529, 59
Skilling J., 2006, Bayesian Anal., 1, 833
Sneep M., Ubachs W., 2005, J. Quant. Spectrosc. Radiat. Transfer, 92, 293
Southworth J. et al., 2012, MNRAS, 426, 1338
Stevenson K. B. et al., 2014, Science, 346, 838
Tennyson J., Yurchenko S. N., 2017, Mol. Astrophys., 8, 1
Thorngren D. P., Fortney J. J., Murray-Clay R., Lopez E. D., 2016, ApJ, 831, 64
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, Contemp. Phys., 49, 71
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
von Braun K. et al., 2012, ApJ, 753, 171
Wakeford H. R. et al., 2017, Science, 356, 628
Wakeford H. R. et al., 2018, AJ, 155, 29
Waldmann I. P., Tinetti G., Rocchetto M., Barton E. J., Yurchenko S. N., Tennyson J., 2015, ApJ, 802, 107
West R. G. et al., 2016, A&A, 585, A126
Yurchenko S. N., Tennyson J., 2014, MNRAS, 440, 1649
Yurchenko S. N., Barber R. J., Tennyson J., 2011, MNRAS, 413, 1828
Yurchenko S. N., Tennyson J., Barber R. J., Thiel W., 2013, J. Mol. Spectrosc., 291, 69
Yurchenko S. N., Al-Refaie A. F., Tennyson J., 2018, A&A, 614, A131

This paper has been typeset from a TEX/LATEX file prepared by the author.