Boletim da Sociedade Astronômica Brasileira, **35**, no. 1, 7-11

Effects of pressure dependency in exoplanetary atmospheric retrieval

A. Novais[1,3], C. Fisher[2], L. Ghezzi[1] & K. Heng[3,4,5]

[1] Observatório do Valongo, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil; e-mail: aline12@ov.ufrj.br
[2] Astrophysics, University of Oxford, Oxford, United Kingdom
[3] University Observatory, Ludwig Maximilian University of Munich, Munich, Germany
[4] ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland
[5] Department of Physics, University of Warwick, Warwick, United Kingdom

Abstract. Although the analysis of exoplanet atmospheres has become one of the most pertinent topics within planetary sciences, characterising these objects directly from their spectra is still a challenge. To interpret the observed spectrum of an exo-atmosphere, one can apply a technique known as atmospheric retrieval, i.e., fitting a model to this data in order to infer the properties of the atmosphere, such as temperature, chemical composition, and presence of clouds. This work considers a retrieval framework which includes H_2O as the main molecular opacity source, and optional opacity features such as additional molecules (e.g., CH_4, CO_2, CO), collision-induced absorption (CIA), Rayleigh scattering, and clouds. Furthermore, our retrieval code accounts for a non-isobaric transit chord, setting atmospheric opacity sources to a full pressure dependency. We perform non-isobaric retrievals in 38 *Hubble Space Telescope* Wide Field Camera 3 (WFC3) near-infrared transmission spectra to establish how the variation of pressure affects the estimation of atmospheric parameters. Our results show that, for WFC3 wavelength range and resolution, retrievals favour atmospheric models mainly including H_2O-only, and cloud-free or constant/grey clouds. We compare our findings with previous analyses in the literature, concluding that Rayleigh scattering is negligible in most of our retrievals, except in the ones where shorter-wavelength WFC3 data is available (e.g., WASP-12b). On the other hand, CIA is strongly dependent on pressure, therefore helping set H_2O abundances, in contrast to former isobaric studies. Additionally, we acknowledge the degeneracy between molecular abundances and the reference transit radius (i.e., where the atmosphere is optically thick), as pointed out by previous studies, may be artificially broken for cloud-free fits, but is still indirectly present in cloudy results. Finally, we suggest new approaches that could help identify additional atmospheric features imprinted in the spectra, considering data in complementary wavelengths, as well as retrieval analyses using higher-quality *James Webb Space Telescope* spectra.

Resumo. Apesar da análise de atmosferas de exoplanetas ter se tornado um dos tópicos mais pertinentes dentro das ciências planetárias, a caracterização desses objetos diretamente de seus espectros continua sendo um desafio. Para interpretar o espectro observado de uma exo-atmosfera, pode-se aplicar uma técnica conhecida como recuperação atmosférica, i.e., ajustando um modelo aos dados espectroscópicos de modo a inferir as propriedades daquela atmosfera, como temperatura, composição química, e presença de nuvens. Este trabalho considera recuperações atmosféricas que incluem H_2O como a principal fonte de opacidade molecular, e opções de opacidades adicionais como moléculas (e.g., CH_4, CO_2, CO), absorção induzida por colisão (CIA), espalhamento Rayleigh, e nuvens. Além disso, nosso código de recuperação considera raios de trânsito não-isobáricos, submetendo as fontes de opacidade a uma total dependência na pressão atmosférica. Realizamos recuperações não-isobáricas em 38 espectros de transmissão no infravermelho próximo, observados pela Wide Field Camera 3 (WFC3) do *Hubble Space Telescope*, para estabelecer como a variação da pressão afeta a estimativa dos parâmetros atmosféricos. Nossos resultados mostram que, considerando a faixa de comprimento de onda e resolução do WFC3, as recuperações favorecem modelos atmosféricos que incluem majoritariamente apenas H_2O, e sem nuvens ou com nuvens cinzas (opacidade constante). Comparamos nossos resultados a análises anteriores da literatura, concluindo que o espalhamento Rayleigh é desprezível na maioria das nossas recuperações, exceto quando dados da WFC3 para comprimentos de onda mais curtos estão disponíveis (e.g., WASP-12b). Por outro lado, CIA é fortemente dependente da pressão, auxiliando a estabelecer as abundâncias de H_2O, ao contrário de estudos isobáricos anteriores. Adicionalmente, reconhecemos que a degenerescência entre a abundância molecular e o raio de trânsito de referência (i.e., onde a atmosfera é opticamente espessa), como apontado por estudos anteriores, pode ser artificialmente quebrada para modelos sem nuvens, mas ainda está presente indiretamente em modelos com nuvens. Finalmente, sugerimos novas abordagens que podem ajudar a identificar características atmosféricas adicionais no espectro, considerando dados em comprimentos de onda complementares, bem como recuperações atmosféricas utilizando espectros de maior qualidade do *James Webb Space Telescope*.

Keywords. Planets and satellites: atmospheres – Planets and satellites: composition – Techniques: spectroscopic

1. Introduction

Transmission spectra can be powerful tools to help estimate characteristics of exoplanetary atmospheres, in particular the temperature structure, chemical abundances of atoms and molecules, clouds/hazes, as well as other properties. Planets with geometrically aligned orbits will pass periodically in front of their stars, which we refer to as a transit. Using the transit technique, if an observed object bears a significant atmosphere, it is possible to obtain an atmospheric transmission spectrum. As the planet passes in front of the star, part of the starlight passes through the planetary atmosphere, and may be absorbed and/or scattered by atmospheric species, depending on their wavelength. This absorbed light spectrum (taken during transit) is subtracted from the spectrum of the star alone (taken outside of transit), providing the transmission spectrum of that atmosphere (Seager 2010; Madhusudhan 2018).

A transmission spectrum can also be interpreted as the dependence of the transit depth on wavelength (Madhusudhan 2018), as can be seen in the atmospheric retrievals of this work. The role of atmospheric retrieval is to use a computational model to interpret an observed transmission spectrum, and therefore

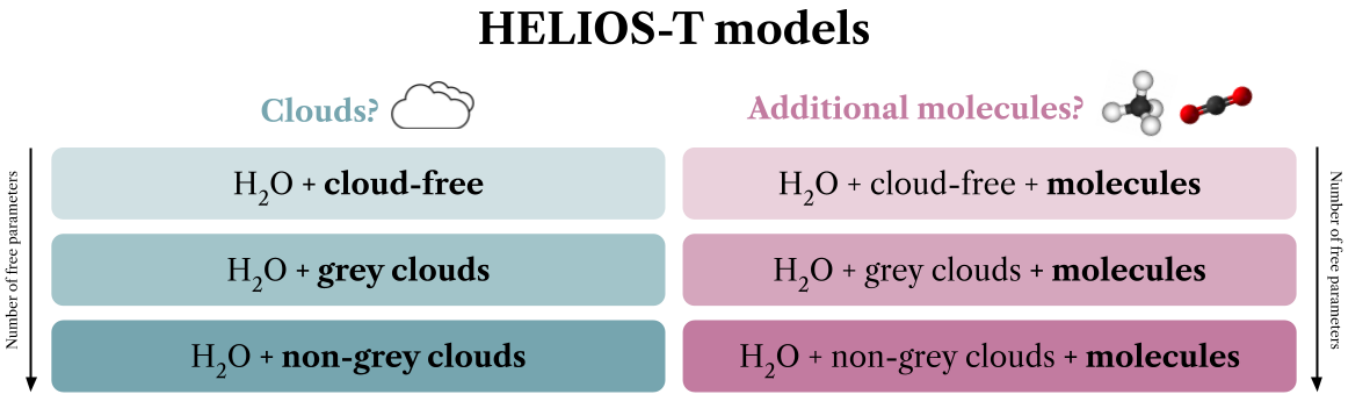

Figure 1. Schematic diagram presenting all possible models in `HELIOS-T`.

try to infer these atmospheric parameters. The present work performs retrievals by using transmission spectral data for a sample of 38 planets, providing estimates on atmospheric temperature, abundance of water, and additional parameters depending on the model.

Modelling the structure of a planetary atmosphere has always been a challenging task. Today, sophisticated computational resources are used by researchers to parameterize the structure and calculate properties of atmospheres, based on the available spectral data detected in the past 30 years. However, we frequently come across computational and observational limitations, caused by the low resolution and/or signal-to-noise ratio and short wavelength range covered by detected data, and by physical/geometric simplifications that need to be assumed to successfully run atmospheric models in a reasonable time, i.e., 1D model, atmosphere divided in several plane-parallel, isothermal, isobaric layers in hydrostatic equilibrium, etc.

To better represent such complex environments, the main purpose of this work is to develop more realistic atmospheric models, considering a retrieval framework. We create an upgraded version of the retrieval model by Fisher & Heng (2018), considering additional molecular opacities, new retrieved parameters, and physical effects not included in Fisher & Heng (2018). Then, we adapt this new version to assume a pressure-dependent (non-isobaric) isothermal model, in contrast to the previous isobaric, isothermal approach. Our intention is to compare the results retrieved by each model, and establish if the modifications in the isobaric code and/or the non-isobaric treatment are required to obtain more accurate values for the included atmospheric parameters, or if the simpler version of the model suffices.

2. Methodology

The current work is split in two parts: first, we studied the results of retrieving the stellar radius ($R_{\rm star}$) and the planetary surface gravity (g), i.e., analysing these two parameters also as posteriors, in contrast to the work of Fisher & Heng (2018), where they are taken as input parameters only (see Table 3 in their work). We used uncertainty values of $R_{\rm star}$ and g, considering their prior distributions as the range between the upper and lower limits of each parameter. We also tested the addition of new opacity sources in `HELIOS-T`, such as additional molecules and Rayleigh scattering. For the second part, our intention was to determine if the previously used isothermal, isobaric approach provided the best fit to the spectral data, in comparison to a non-isobaric numerical treatment. Therefore, we adapted the code to consider a non-isobaric, isothermal approach, using different opacities that may vary with pressure.

We have analysed 38 currently available *Hubble Space Telescope* Wide Field Camera 3 (WFC3) spectra in the near-infrared region (0.8–1.7 μm) using `HELIOS-T`. Our sample consists of the following planets: GJ 436b, GJ 1214b, GJ 3470b, HAT-P-1b, HAT-P-3b, HAT-P-11b, HAT-P-12b, HAT-P-17b, HAT-P-18b, HAT-P-26b, HAT-P-32b, HAT-P-38b, HAT-P-41b, HD 97658b, HD 149026b, HD 189733b, HD 209458b, TRAPPIST-1d, TRAPPIST-1e, TRAPPIST-1f, TRAPPIST-1g, WASP-12b, WASP-17b, WASP-19b, WASP-29b, WASP-31b, WASP-39b, WASP-43b, WASP-52b, WASP-63b, WASP-67b, WASP-69b, WASP-74b, WASP-76b, WASP-80b, WASP-101b, WASP-121b, and XO-1b. Spectral data were provided by Chloe Fisher and Kevin Heng, who previously studied the same 38-planet sample (see Fisher & Heng 2018 for references on each object).

Each planet was examined by a range of models, which consider different atmospheric parameters in the retrieval code. These parameters contemplate the temperature, treated as constant, a reference pressure/transit radius, and opacity sources, such as collision-induced absorption (CIA), molecules (H_2O, and HCN, NH_3, CH_4, CO, depending on the model), and clouds. While CIA and water are included in all models, the presence of clouds and the remaining molecules is optional. When models account for clouds, they are treated as grey or non-grey, with "grey" clouds referring to the concept of a grey atmosphere: a simplification where the opacity has no dependency on wavelength. The models can assume isobaric or non-isobaric atmospheres, with small differences in the molecular lists and cloud properties. A flat, one-parameter model is also assumed in both approaches, to represent a flat continuum with no spectral features. Figure 1 represents the different models accounted for in `HELIOS-T`, for isobaric or non-isobaric atmospheres.

All models except the flat line require the stellar radius $R_{\rm star}$, the planetary surface gravity g, and the reference transit radius R_0 as input parameters. The reference transit radius is defined as the radius at which the planet is optically thick, and has a reference pressure of P_0 (de Wit & Seager 2013; Griffith 2014). Calculation of the transit chord for low-resolution WFC3 data requires the values of R_0 and P_0, along with the scale height and opacity values (see Heng & Kitzmann 2017; Fisher & Heng 2018; Novais et al., submitted to MNRAS, for equations and details), causing a dependency relationship between these parameters (also referred to as a normalization degeneracy, see Section

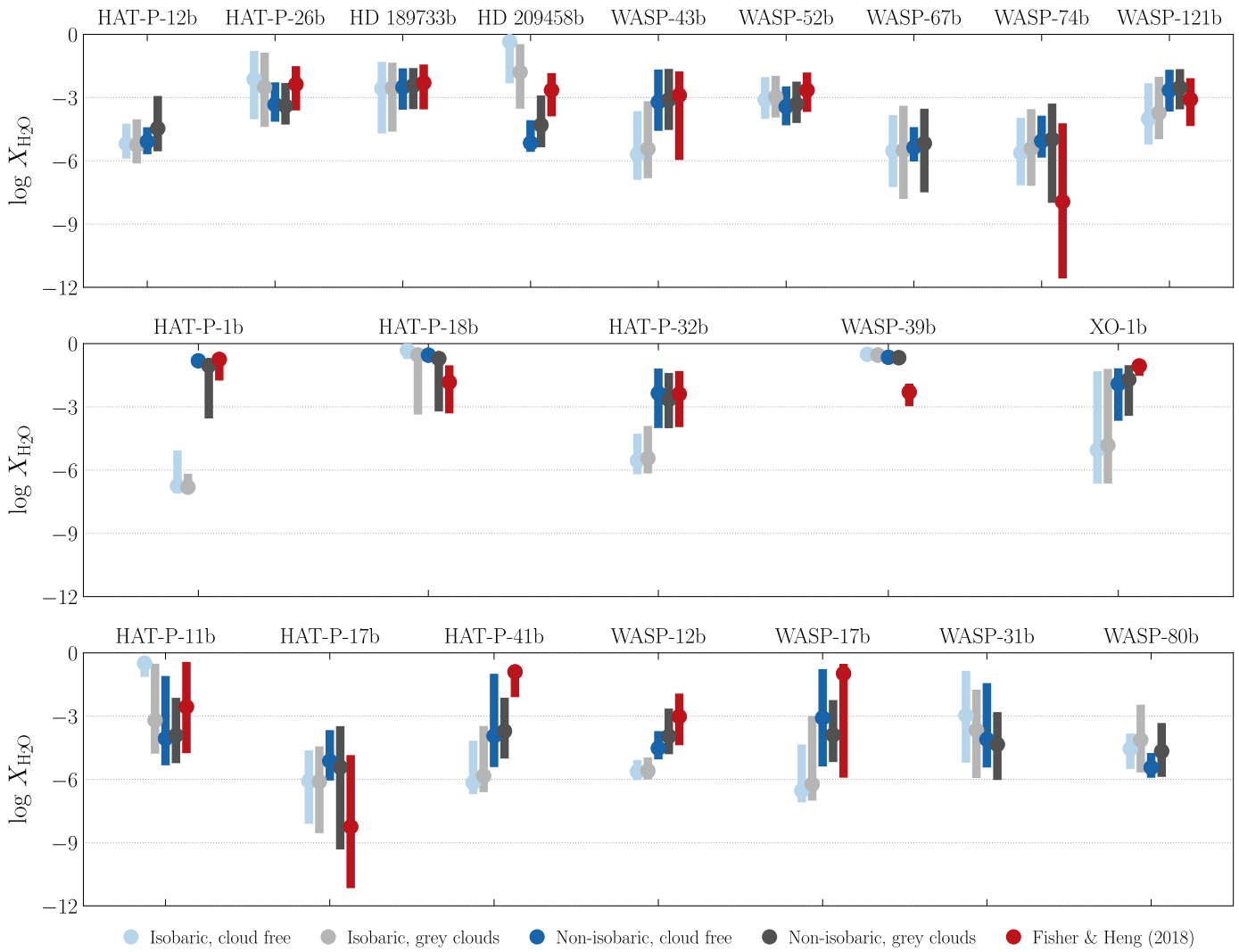

Figure 2. Comparison between retrieved water abundance values for all planets considering cloud-free (blue) and grey cloud (grey) models, using fully isobaric (dark-coloured) and non-isobaric (light-coloured) approaches. Red circles correspond to values found by Fisher & Heng (2018). Planets that are not in this plot are considered flat lines or outliers. Figure from Novais et al. (submitted to MNRAS).

3.3). Consequently, a reference parameter (P_0 or R_0) needs to be retrieved together with the main parameters (T, X_{H_2O}) in all multi-parameter models, as discussed in Fisher & Heng (2018). Overall, R_0 may be roughly approximated to the observed planetary radius.

Input values for R_{star} and g are from the literature, while R_0 values were calculated by Fisher & Heng (2018). Differing from Fisher & Heng (2018), this work considers R_{star}, g, and R_0 as free parameters in all non-flat-line models, allowing their priors to vary within their uncertainty range, and retrieving their posterior distributions.

3. Results and Discussion

3.1. Selecting best-fit models using Bayesian inference

By applying a range of isobaric and non-isobaric models, we designate which retrieval model best represents the atmosphere of each planet through a Bayesian inference analysis (Trotta 2008), which uses Bayes theorem to determine how well a model fits the given data by estimating its probability. The method is based on the comparison of the Bayesian evidence of each model, calculated in each retrieval by a nested-sampling algorithm (see Buchner et al. 2014 and Buchner 2021 for more details). For each planet, we select the model with the highest Bayesian evidence as the best-fit model. We then compute the Bayes factor of each model, which is the difference between the highest Bayesian evidence (amongst all models) and the Bayesian evidence of the given model, both in logarithm scale. One can easily conclude that the Bayes factor of the best-fit model for a planet is always null. For planets with two or more best models (i.e., two or more models with maximum Bayesian evidence), the simplest model is considered the best.

If the analysis of a model is inconclusive, i.e., its Bayes factor is less than unity, we cannot rule out the possibility of that model being the best fit, despite its Bayesian evidence not being the highest. This is especially taken into account for flat-line models: if the flat-line model of a planet has Bayes factor of less than unity, we assume the flat-line is the best fit for the atmosphere of that planet.

Bayesian model comparison leads us to three broad classes of outcomes for the 38 objects: possibly cloud-free model atmospheres with trace amounts of water; possibly cloud-free model atmospheres with enhanced mean molecular weight because of elevated amounts of water; and possibly cloudy model atmospheres with water and grey clouds. Note that for all three classes of outcomes, both the cloud free and grey cloud interpretations

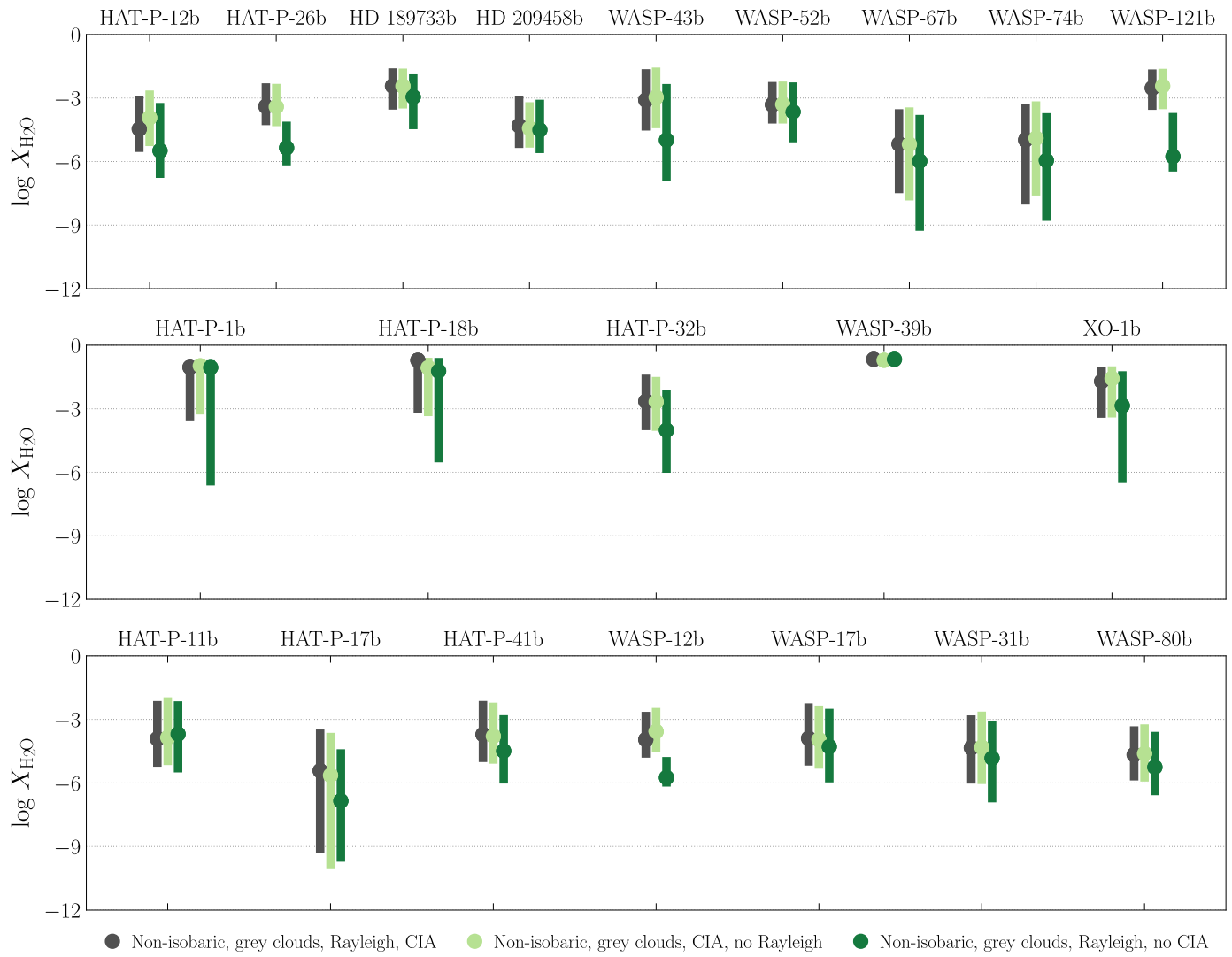

Figure 3. Comparison between retrieved water abundance values for all planets considering a non-isobaric, grey cloud model (grey), non-isobaric, grey cloud model without Rayleigh scattering (light green), and non-isobaric, grey cloud model without CIA (dark green). Planets that are not in this plot are considered flat lines or outliers. Figure from Novais et al. (submitted to MNRAS).

are valid (i.e., Bayesian model comparison produces Bayes factors below unity).

3.2. Trends

Figure 2 provides an executive summary of the retrieved water abundances for all of the objects in our sample. It compares the isobaric versus non-isobaric treatments and the inclusion or exclusion of grey clouds. For continuity to past work, the retrieved water abundances of Fisher & Heng (2018) are also included. We recall the main differences between the current work and Fisher & Heng (2018) are the inclusion of Rayleigh scattering, variable $R_{\rm star}$ and g, and non-isobaric transit chords, which reverberate in the correct implementation of cross sections of molecular species and CIA.

For most objects, the isobaric treatment yields water abundances that are consistent with the non-isobaric treatment, considering the wide uncertainty ranges. HD 209458b, HAT-P-1b, and HAT-P-32b are some exceptions, for which the inconsistency between isobaric and non-isobaric may be explained by high retrieved H_2O abundances. In most cases, the retrieved water abundances are broadly consistent with those of Fisher & Heng (2018).

Figure 3 does a different comparison, which is to examine the retrieved water abundances using the non-isobaric treatment with grey clouds. Rayleigh scattering and CIA are excluded in turn. The comparison highlights the general importance of including CIA in non-isobaric retrievals, which produces retrieved water abundances that are more precise. When CIA is removed, it no longer sets a continuum in the spectra, hence lowering the water abundance.

3.3. Parameter degeneracies

Due to the limited spectral range and resolution, our study also analyses degeneracies between retrieved parameters, which are illustrated using two representative case studies: HD 209458b and WASP-12b. Their data favour a cloud-free and a grey-cloud model, respectively, as represented in Figure 4.

The first thing to notice is, according to Benneke & Seager (2012); Griffith (2014); Fisher & Heng (2018), a "normalization degeneracy" (i.e., between the reference radius R_0 and the water abundance X_{H_2O}) is expected in retrievals using WFC3 data. Nonetheless, as the stellar radius $R_{\rm star}$ is a free parameter in our retrievals, the normalization degeneracy is subdued, and a degeneracy between $R_{\rm star}$ and R_0 appears instead.

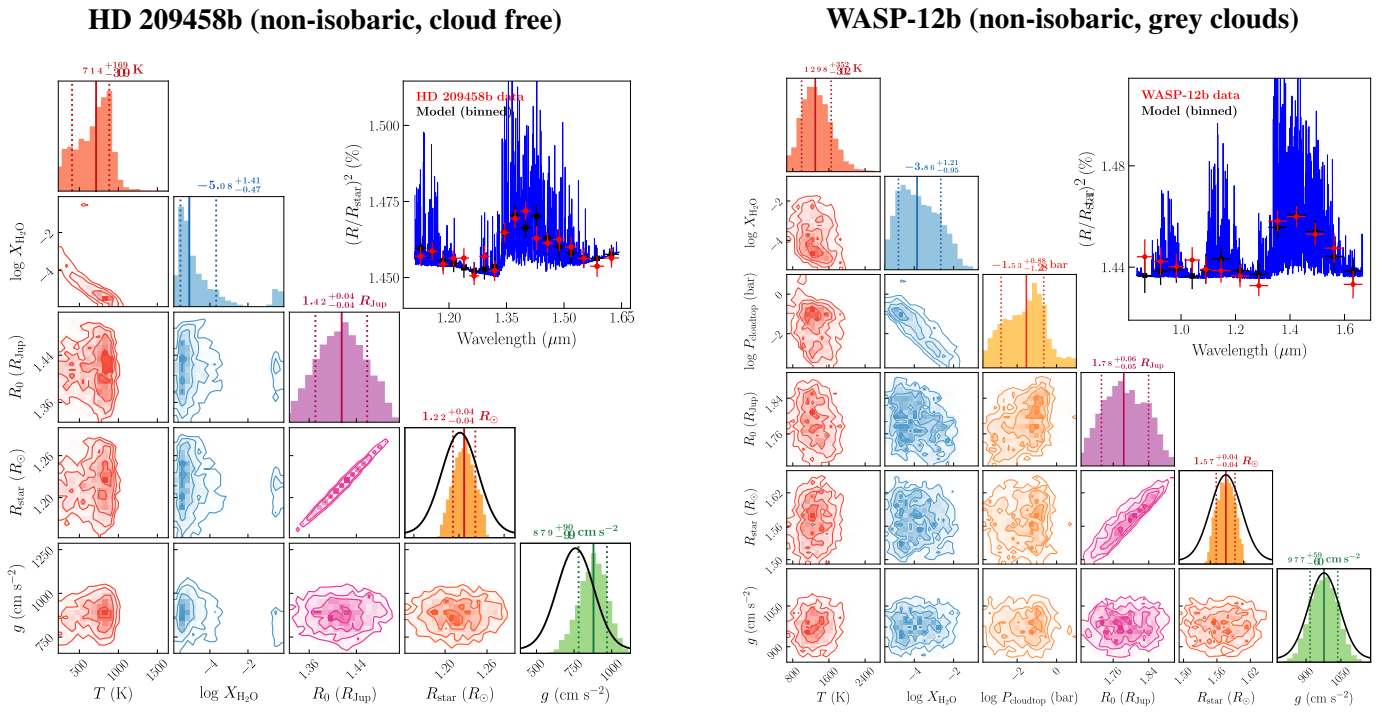

Figure 4. Retrieval outcomes for HD 209458b using a non-isobaric cloud-free model (left) and WASP-12b using a non-isobaric grey-cloud model (right). Both outcomes include H_2O, Rayleigh scattering, and CIA. Corner plots show histograms for posterior distribution of each parameter, where the coloured solid line is the mean value, and the dashed lines correspond to the 1σ uncertainties. These values are precisely represented on top of each histogram. Black solid lines in distributions represent Gaussian priors for R_{star} and g. On the top right corner of each figure, the best-fit spectrum is presented in blue in a resolution of 2 cm^{-1}. Red circles show WFC3 observed data, and black circles correspond to the best-fit model binned in the resolution of the data. Figure from Novais et al. (submitted to MNRAS).

When Rayleigh scattering by H_2 is added (an effect ignored by Fisher & Heng 2018), an absolute theoretical normalization is introduced via a spectral slope (Benneke & Seager 2012). This is generally the case for atmospheres where H_2 is the dominant background gas. However, if CIA is included, there is little difference in outcomes when Rayleigh scattering is or is not present. This is because CIA is providing a higher continuum, and Rayleigh scattering becomes negligible. Note that this is only true for WFC3 data coverage, since the Rayleigh slope extends to the optical range, meaning the effects of Rayleigh scattering are most seen in optical data.

When the non-isobaric best-fit model is cloud free (e.g., left panel of Figure 4), the spectral continuum is sourced by CIA. CIA is associated with the correct pressure, which has the effect of producing lower water abundances–regardless of the exclusion or inclusion of Rayleigh scattering.

When the non-isobaric best-fit model includes grey clouds (e.g., right panel of Figure 4), the spectral continuum is sourced by CIA and grey clouds, and a degeneracy between the cloud-top pressure $P_{cloudtop}$ and X_{H_2O} is present alongside the R_0–R_{star} degeneracy. In all retrievals, there is a degeneracy between the temperature T and X_{H_2O}, which only shows up strongly when the $P_{cloudtop}$–X_{H_2O} degeneracy is suppressed.

Acknowledgements. A.N. acknowledges financial support from Coordination of Improvement of Higher Education Personnel (CAPES) and partial financial support from Ludwig-Maximilians-Universität München (LMU-Munich). L.G. acknowledges financial support from Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), through the ARC research grant E-26/211.386/2019.

References

Benneke, B. and Seager, S., 2012, The Astrophysical Journal, vol. 753, number 2, p. 100
Buchner, J. et al., 2014, Astronomy & Astrophysics, vol. 564, id. A125, 25 pp.
Buchner, J., 2021, eprint 2101.09675
de Wit, J. & Seager, S., 2013, Science, vol. 342, number 6165, p. 1473-1477
Fisher, C. & Heng, K., 2018, Monthly Notices of the Royal Astronomical Society, vol. 481, issue 4, p. 4698-4727
Griffith, C., 2014, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 372, issue 2014, pp. 20130086-20130086
Heng, K. & Kitzmann, D., 2017, Monthly Notices of the Royal Astronomical Society, vol. 470, issue 3, p. 2972-2981
Madhusudhan, N., 2018, Handbook of Exoplanets, eid 104, p. 104
Seager, S., 2010, isbn 978-0-691-14645-4
Trotta, R., 2008, Contemporary Physics, vol. 49, issue 2, pp. 71-104