



YunMa: Enabling Spectral Retrievals of Exoplanetary Clouds

Sushuang Ma¹ , Yuichi Ito^{2,1} , Ahmed Faris Al-Refaie¹ , Quentin Changeat^{3,1} , Billy Edwards^{4,5,1} , and
Giovanna Tinetti¹

¹ Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT UK; sushuang.ma.20@ucl.ac.uk

² National Astronomical Observatory of Japan, 2 Chome-21-1, Osawa, Mitaka, Tokyo 181-8588, Japan

³ European Space Agency (ESA) ESA Office, Space Telescope Science Institute (STScI), 3700 San Martin Drive, Baltimore, MD 21218, USA

⁴ SRON, Netherlands Institute for Space Research, Niels Bohrweg 4, NL-2333 CA, Leiden, The Netherlands

⁵ AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris, F-91191 Gif-sur-Yvette, France

Received 2023 January 19; revised 2023 July 24; accepted 2023 September 1; published 2023 November 3

Abstract

In this paper, we present *YunMa*, an exoplanet cloud simulation and retrieval package, which enables the study of cloud microphysics and radiative properties in exoplanetary atmospheres. *YunMa* simulates the vertical distribution and sizes of cloud particles and their corresponding scattering signature in transit spectra. We validated *YunMa* against results from the literature. When coupled to the *TauREx 3* platform, an open Bayesian framework for spectral retrievals, *YunMa* enables the retrieval of the cloud properties and parameters from transit spectra of exoplanets. The sedimentation efficiency (f_{sed}), which controls the cloud microphysics, is set as a free parameter in retrievals. We assess the retrieval performances of *YunMa* through 28 instances of a K2-18 b-like atmosphere with different fractions of H₂/He and N₂, and assuming water clouds. Our results show a substantial improvement in retrieval performances when using *YunMa* instead of a simple opaque cloud model and highlight the need to include cloud radiative transfer and microphysics to interpret the next-generation data for exoplanet atmospheres. This work also inspires instrumental development for future flagships by demonstrating retrieval performances with different data quality.

Unified Astronomy Thesaurus concepts: Exoplanets (498); Atmospheric clouds (2180); Exoplanet atmospheres (487); Transmission spectroscopy (2133)

1. Introduction

Thousands of exoplanets have been detected since the late 20th century. During the past decade, transit spectroscopy has become one of the most powerful techniques for studying exoplanets' atmospheres in-depth (e.g., reviews by Tinetti et al. 2013; Burrows 2014; Madhusudhan 2019). Data recorded from space-borne instruments (e.g., Hubble, Spitzer, and James Webb Space Telescopes) or from the ground have revealed important information about exoplanet atmospheric chemistry and dynamics (e.g., Sing et al. 2016; Tsiaras et al. 2018; Welbanks et al. 2019; Venot et al. 2020; Roudier et al. 2021; Changeat et al. 2022a; Edwards et al. 2022; JWST Transiting Exoplanet Community Early Release Science Team 2023) and may provide insight into planetary interior composition and formation (Madhusudhan et al. 2020; Tsai et al. 2021; Yu et al. 2021; Charnay et al. 2022).

A number of spectral retrieval models have been developed by different teams to interpret the atmospheric data and quantify their information content; these include, e.g., Madhusudhan & Seager (2009), Lee et al. (2012), TauREx 3 (Al-Refaie et al. 2021), NEMESIS (Irwin et al. 2008), CHIMERA (Line et al. 2013), ARtful modeling Code for exoplanet Science (ARCiS; Ormel & Min 2019; Min et al. 2020), PICASO (Batalha et al. 2019; Robbins-Blanch et al. 2022), BART (Harrington et al. 2022), petitRADTRANS (Mollière et al. 2020), HELIOS (Kitzmann et al. 2020), POSEIDON (MacDonald & Madhusudhan 2017), HyDRA (Gandhi &

Madhusudhan 2018), SCARLET (Benneke 2015), PLATON II (Zhang et al. 2020), and Pyrat-Bay (Cubillos & Bleicic 2021). Up to date, most of the retrieval studies of exoplanetary atmospheres are highly parameterized. This approach has been very sensible given the relatively poor information content of current atmospheric data. However, a number of papers in the literature (e.g., Caldas et al. 2019; Changeat et al. 2021, 2022a) have cautioned against this approach when applied to data recorded with next-generation facilities.

Clouds are omnipresent in planetary, exoplanetary, and brown dwarf atmospheres (see, e.g., review by Helling 2023a) and have often been detected in exoplanet atmospheric data (Kreidberg et al. 2014; Sing et al. 2016; Stevenson 2016; Tsiaras et al. 2018). Their presence imposes additional complexity and uncertainties in the interpretation of exoplanet atmospheric spectra (e.g., Mai & Line 2019; Tsiaras et al. 2019; Changeat et al. 2022b).

Models simulating the formation and radiative properties of clouds and hazes have been published in the literature, e.g., Exoplanet Radiative-convective Equilibrium Model (Baudino et al. 2015; Charnay et al. 2018), Gao et al. (2020), Windsor et al. (2023), and Kawashima & Ikoma (2018).

Due to the—currently limited—observational constraints and computational resources available to simulate the complexity of clouds, retrieval studies of cloudy atmospheres are still in their infancy (see, e.g., Fortney et al. 2021). For instance, many studies have adopted wavelength-independent opaque clouds, where all the radiation beneath the cloud top is blocked from reaching the telescope, and retrieves the vertical location of clouds (Brogi & Line 2019; Boucher et al. 2021). Wakeford et al. (2018) used a gray, uniform cloud in the ATMO Retrieval Code (Tremblin et al. 2015; Drummond et al. 2016;

Goyal et al. 2018). Other models constrain from radiative transfer the uniform cloud particle sizes without being estimated through cloud microphysics models. For instance, Benneke et al. (2019) have initially estimated the particle sizes in the atmosphere of GJ 3470 b using Mie-scattering theory. Extended from this highly parametric approach, cloud scattering parameters and inhomogeneous coverage were also retrieved: NEMESIS was used by Barstow (2020) and Wang et al. (2022) to retrieve the cloud's opacity, scattering index, top and base pressures, particle sizes, and shape factor. Pinhas et al. (2019) ran POSEIDON to constrain the cloud's top pressure and coverage fraction. Wang et al. (2022) adopted PICASO to extract the cloud's base pressure, optical thickness, single scattering albedo, scattering asymmetry, and coverage. Lueber et al. (2022) extended the use of Helios-r2 to retrieve nongray clouds, with extinction efficiencies estimated from Mie theory calculations. The model Aurora (Welbanks & Madhusudhan 2021) presents inhomogeneities in cloud and haze distributions by separating the atmosphere horizontally into four distinct areas.

The data provided by the next-generation telescopes will be greatly superior in quality and quantity, allowing us to obtain more stringent constraints to our understanding of clouds in exoplanetary atmospheres. Transit spectra of exoplanets recorded from space by the James Webb Space Telescope (JWST; 0.6–28.3 μm , Gardner et al. 2006; Greene et al. 2016; Bean et al. 2018), Ariel (0.5–7.8 μm , Tinetti et al. 2018, 2021), and Tinkle (0.5–4.5 μm , Edwards et al. 2019) at relatively high spectral resolution and/or broad wavelength coverage will open the possibility of integrating self-consistent, cloud microphysics approaches into atmospheric retrieval codes. A good example of such models is ARCiS (Ormel & Min 2019; Min et al. 2020), which simulates cloud formation from diffusion processes and parametric coagulation. ARCiS also generates cloudy transit spectra from Mie theory (Fleck & Canfield 1984) and distribution of hollow spheres (Min et al. 2005; Mollière et al. 2019), and can be used to retrieve the cloud diffusivity and nuclei injection from transit spectra.

In this work, we present a new optimized model to study cloud microphysical processes directly integrated into a spectral retrieval framework. We consider clouds as a thermochemical product, i.e., the aggregation of condensates in the atmosphere, while hazes form photochemically (Kawashima & Ikoma 2018). The cloud distribution depends on the atmospheric conditions. Being generated thermochemically, clouds form and diffuse depending on the atmospheric thermal structure and, in return, contribute to it. They also depend on the mixing profiles of the condensable gases in the atmosphere. Clouds act as absorbers and/or scatterers and therefore may dampen the atomic and molecular spectroscopic features and change the continuum.

Based on studies of the Earth and solar system's planetary atmospheres, Lewis (1969) published a 1D cloud model optimized to describe tropospheric clouds in giant planets. This model assumes that the fall speeds of all condensates are equivalent to the updraft velocities, and only vapor is transported upward. Lunine et al. (1989) included a correlation between cloud particle sizes, downward sedimentation, and upward turbulent mixing. Based on previous models by Lewis (1969), Carlson et al. (1988), Lunine et al. (1989), and Marley et al. (1999); Ackerman & Marley (2001) proposed a new method to estimate the mixing ratio and vertical size

distribution of cloud particles (A-M model hereafter). In the A-M model, the sedimentation timescale is estimated through cloud microphysics, taking into account the atmospheric gas kinetics and dynamical viscosity. The model assumes an equilibrium between upward turbulent mixing and sedimentation, where the turbulent mixing is derived from the eddy diffusion in the atmosphere. The key assumptions of the A-M model are as follows:

1. Clouds are distributed uniformly in the horizontal direction.
2. Condensable particles rain out at (super)saturation while maintaining a balance of the upward and downward drafts.
3. It does not consider the cloud cover variations caused by precipitation or the microphysics between different types of clouds.

The A-M model was originally proposed for giant exoplanets and brown dwarfs and was tested on Jupiter's ammonia clouds, demonstrating that this approach is applicable to a broad range of temperatures and planetary types.

Another popular 1D cloud microphysics model is the Community Aerosol and Radiation Model for Atmospheres (CARMA), initially developed for the Earth's stratospheric sulfate aerosols (Toon et al. 1979; Turco et al. 1979). CARMA is a time-dependent cloud microphysics model, which solves the discretized continuity equations for aerosol particles starting from nucleation. Gao et al. (2018) extended the use of CARMA to simulate clouds on giant exoplanets and brown dwarfs by including additional condensates predicted to form in hot atmospheres and compared the results with the A-M model. The A-M model, while able to provide the cloud particle sizes and number density distributions, is of intermediate numerical complexity and, therefore, potentially adaptable to be included in retrieval codes. In addition to the original implementation by Ackerman & Marley (2001), Virga (Rooney et al. 2022) simulates the cloud's particle size distribution from the A-M approach and estimates separately the sedimentation efficiency. PICASO (Batalha et al. 2019; Robbins-Blanch et al. 2022) adopts Virga to simulate cloudy exoplanetary atmospheres. Adams et al. (2022) couples MIT global circulation model (GCM) and Virga to include clouds in 3D models. The above are forward simulations only. One step further, Xuan et al. (2022) present retrieval studies on HD 4747 B with clouds using petitRADTRANS (Mollière et al. 2020), where the cloud simulation in retrieval is motivated by the A-M approach. To estimate the cloud mixing ratio in the retrieval iterations, this model does not solve the full ordinary differential equation (ODE; see Equation (2)). Instead, it adopts an approximation of the A-M approach, assuming the mixing ratio of condensable gas above the cloud base is negligible.

To simulate inhomogeneities for cloud formation in the horizontal direction, we would need to consider global circulation atmospheric effects, such as those modeled in Cho et al. (2021). An example of a 3D atmospheric model with clouds is Aura-3D (Welbanks & Madhusudhan 2021; Nixon & Madhusudhan 2022). The retrieval part for Aura-3D is highly parameterized, both for the atmospheric and cloud parameters. Helling et al. (2019, 2023b) have simulated global cloud distributions by generating inputs to their kinetic cloud model from precalculated 3D GCMs. Unfortunately, these complex models require excessive computing time. In addition, the data

expected to be observed in the near future are unlikely to constrain the large number of parameters needed in a 3D model. Therefore, while theoretical studies with 3D models are very important to progress in our understanding of clouds in exoplanetary atmospheres and as benchmarks, they are currently less useful for interpreting available data.

In this paper, we present a new cloud retrieval model, *YunMa*, optimized for transit spectroscopy. In *YunMa*, we built the cloud model based on Ackerman & Marley (2001) and simulated the cloud contribution to transit spectra using extinction coefficients as calculated by the open-source Bohren & Huffman (2008; hereafter BH-Mie) code. *YunMa* is fully integrated into the *TauREx 3* retrieval platform (Al-Refaie et al. 2021, 2022a) and, for the first time, provides cloud microphysics capabilities into a retrieval model. We describe the model in Section 2. In Section 3, we detail the experimental setups. In Section 4, we validate particle size distributions and spectroscopic simulations against previous studies published in the literature. After validation, we show new spectral and retrieval simulations obtained with *YunMa*. In Section 5, we discuss our results and assumptions and identify possible model improvements to be considered in future developments.

2. *YunMa* Description

YunMa estimates the vertical distribution of the cloud particle sizes (VDCP hereafter) based on A-M model and their contribution to the radiative transfer calculations. The *YunMa* module has been integrated into the *TauREx 3* retrieval platform: the combined *YunMa-TauREx* model is able to constrain the VDCP from observed atmospheric spectra, as described in detail below.

2.1. Modeling the Cloud Particle Size Distribution

YunMa model contains a numerical realisation of the A-M microphysical approach to simulate the VDCP. We show in Figure 1 a pictorial representation of the A-M approach: it assumes that clouds form with different VDCP to maintain the balance between the upward turbulent mixing and downward sedimentation of the condensable species. Depending on the atmospheric $T-p$ profile, multiple cloud layers may form.

2.1.1. Cloud Mixing Profile

Cloud particles start forming when the partial pressure of a certain gas exceeds the saturation vapor pressure (SVP): the formation strongly depends on the atmospheric thermal structure. The condensation process, occurring when the partial pressure exceeds the SVP, is estimated by comparing the molecular mixing ratio of the gas phase with its saturation vapor mixing ratio:

$$q_c(z) = \max[0, q_t(z) - (S_c + 1) q_s(z)], \quad (1)$$

where q_c is the mixing ratio of the condensed species, z is the altitude, q_s is the mixing ratio where the condensable gas saturates, q_t is the total mixing ratio of a condensable chemical species, including both the condensate and gas phases, and S_c is the supersaturation factor, which persists after condensation. q_s can be estimated from the ratio between the SVP of a certain chemical species and the atmospheric pressure at the same altitude. Note that, in this paper, the mixing ratio refers to the volume fraction of a chemical species in the atmosphere.

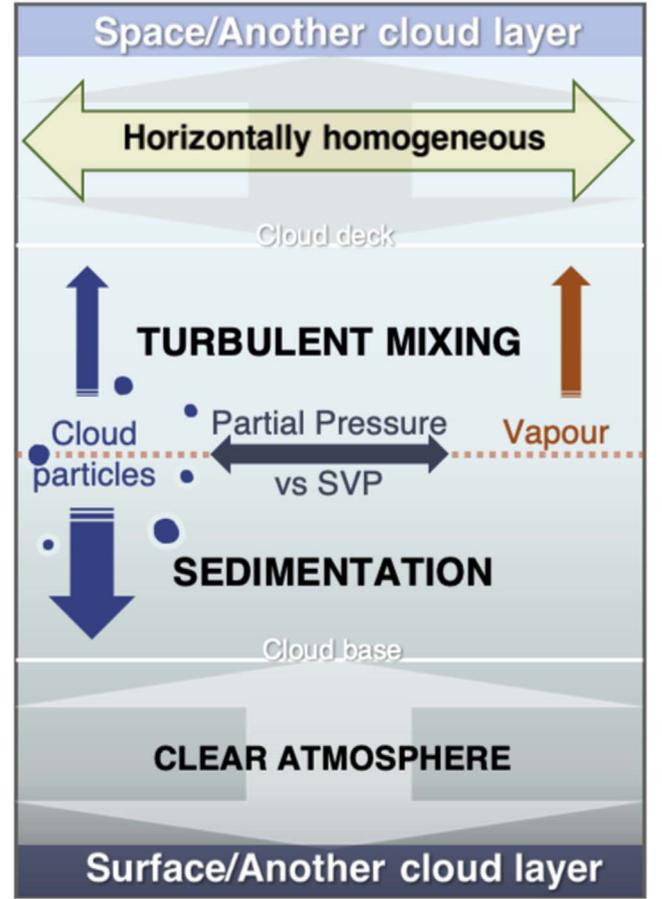


Figure 1. Sketch of the A-M microphysical approach adopted in *YunMa*. Cloud particles may form when the mixing ratio of the condensable gas exceeds the saturation mixing ratio, which is derived from its saturation vapor pressure (SVP). The vertical distribution of the cloud particle sizes (VDCP) is derived from the balance between the sedimentation of the cloud particles and the atmospheric turbulent mixing.

In the A-M approach, the turbulent mixing of the condensate and vapor is assumed to be in equilibrium with the sedimentation of the condensate:

$$-K(z) \frac{\partial q_t(z)}{\partial z} - f_{\text{sed}} w_*(z) q_c(z) = 0, \quad (2)$$

where K ($\text{m}^2 \text{s}^{-1}$) represents the vertical eddy diffusion coefficient, and w_* (m s^{-1}) is the convective velocity. f_{sed} is the ratio between the mass-weighted droplet sedimentation velocity and w_* , defined as follows:

$$f_{\text{sed}} = \frac{\int_0^\infty v_f \frac{dm}{dr} dr}{\varepsilon \rho_a w_* q_c}; \quad (3)$$

here, ρ_a is the atmospheric mass density, which can be estimated through the ideal gas law, ρ_p is the mass density of a condensed particle, ε is the ratio between the molecular weights of the condensates and the atmosphere, and v_f is the sedimentation velocity, which will be explained later. The first term in Equation (2) describes the upward vertical draft derived from the macroscopic eddy diffusion equation. The second term describes the downward sedimentation, which is in equilibrium with the first term.

The eddy diffusion coefficient (K) is one of the key parameters affecting cloud formation. In free convection (Gierasch & Conrath 1985), it can be estimated as follows:

$$K = \frac{H}{3} \left(\frac{L}{H} \right)^{\frac{4}{3}} \left(\frac{RF}{\mu \rho_a c_p} \right)^{\frac{1}{3}}, \quad (4)$$

where H , μ , and c_p are, respectively, the atmospheric scale height, mean molecular weight, and specific heat capacity. $F = \sigma T_{\text{eff}}^4$ is the approximated radiative flux. The turbulent mixing length (L) is the scale height of the local stability in eddy diffusion, as opposed to the atmospheric scale height (H). *YunMa* has an application programming interface for K , and it can use the values provided by disequilibrium chemistry models plugged by the users into *TauREx 3*—e.g., the kinetic model plugin of *TauREx 3* (Al-Refaie et al. 2022b), and derive L accordingly, from Equation (4). The convective velocity scale (w_*) mentioned above can also be estimated as a ratio between K and L . There are different ways to constrain K from the atmospheric chemical and vertical advective timescales (e.g., Parmentier et al. 2013; Zhang & Showman 2018; Komacek et al. 2019; Baeyens et al. 2021). However, the current estimations of K in the literature lack validation from observations. While *YunMa* is designed to be self-consistent with these approaches, this paper uses constant K in the experiments as a first-order estimation. Free convection is justified by assuming the cloud forms in the deep convective layer of the atmosphere and by neglecting 3D effects. These approximations in retrieval studies will need to be revisited with the improved quality of the data available and computing facilities.

The sedimentation velocity, denoted by v_f is the speed at which a cloud particle settles within a heterogeneous mixture due to the force of gravity. v_f can be estimated through viscous fluid physics:

$$v_f = \frac{2}{9} \frac{\beta g r^2 \Delta \rho}{\eta}, \quad (5)$$

where $\Delta \rho$ is the difference between ρ_p and ρ_a , β is the Cunningham slip factor, and η is the atmospheric dynamical viscosity (see Appendix A for more details of SVP, β , and η).

2.1.2. Particle Size and Number Density

Following the A-M approach, we assume spherical cloud particles with radii r . The particle radius at w_* , denoted as r_w , can be obtained using these relationships between v_f and w_* :

$$v_f(r_w) = w_*, \quad (6)$$

and

$$v_f = w_* \left(\frac{r}{r_w} \right)^\alpha, \quad (7)$$

where α corresponds to the sedimentation velocity decrease in viscous flows. In A-M, the particle size distribution was constrained by in situ measurements of Californian stratocumulus clouds, which followed a broad lognormal distribution. The assumption of lognormal distribution allows estimating the geometric mean radius (r_g), the effective radius (r_{eff}), and the

total cloud particle number density (N), using the detailed definitions and derivations listed in Appendix B.

2.2. Cloud Contribution in Transit Spectra

To estimate the wavelength-dependent cloud contribution to transit spectra, *YunMa* adopts the scattering theory and absorption cross sections as described in BH-Mie, assuming spherical cloud particles. The cross section of the cloud particles (k_λ) at each wavelength (λ) and particle size are estimated through the extinction coefficient (Q_{ext}) of the corresponding wavelength and particle size, derived by BH-Mie from the refractive indices of the cloud particles:

$$k_\lambda = Q_{\text{ext}} \pi r^2. \quad (8)$$

We used the water ice refractive indices reported in Warren & Brandt (2008) for our simulations of the temperate super-Earth. We show some examples of atmospheres with water ice clouds in Section 4. Post-experimental tests were conducted to avoid the contamination of liquid water particles.

We simulate the cloud optical depths from the particle sizes and number densities along each optical path, which passes the terminator at altitude z_{ter} , with a path length $s_{z_{\text{ter}}}$ of each atmospheric layer:

$$\tau_\lambda = \int_{z_{\text{ter}}}^{z_{\text{top}}} \int_0^\infty k_\lambda \frac{dn}{dr} dr \frac{ds_{z_{\text{ter}}}}{dz} dz, \quad (9)$$

where n is the accumulated number density of particles with a radius smaller than r , and z_{top} is the altitude at the top of the atmosphere. The contribution of the clouds to the transit spectra, ΔF_c , can be estimated as follows:

$$\Delta F_c = \frac{2 \int_{z_{\text{bottom}}}^{z_{\text{top}}} (R_p + z)(1 - e^{-\tau_\lambda}) dz}{R_s^2}, \quad (10)$$

where z_{bottom} is the altitude at R_p . While *YunMa* has the capability to include any customised cloud particle size distribution in the spectral simulations, in this paper, we aim at model testing, and for simplicity, we use a single radius bin, i.e., uniform cloud particle size $r_c = r_g$ (see Equation (B1) in Appendix B) for each atmospheric layer in the radiative-transfer simulation.

2.3. Cloud Simulation Validation

We validated our implementation of A-M model against Ackerman & Marley (2001) by comparing the condensate mixing ratios of Jovian ammonia clouds (q_c) with different values of f_{sed} , as shown in Figure 2. The two sets of results are consistent, and the small differences in q_c translate into $\sim 10^{-2}$ ppm in transit depth, which is completely negligible compared to typical observational noise.

We further validated our implementation against the results from Gao et al. (2018) by comparing the KCl cloud molecular mixing ratio and particle size profile (Figure 3). Gao et al. (2018) used CARMA to simulate cloud microphysics in exoplanets and brown dwarfs with $T_{\text{eff}} = 400$ K, and $\log g = 3.25, 4.25, \text{ and } 5.25$ (in cgs units), corresponding to planetary masses of $0.72 M_J, 8.47 M_J, \text{ and } 44.54 M_J$. In the best fit between CARMA to A-M model, the $f_{\text{sed}} = 0.125, 0.093, \text{ and } 0.025$, for the cases with $K = 10^4, 10^3, \text{ and } 10^2 \text{ m}^2 \text{ s}^{-1}$, respectively. The mixing length is derived from constant eddy

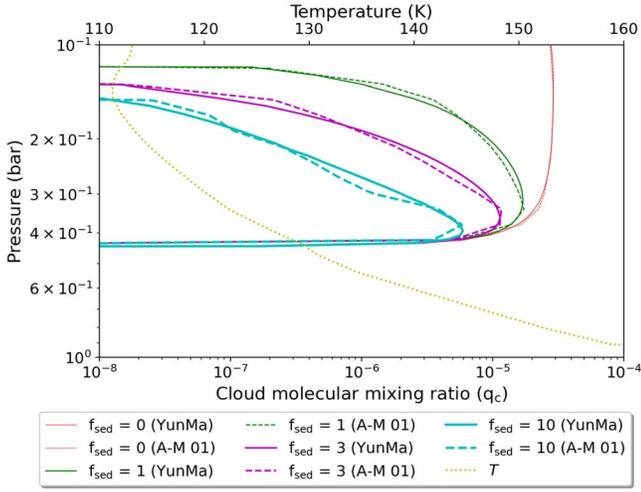


Figure 2. Validation of the *YunMa* cloud microphysics model against the Jovian ammonia clouds in Ackerman & Marley (2001). q_c corresponding to different sedimentation efficiencies (f_{sed}) are shown. Solid lines: results from *YunMa*. Dashed lines: results from Figure 1 in Ackerman & Marley (2001). Dotted line: T - p profile.

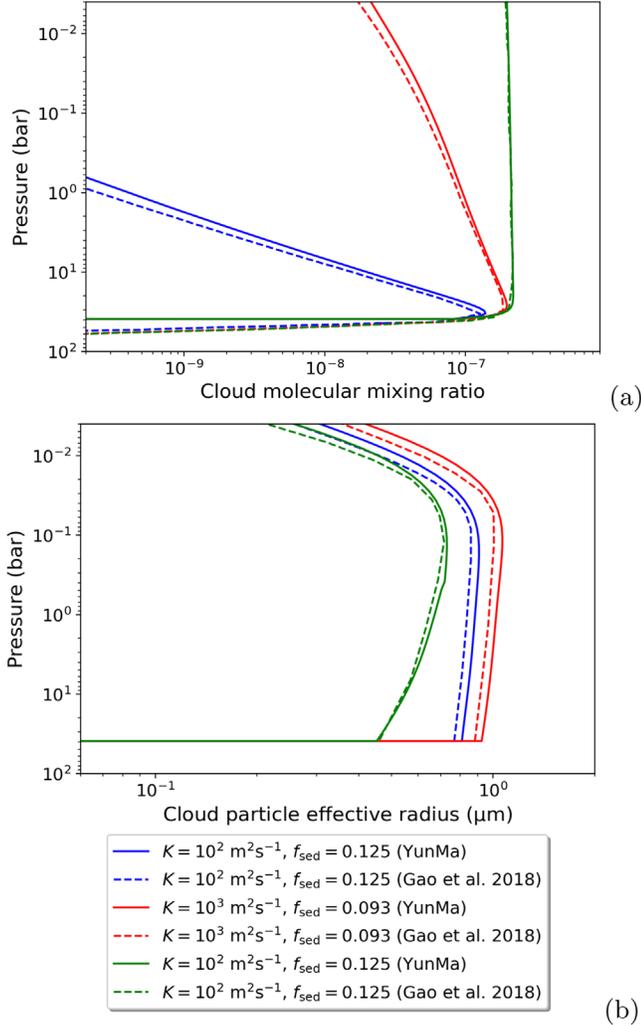


Figure 3. Validation of *YunMa* (solid lines) against the A-M model in Gao et al. (2018; dashed lines). Top: condensate mixing ratios. Bottom: cloud particle effective mean radii.

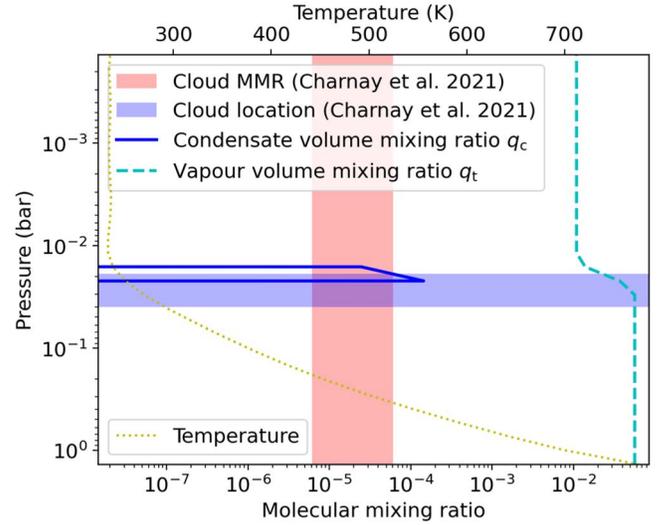


Figure 4. Comparison of *YunMa*'s results against simulations by Charnay et al. (2021b) for the $100\times$ solar metallicity scenario at the substellar point (see Figure 6(a) in the original paper). The condensate (blue) and vapor (cyan) MMRs are simulated by *YunMa* using the A-M approach with $K = 10^2 \text{ m}^2 \text{ s}^{-1}$. The cross section of the two shaded areas indicates the range of cloud MMR and location simulated by Laboratoire de Météorologie Dynamique Generic GCM (LMDG) in Charnay et al. (2021b). LMDG is derived from the LMDZ Earth (Hourdin et al. 2006) and Mars (Forget et al. 1999) GCMs.

diffusion, as described in Equation (4). The results agree with each other within 8%, i.e., $\sim 10^{-3}$ ppm difference in the transit depth.

We compared simulations from *YunMa* with the results from Charnay et al. (2021b), which include the horizontal effects generated by global circulation. While very precise comparison and validation are not possible in this case due to the two approaches' very different natures and assumptions, it is useful to test whether we can reproduce similar results when we use consistent assumptions. In the comparison shown in Figure 4, we used the value of $K = 10^2 \text{ m}^2 \text{ s}^{-1}$ estimated in Charnay et al. (2021b) assuming $100\times$ solar composition. When setting f_{sed} to 3, *YunMa* produced similar results to those reported by Charnay et al. (2021b), with clouds forming in the region between 3×10^{-2} and 1×10^{-2} bar and a cloud molecular mixing ratio of approximately 10^{-4} .

We have validated the BH-Mie module in *YunMa* against PyMieScatt, an open-source model simulating atmospheric particle scattering properties (Sumlin et al. 2018), as shown in Figure 5. Cloud particle radii were selected in the range of 0.1 – $100 \mu\text{m}$. The largest discrepancy in Q_{ext} is within ± 0.002 , which corresponds to an average of 0.01 ppm in the planetary transit depth of the nominal scenario in our experiments.

2.4. *YunMa*–*TauREx*: Retrieval of Cloudy Atmospheres

We integrate the *YunMa* VDCP and τ_λ simulations in the Tau Retrieval of Exoplanets framework (*TauREx 3*; Al-Rafaie et al. 2021, 2022a), which allows atmospheric retrieval simulations. *TauREx 3* combined with *YunMa* allow us to perform retrievals, which include cloud microphysical processes and cloud scattering properties. The parameters estimated by *TauREx 3* include atmospheric T - p and chemical profiles, planetary parameters (e.g., mass and radius), and stellar parameters (e.g., temperature and metallicity). The radiative transfer calculations executed by *TauREx 3* consider

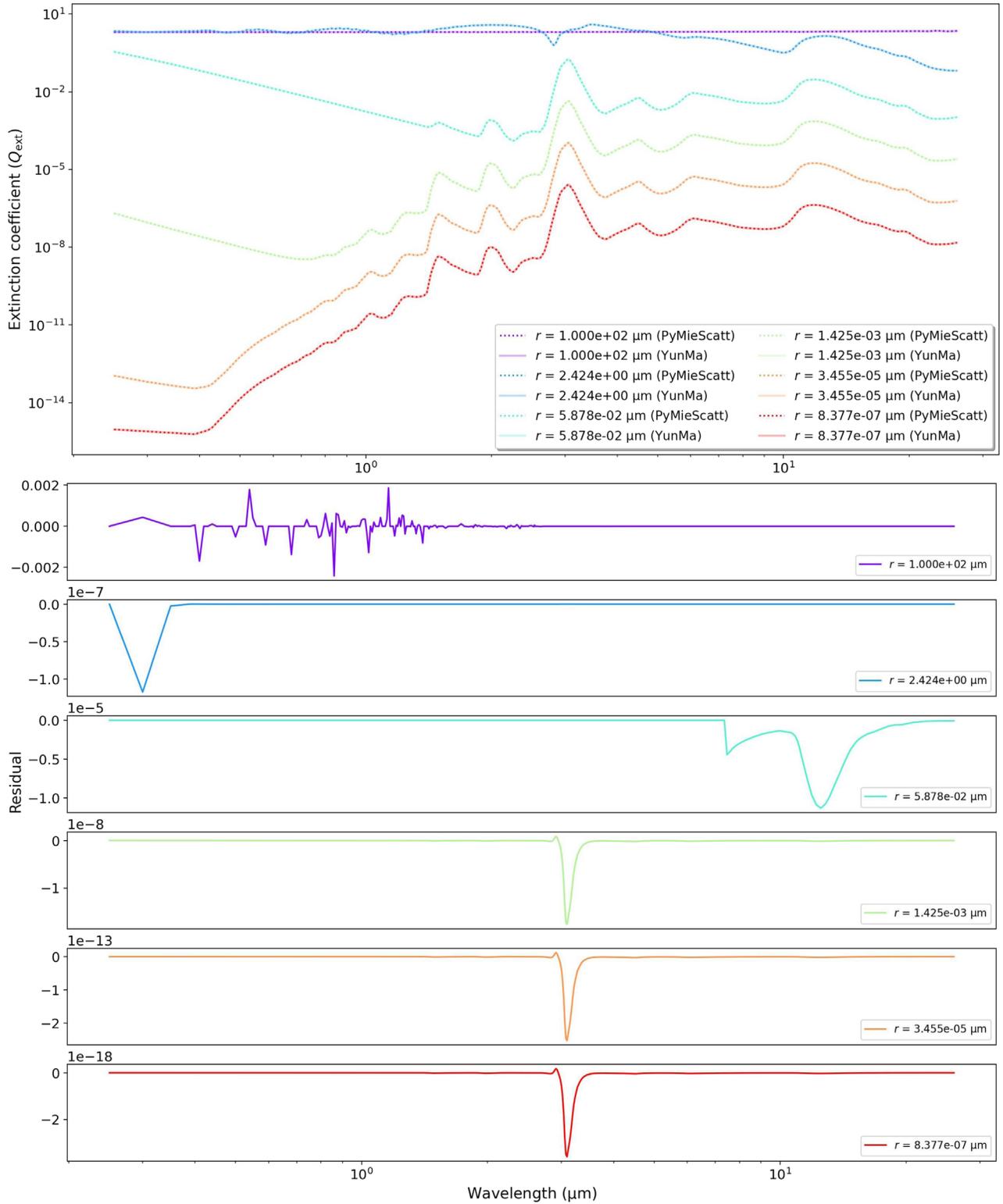


Figure 5. Validation of radiative-transfer simulations obtained with *YunMa* against the open-source code *PyMieScatt* (Sumlin et al. 2018). The extinction coefficients for cloud particles with different sizes are estimated from the theory of Bohren & Huffman (2008). To address the computational limitations of retrievals, we precalculated the extinction coefficients used in Equation (8) to estimate the cross sections of the cloud particles. The precalculated list includes values for particle radii from 1×10^{-7} to $1 \times 10^{-2} \mu\text{m}$, equally spaced in the logarithm space. Here, we show only six examples in the top panel. Bottom panels: residuals obtained by subtracting the extinction coefficients as estimated by the two codes, *YunMa* and *PyMieScatt*. The maximum discrepancy, corresponding to the largest particle radius simulated here ($r = 1 \times 10^2$), is negligible, i.e., ± 0.002 .

molecular and atomic absorptions, Rayleigh scattering and collisionally induced absorptions (CIA) of $\text{H}_2\text{--H}_2$ and $\text{H}_2\text{--He}$ pairs from Cox (2015).

YunMa uses as the initial condition the gas mixing ratio profiles provided by *TauREx 3* chemistry models (q_t , Equations (1), (2)). In this paper, for simplicity, we assume

Table 1Priors for Spectral Retrieval Experiments Using *YunMa* of All the Cases Listed in Table 2

Parameter	Unit	Ground	Mode	Priors
R_p	\mathcal{R}_e^N	0.20	factor	0.75–1.25
f_{sed}	...	Table 2	log	10^{-3} – 10^2
$X_{\text{H}_2\text{O}}$...	Table 2	log	10^{-12} –1
p_c	bar	Table 2	log	10^{-4} –1
T_c	K	200	linear	0–500
T_{surf}	K	1000	linear	500–2000
X_{N_2}	...	Table 2	linear	10^{-12} –1

Note. $X_{\text{H}_2\text{O}}$ represents the water vapor mixing ratio.

the baseline chemical abundances are constant with altitude instead of a more complex chemical structure. *YunMa* then adjusts the gas phase mixing ratios, atmospheric mean molecular weight, and atmospheric density in the *TauREx 3* chemistry models as a result of the formation of clouds. To simulate transit spectra and perform retrievals, we use the atmospheric grids and optical paths defined in *TauREx 3* and add the cloud opacities as estimated by *YunMa* BH-Mie to the absorptions caused by the chemical species, using the methods explained in Section 2.2. The retrievals were tested on 80 Intel (R) Xeon(R) Gold 6248 CPU @ 2.50 GHz.

3. Methodology

In this paper, we use *YunMa* to perform retrieval simulations of small temperate planets, where we expect a considerable amount of H_2O to be present in the atmosphere. For simplicity, we consider only water clouds forming in the atmosphere, and we do not consider supersaturation cases. The planetary parameters are inspired by K2-18 b (Tsiaras et al. 2019; Charnay et al. 2021b; Yu et al. 2021), which is a suitable candidate for cloud model testing. We list all the priors of our experiment in Table 1. In this work, we estimate η using the approximation proposed by Rosner (2012; Equation (A4)). We include both scattering and absorption due to water clouds based on BH-Mie, Rayleigh scattering of all the gas species, and CIA of H_2 – H_2 and H_2 – He pairs, which are enabled by *TauREx 3*. We use N_2 as a representative-inactive gas undetectable spectroscopically but that contributes to the increase of the atmospheric mean molecular weight, m , and decrease of scale heights $H = k_B T / mg$. H_2 and He act as the filling gases. We use the POKAZATEL data set for $^1\text{H}_2^{16}\text{O}$ (Polyansky et al. 2018) from the ExoMol database⁶ (Tennyson & Yurchenko 2012; Chubb et al. 2021; Tennyson & Yurchenko 2021) to estimate the water vapor absorption and Rayleigh scattering. The CIA data is from HITRAN⁷ (Karman et al. 2019). We use the PHOENIX library (Husser et al. 2013) to simulate the stellar atmospheres spectra.

For the numerical parameter settings, after a number of tests, we decided to use the explicit Runge–Kutta method of order 8 (DOP853, Hairer et al. 1993) with relative tolerance (*rtol*) of 1×10^{-13} and absolute tolerance (*atol*) of 1×10^{-16} to solve the partial differential of Equation (2) for all the experiments presented in Section 4. We have opted for a logarithm sampling

to retrieve most of the atmospheric parameters, e.g., f_{sed} , $X_{\text{H}_2\text{O}}$, and p_c . We have used a linear sampling, instead, for N_2 to obtain a better numerical performance. The priors are sufficiently unconstrained to avoid biases generated by excessive preknowledge, as discussed, e.g., in Changeat et al. (2021). After a number of tests, we have chosen to use 400 live points for 3D retrievals and 1000 for more dimensions.

To begin with, we run a sensitivity study with *YunMa* about the planetary and instrumental parameters. We set the planetary radius (R_p), f_{sed} , and $X_{\text{H}_2\text{O}}$ as free parameters in our 3D retrieval tests. We list in Table 1 the planetary parameters adopted in the simulations, the prior ranges, and the sampling modes. The simulations are conducted with 80 atmospheric layers, from 10 bar to 10^{-6} bar, which encompass the typical observable atmospheric range for super-Earths. We select case (2) in Table 2 as the nominal case, and test the model sensitivity to the key parameters in the retrievals. In case (2), the clouds dampen the gas spectroscopic features but do not obscure them entirely (see Figure 6). The nominal f_{sed} refers to the value adopted in A-M. The water SVP (both liquid and ice) used are taken from Appendix A in Ackerman & Marley (2001). In this experiment, we will perform sets of retrievals with the aim of the following:

1. sensitivity studies to key atmospheric parameters (cases (1)–(9));
2. sensitivity studies to data quality (cases (2), (10)–(13));
3. retrievals of atmospheric thermal profiles (cases (14)–(17));
4. addition of N_2 in retrievals (cases (18)–(20));
5. degeneracy between clouds and heavy atmosphere (cases (21)–(23));
6. comparison of cloud retrieval models (cases (22), (24)–(25)),
7. retrievals of featureless spectra (cases (26)–(28)).

3.1. T–p Profile

Isothermal T – p profiles, as commonly used in transit retrieval studies, are too simplistic for cloud studies. In our experiments, we first assume T – p profiles with a dry adiabatic lapse rate (DALR) in the troposphere, a moist adiabatic lapse rate (MALR) in the cloud-forming region, and a colder isothermal profile above the tropopause. However, to be compatible with the computing requirements in retrieval, we simplify it to a *TauREx* “two-point” profile (“N-point” performances evaluated in Changeat et al. 2021). We define as T_c and p_c the temperature and pressure at the tropopause and T_{surf} the temperature at 10 bar. The two-point profile is a fit to the points (T_c, p_c) and $(T_{\text{surf}}, p_{\text{surf}})$. One condition of cloud formation is that the atmospheric pressure exceeds the SVP, which is influenced by the thermal gradient in the lower atmosphere controlled by T_{surf} and T_c . These factors determine the location where the pressure exceeds the SVP and therefore the cloud formation.

3.2. Instrumental Performance

The new generation of space-based facilities, such as JWST and Ariel, will deliver unprecedentedly high-quality data in terms of wavelength coverage, signal-to-noise ratio, and spectral resolution. We select as nominal case transit spectra covering 0.4–14 μm , at a spectral resolution of 100, with

⁶ <https://exomol.com>

⁷ <https://hitran.org>

Table 2
YunMa Retrieval Experimental Results

Case	GTPs						Posteriors							
	$\log(f_{\text{sed}})$	$\log(p_c)$ log(bar)	$\log(X_{\text{H}_2\text{O}})$	X_{N_2}	Error (ppm)	λ (μm)	Res.	R_p \mathcal{R}_{J_e}	$\log(f_{\text{sed}})$	$\log(X_{\text{H}_2\text{O}})$	X_{N_2}	$\log(p_c)$ log(bar)	T_c (K)	T_{surf} (K)
1	0.48	-3	-1	...	10	0.4-14	100	$0.20^{+6.98e-06}_{-4.06e-06}$	$0.29^{+0.00}_{-0.00}$	$-1.10^{+0.00}_{-0.00}$
2	0.48	-2.7	-1	...	10	0.4-14	100	$0.20^{+1.72e-03}_{-1.64e-03}$	$0.48^{+0.03}_{-0.02}$	$-0.99^{+0.17}_{-0.18}$
3	0.48	-2.3	-1	...	10	0.4-14	100	$0.20^{+3.54e-03}_{-3.31e-03}$	$-0.85^{+0.93}_{-1.56}$	$-1.34^{+0.47}_{-1.14}$
4	0.48	-2	-1	...	10	0.4-14	100	$0.20^{+5.30e-03}_{-1.75e-03}$	$-1.11^{+0.97}_{-0.75}$	$-1.57^{+0.79}_{-0.85}$
5	0.48	-1	-1	...	10	0.4-14	100	$0.20^{+1.63e-03}_{-2.39e-03}$	$-0.76^{+0.97}_{-0.81}$	$-1.22^{+0.31}_{-0.96}$
6	-2	-2.7	-1	...	10	0.4-14	100	$0.20^{+2.80e-03}_{-2.10e-03}$	$-2.02^{+0.09}_{-0.11}$	$-0.91^{+0.28}_{-0.22}$
7	1	-2.7	-1	...	10	0.4-14	100	$0.20^{+2.12e-04}_{-7.09e-04}$	$1.29^{+0.43}_{-0.28}$	$-0.92^{+0.02}_{-0.07}$
8	0.48	-2.7	-2	...	10	0.4-14	100	$0.20^{+2.50e-05}_{-2.63e-05}$	$0.44^{+0.00}_{-0.01}$	$-2.02^{+0.01}_{-0.01}$
9	0.48	-2.7	-0.3	...	10	0.4-14	100	$0.20^{+4.43e-03}_{-3.44e-03}$	$-1.22^{+0.95}_{-1.06}$	$-0.75^{+0.34}_{-0.37}$
10	0.48	-2.7	-1	...	1	0.4-14	100	$0.20^{+1.62e-04}_{-1.63e-04}$	$0.48^{+0.00}_{-0.00}$	$-1.00^{+0.02}_{-0.02}$
11	0.48	-2.7	-1	...	30	0.4-14	100	$0.21^{+8.61e-04}_{-3.76e-03}$	$0.77^{+0.85}_{-0.28}$	$-0.50^{+0.11}_{-0.37}$
12	0.48	-2.7	-1	...	10	0.4-14	10	$0.21^{+7.33e-04}_{-2.85e-03}$	$0.96^{+0.70}_{-0.44}$	$-0.47^{+0.10}_{-0.29}$
13	0.48	-2.7	-1	...	10	1-14	100	$0.20^{+2.47e-03}_{-3.95e-04}$	$0.51^{+0.10}_{-0.01}$	$-0.76^{+0.25}_{-0.04}$
14	0.48	-2.7	-1	...	10	0.4-14	100	$0.20^{+7.53e-04}_{-9.24e-04}$	$0.42^{+1.04}_{-1.15}$	$-0.62^{+0.07}_{-0.08}$...	$-2.39^{+0.35}_{-0.33}$	$124.55^{+56.56}_{-82.38}$	$1264.49^{+151.39}_{-121.10}$
15	0.48	-2.7	-1	...	1	0.4-14	100	$0.20^{+1.39e-05}_{-1.44e-05}$	$0.20^{+0.01}_{-0.02}$	$-0.71^{+0.00}_{-0.00}$...	$-2.54^{+0.02}_{-0.01}$	$194.23^{+1.25}_{-0.88}$	$1043.45^{+4.14}_{-1.91}$
16	0.48	-2.3	-1	...	10	0.4-14	100	$0.21^{+7.71e-04}_{-1.02e-03}$	$0.25^{+1.16}_{-1.24}$	$-0.09^{+0.06}_{-0.10}$...	$-1.21^{+0.52}_{-0.50}$	$175.25^{+31.87}_{-86.34}$	$1237.49^{+508.48}_{-528.66}$
17	0.48	-2	-1	...	10	0.4-14	100	$0.20^{+6.21e-03}_{-4.08e-03}$	$-1.84^{+2.85}_{-0.97}$	$-2.88^{+2.82}_{-3.98}$...	$-0.95^{+0.55}_{-1.03}$	$150.58^{+44.47}_{-122.39}$	$1000.70^{+706.26}_{-352.40}$
18	0.48	-2.7	-1	10^{-12}	10	0.4-14	100	$0.21^{+2.88e-04}_{-1.72e-04}$	$0.62^{+0.03}_{-0.02}$	$-0.86^{+0.11}_{-0.12}$	$0.11^{+0.03}_{-0.03}$
19	0.48	-2.7	-1	0.1	10	0.4-14	100	$0.20^{+4.02e-04}_{-6.63e-04}$	$0.67^{+0.59}_{-0.12}$	$-0.96^{+0.23}_{-0.24}$	$0.37^{+0.09}_{-0.07}$
20	0.48	-2.7	-1	0.5	10	0.4-14	100	$0.20^{+1.82e-04}_{-2.84e-04}$	$0.39^{+0.12}_{-0.10}$	$-1.03^{+0.33}_{-0.46}$	$0.74^{+0.13}_{-0.13}$
21	0.48	-2.7	-1	0.5	10	0.4-14	100	$0.20^{+3.96e-04}_{-4.70e-04}$	$0.56^{+0.96}_{-1.19}$	$-0.69^{+0.23}_{-0.33}$	fixed to 0.5	$-1.69^{+0.32}_{-0.36}$	$103.11^{+70.32}_{-69.47}$	$1443.65^{+373.77}_{-310.00}$
22	0.48	-2.7	-1	0.5	10	0.4-14	100	$0.20^{+3.87e-04}_{-3.48e-04}$	$0.60^{+0.91}_{-0.98}$	$-0.82^{+0.34}_{-0.44}$	$0.68^{+0.15}_{-0.21}$	$-1.77^{+0.32}_{-0.36}$	$111.53^{+67.82}_{-73.41}$	$1553.21^{+276.23}_{-399.58}$
23	0.48	-2.7	-1	0.5	10	0.4-14	100	$0.20^{+7.03e-04}_{-6.06e-04}$	$0.42^{+1.00}_{-1.02}$	$-0.11^{+0.07}_{-0.09}$	fixed to 0	$-1.62^{+0.42}_{-0.42}$	$130.80^{+52.94}_{-86.47}$	$1071.37^{+399.26}_{-381.53}$
24	0.48	-2.7	-1	0.5	10	0.4-14	100	$0.20^{+6.13e-04}_{-9.64e-04}$...	$-1.90^{+0.81}_{-1.36}$	$0.78^{+0.12}_{-0.18}$	$-2.27^{+1.22}_{-1.13}$	$327.32^{+120.86}_{-245.79}$	$1370.26^{+430.51}_{-535.90}$
25	0.48	-2.7	-1	0.5	10	0.4-14	100	$0.20^{+3.80e-04}_{-3.47e-04}$...	$-0.95^{+0.32}_{-0.47}$	$0.74^{+0.14}_{-0.16}$	$-1.72^{+0.38}_{-0.35}$	$48.60^{+51.83}_{-32.93}$	$1577.37^{+285.82}_{-401.02}$
26	0.48	-2.3	-1	0.1	10	0.4-14	100	$0.21^{+4.38e-04}_{-8.41e-04}$	$-2.27^{+0.67}_{-0.46}$	$-3.80^{+2.23}_{-3.28}$	$0.78^{+0.13}_{-0.18}$	$-1.31^{+0.82}_{-1.06}$	$86.15^{+65.71}_{-61.63}$	$1136.49^{+542.96}_{-444.41}$
27	0.48	-2.3	-1	0.5	10	0.4-14	100	$0.20^{+4.50e-04}_{-6.77e-04}$	$-2.20^{+0.60}_{-0.48}$	$-4.33^{+2.36}_{-3.92}$	$0.72^{+0.17}_{-0.21}$	$-1.29^{+0.80}_{-1.11}$	$62.56^{+65.77}_{-45.02}$	$1052.87^{+540.75}_{-342.32}$
28	0.48	-2.3	-1	0.5	1	0.4-14	100	$0.20^{+4.66e-05}_{-4.72e-05}$	$0.46^{+0.08}_{-0.17}$	$-1.20^{+0.08}_{-0.10}$	$0.90^{+0.02}_{-0.01}$	$-2.31^{+0.08}_{-0.11}$	$187.18^{+15.16}_{-13.86}$	$976.55^{+33.41}_{-41.97}$

Note. The GTPs assumed in the simulations are listed in the left columns of the table, and the retrieved posteriors are on the right. The retrieval priors are listed in Table 1. In each case, we test the model sensitivity of the atmospheric parameters of f_{sed} , p_c , $X_{\text{H}_2\text{O}}$, X_{N_2} , and the observational parameters of the error, wavelength coverage (λ), and the spectral resolution. We choose the nominal value of p_c as 2×10^{-3} bar, where the cloud is not either too thin to be detected or too thick to block the spectroscopic features.

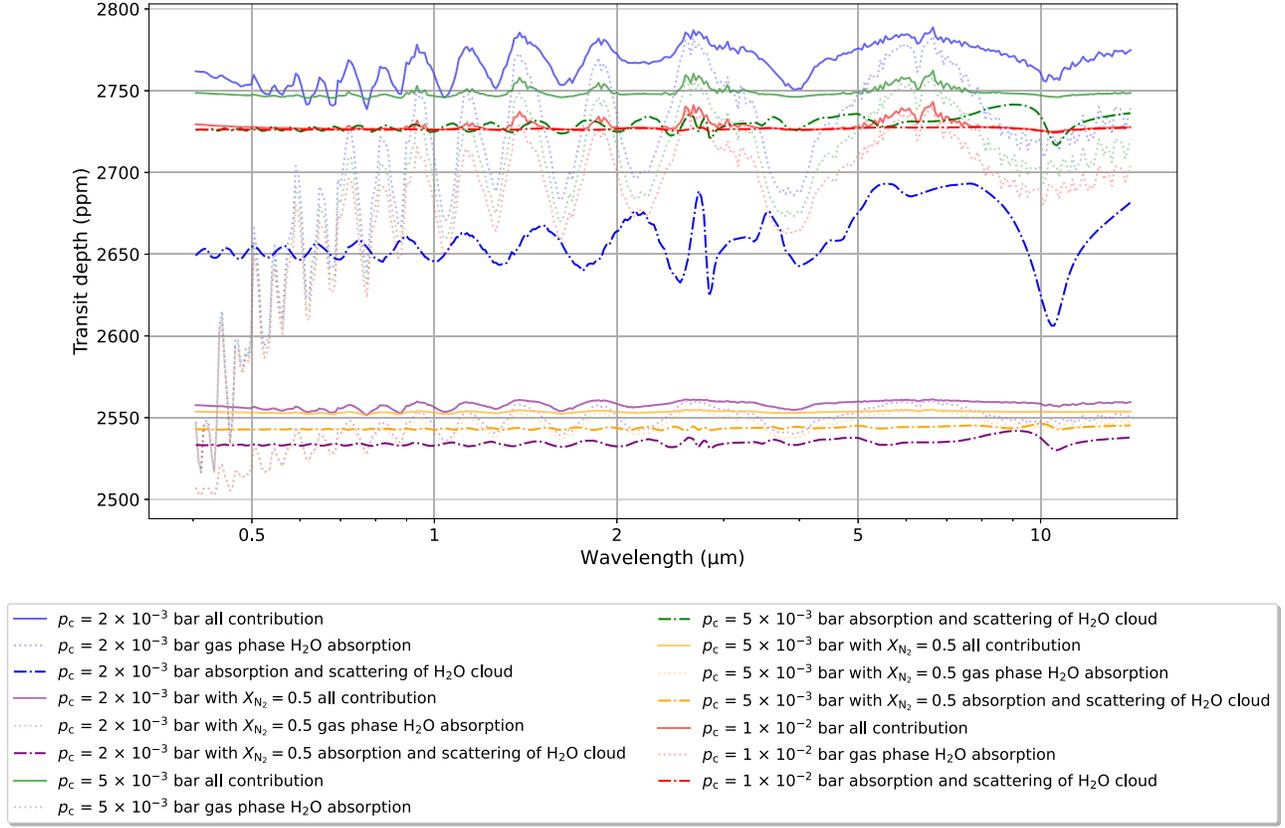


Figure 6. Simulated transit spectra of cloudy super-Earths using *YunMa*. Solid line: total transit depth with all contributions included. Dotted–dashed line: water ice clouds. Faint-dotted line: water vapor. Blue, green, and red lines: H₂/He-dominated, cloudy atmospheres with different p_c (see legend). Purple and yellow lines: heavier cloudy atmospheres with 50% N₂.

10 ppm uncertainty across wavelengths. We chose $0.4 \mu\text{m}$ as the blue cutoff to maximize the information content about Rayleigh scattering and $14 \mu\text{m}$ as the red cutoff to maximize the information content about the water vapor and atmospheric temperature for the type of planets considered here. The choice of wavelength coverage and precision is inspired by current and planned instrumentation, while not trying to reproduce a specific observatory with its own limitations. The focus of this paper is on the retrievability of clouds and not on the performance of a specific facility.

4. Results

4.1. Simulated Transit Spectra with *YunMa*

We present here the transmission spectra generated with *YunMa* of a cloudy super-Earth. Figure 6 shows five examples at resolving power of 100, which corresponds to the ground truths of some of the retrieval cases in Table 2: $p_c = 2 \times 10^{-3}$ bar with $X_{\text{N}_2} = 0$ (blue, cases (2), (10), (11), (14), and (15)), and $= 0.5$ (purple, cases (20)–(25)); $p_c = 1 \times 10^{-2}$ bar with $X_{\text{N}_2} = 0$ (green, cases (3) and (16)), and $= 0.5$ (yellow, cases (27) and (28)) and one case of opaque cloud with $p_c = 1 \times 10^{-2}$ bar in the H₂/He-dominated atmosphere (red, cases (4) and (17)). The planetary and atmospheric parameters are listed in Tables 1 and 2. All the simulations contain baseline 10% H₂O abundance across the atmosphere, which is then altered by the cloud formation. The rest of the atmosphere is N₂ and H₂/He. We select $f_{\text{sed}} = 3$ for all the scenarios. The simulation results are summarized in Table 3.

In the experiment without N₂ and p_c set to 2×10^{-3} bar, clouds form at high altitude, where the atmospheric density (ρ_a) is low compared to the cases where $p_c = 5 \times 10^{-3}$ bar and $= 1 \times 10^{-2}$ bar. Here, the sedimentation velocity (v_f) is small with small cloud particle radii and number density. The cloud contribution (blue dashed–dotted line) has a mean transit depth of 2660 ppm and σ_{spec} of 17 ppm. It is an optically thin cloud, which does not completely block the spectral features shaped by water vapor absorption (blue dotted line). When $p_c = 1 \times 10^{-2}$ bar, clouds form at relatively low altitudes, where the atmospheric density (ρ_a) is high. Here, v_f is large, and the cloud particles have relatively large radii and number density, which increase the opacity. The cloud contribution (red dashed–dotted line) has a mean flux depth of 2727 ppm and σ_{spec} of 0.62 ppm. Since the clouds are optically thick, they contribute significantly to the mean transit depth and obscure the spectral features of water vapor (red dotted line). However, the water vapor features are still able to show due to the low altitude of the clouds. Still, the spectral deviation is only 12.16 ppm, where the spectroscopic features have a high chance of being hidden by the observational uncertainty. $p_c = 5 \times 10^{-3}$ bar is an intermediate case regarding the simulated cloud altitude and opacity. The simulation suggests that the intermediate combination of these two cloud properties does not result in more significant atmospheric features than in other cases.

In atmospheres with relatively high mean molecular weight—and therefore small scale height—for the same value of p_c , the transit depth is smaller, as expected. In the case with $X_{\text{N}_2} = 0.5$, and $p_c = 2 \times 10^{-2}$ and $= 5 \times 10^{-2}$ bar, the mean

Table 3
Atmospheric and Cloud Parameters Included in *YunMa* Simulations

	p_c	2×10^{-3} (bar)	5×10^{-3} (bar)	1×10^{-2} (bar)	2×10^{-3} (bar)	5×10^{-3} (bar)
	X_{N_2}	0.5	0.5
Parameter	Unit					
All contribution, mean	ppm	2770.17	2750	2729.82	2558.65	2553.71
All contribution, std (σ_{spec})	ppm	15.72	11.59	12.16	2.58	0.62
Cloud contribution, mean	ppm	2659.65	2729.60	2726.53	2534.78	2542.78
Cloud contribution, std	ppm	17.21	4.39	0.62	2.15	0.85
MMW (bottom of the atmosphere)	g mol^{-1}	3.88	3.88	3.88	16.73	16.73
Atmospheric pressure (cloud base)	bar	2.34×10^{-3}	6.40×10^{-3}	1.43×10^{-2}	2.34×10^{-3}	6.40×10^{-3}
Atmospheric pressure (cloud deck)	bar	8.54×10^{-4}	2.34×10^{-3}	4.28×10^{-3}	8.54×10^{-4}	2.33×10^{-3}
Cloud MMR (cloud base)	...	6.35×10^{-4}	9.51×10^{-5}	2.00×10^{-4}	2.92×10^{-4}	4.35×10^{-4}
Cloud MMR (cloud deck)	...	7.56×10^{-8}	2.92×10^{-7}	7.28×10^{-8}	3.49×10^{-7}	1.33×10^{-6}
v_f	m s^{-1}	6.54–7.53	6.07–6.67	4.43–6.41	7.14–7.97	4.37–7.28
r_c	μm	9.03–6.39	9.79–8.88	12.13–9.35	13.99–11.09	25.53–13.66
N (cloud base)	m^{-3}	1.03×10^4	3.19×10^4	7.47×10^4	1.28×10^4	8.22×10^3
N (cloud deck)	m^{-3}	1.45×10^1	5.72×10^1	2.24×10^1	1.28×10^1	7.16×10^1

Note. The corresponding transit depths are also reported. Molecular mixing ratio (MMR) and mean molecular weight (MMW). The pressure at the bottom of the atmosphere is assumed to be 10 bar.

value of the transit depths are ~ 200 ppm smaller than those in a H_2/He -dominated atmosphere. Here, the cloud particles form at higher ρ_a and therefore have larger particle size and larger number density compared to those formed in the H_2/He -dominated atmospheres. The spectrum has σ_{spec} of 2.58 and 0.62 ppm, which are negligible compared to the observational uncertainty.

Besides p_c and X_{N_2} , we also have tested different f_{sed} to understand how this parameter controls the cloud microphysics. The particle radii, r_c , number density, and transit spectra across all the cloud pressure levels and obtained with different f_{sed} are shown in Figure 7. Here, we note that, from the results, the cloud particle sizes increase with f_{sed} while the number densities at each layer behave reversely. Also, which is easy to understand, the more atmospheric layers with clouds, the larger the optical depth.

4.2. Retrieval Results

We show in this section how *YunMa* performs with different model assumptions and ground truth parameters (GTPs), following the approach described in Section 3. GTPs and priors are listed in Table 1. The retrieved values and one standard deviation (1σ) of the posterior distributions obtained for all the simulated cases are summarized in Table 2.

4.2.1. Sensitivity Studies to Key Atmospheric Parameters

We have performed sensitivity studies to test how the model behaves when changing the key atmospheric parameters, including p_c , f_{sed} , $X_{\text{H}_2\text{O}}$, X_{N_2} (cases (1)–(9)). Cases (1)–(5) test the effects of different p_c to the transit spectra and retrievals. Tuning p_c alters the cloud altitude and the optical thickness: in these cases, the larger is p_c ; the more opaque, meanwhile, the lower altitude becomes the clouds. The significance of spectroscopic features owns to both factors. Generally speaking, the more significant features are, the easier it is to retrieve the atmospheric parameters. Here, we test p_c from 10^{-3} to 10^{-1} bar, which is a much broader range than the one considered in previous literature about K2-18 b. The transit spectra of cases (2) and (4) are shown in Figure 6 (blue and red solid lines,

respectively). We choose $p_c = 2 \times 10^{-3}$ bar as nominal case and show the corresponding posterior distributions in Figure 8.

Cases (2), (6), and (7) test the impact on the transit spectra and retrievals of the sedimentation efficiency (f_{sed}), which controls the cloud microphysics in the model. In case (6), we set the sedimentation efficiency f_{sed} to 0.01, i.e., the downward sedimentation of the cloud particles is relatively slow compared to the net upward molecular mixing of the condensable species. By contrast, in case (7) where $f_{\text{sed}} = 10$, we have a larger downward draft velocity scale compared to the upward one. In analyzing the results, we utilize the terms *accuracy* to indicate that our retrieved result is in a certain range of the ground truth and *precision* to indicate the 1σ of the posteriors. In the simple cases, when $f_{\text{sed}} = 0.01$ and 3 (cases (2) and (6)), the accuracy levels of f_{sed} , $X_{\text{H}_2\text{O}}$ and R_p are $> 90\%$. In comparison, in the scenario where $f_{\text{sed}} = 10$ (case (7)), the accuracy level is $> 70\%$. This result is not unexpected, as high f_{sed} scenarios tend to have negligible impact on the planetary transit depth due to thinner cloud layers and smaller number density (N) compared to a low f_{sed} scenario, e.g., the $f_{\text{sed}} = 10$ and 100 cases in Figure 7.

In cases (8) and (9), we modulate the amount of condensable gas, here, represented by the water vapor mixing ratio ($X_{\text{H}_2\text{O}}$). In the cases of our experiment, the clouds start to form when $X_{\text{H}_2\text{O}}$ reaches the significance of 1×10^{-3} . When $X_{\text{H}_2\text{O}} = 0.01$ (case (8)), a thin cloud with low opacity may form, the water vapor spectral features are visible, and the retrieved values of $\log(X_{\text{H}_2\text{O}})$ and $\log(f_{\text{sed}})$ have $> 90\%$ accuracy. By contrast, a higher mixing ratio of the condensable gas increases the partial pressure and contributes to the condensation process. This is, for instance, the case of $X_{\text{H}_2\text{O}} = 0.5$ (case (9)) where the cloud is thick and largely blocks the spectral features, making the retrieval of the atmospheric parameters difficult.

4.2.2. Sensitivity Studies to Data Quality

We then test how *YunMa*'s performances degrade when we compromise with the data quality, for instance, as follows: uncertainties of 30 ppm in case (11) and spectral resolution of 10 in case (12), which should have similar effects. We move the blue cutoff at longer wavelengths in case (13). From the experiments on observational data quality, case (10)'s Bayesian

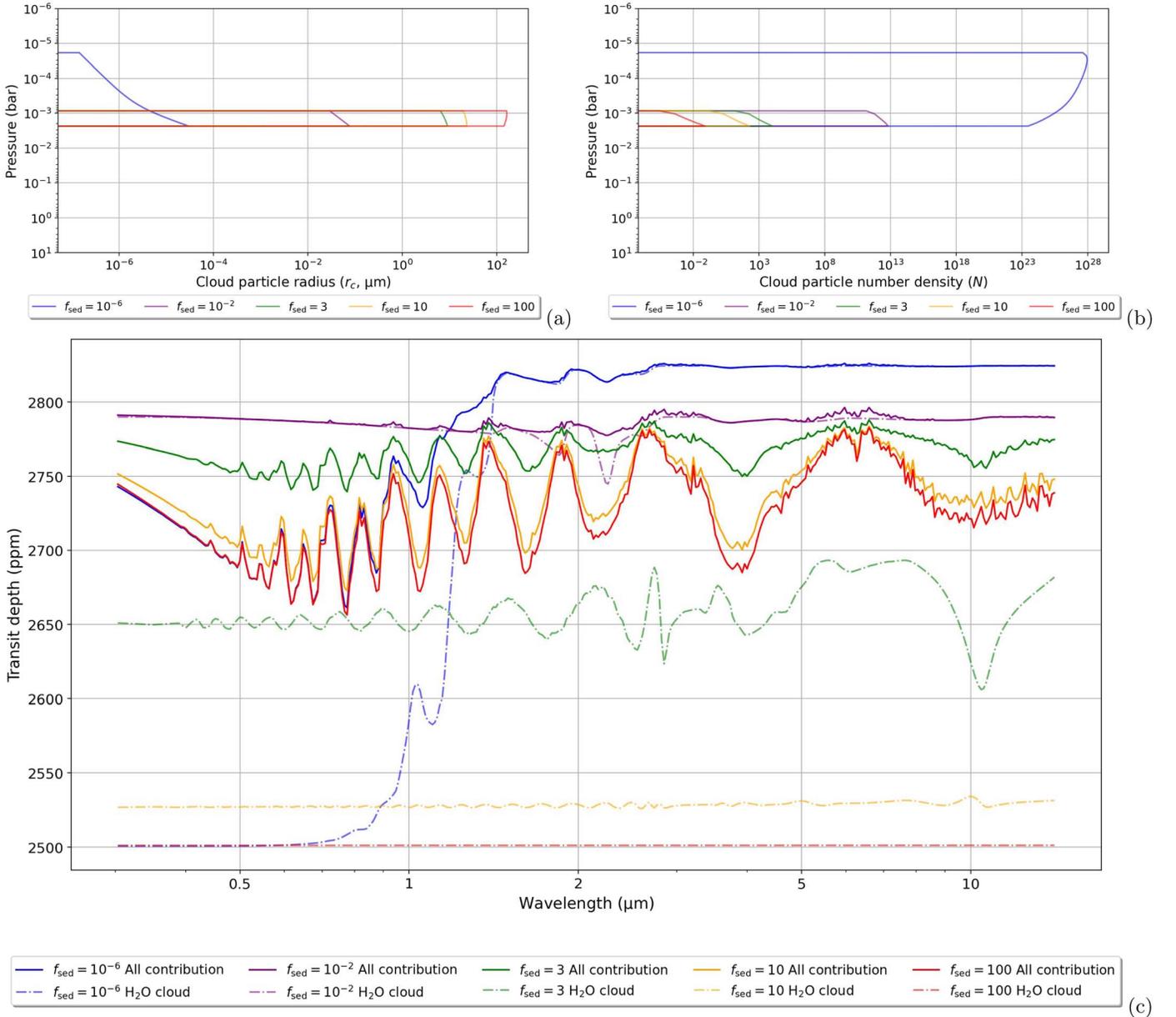


Figure 7. Test results for r_c (top left), N (top right), and the transit spectrum (bottom) with different sedimentation efficiency f_{sed} . The cloud layer shrinks with the increasing of f_{sed} .

evidence (4568.52) compared to the ones calculated for case (2) (3755.80) and case (11) (3368.29) showcases how the retrieval performance improves when the observational uncertainties are small. Case (2)’s performance surpasses that of case (12), as the spectral resolution of the transit spectrum used as input to the retrieval is higher. In case (13), we omitted the information in the optical wavelengths, which means we have less information about the cloud scattering properties. The retrieval performances are degraded compared to case (2), which includes the optical wavelengths.

4.2.3. Retrievals of Atmospheric Thermal Profiles

In cases (14)–(17), we retrieved the T – p profiles as free parameters for different p_c to test how *YunMa* performs with increasingly complex model assumptions and which parameters may be problematic in these retrievals. Our results show

that both the T – p profiles and cloud parameters can be constrained, although the retrieved gas phase mixing ratio and f_{sed} distribution may have large standard deviations in some of the cases.

4.2.4. Addition of N_2 in Retrievals

Inactive, featureless gases, such as N_2 , inject much uncertainty in the retrieval. In cases (18)–(20), we include different amounts of the inert and featureless gas N_2 in the atmosphere. N_2 may exist in super-Earths’ atmospheres, as it happens for solar system planets at similar temperatures. Being N_2 heavier than H_2O , we adjust the mean molecular weight and the scale heights accordingly by modulating X_{N_2} . Heavier atmospheres have smaller scale heights: the spectral features are less prominent and harder to detect. We first try to retrieve only R_p , f_{sed} , X_{H_2O} , and X_{N_2} . The results show that, despite the

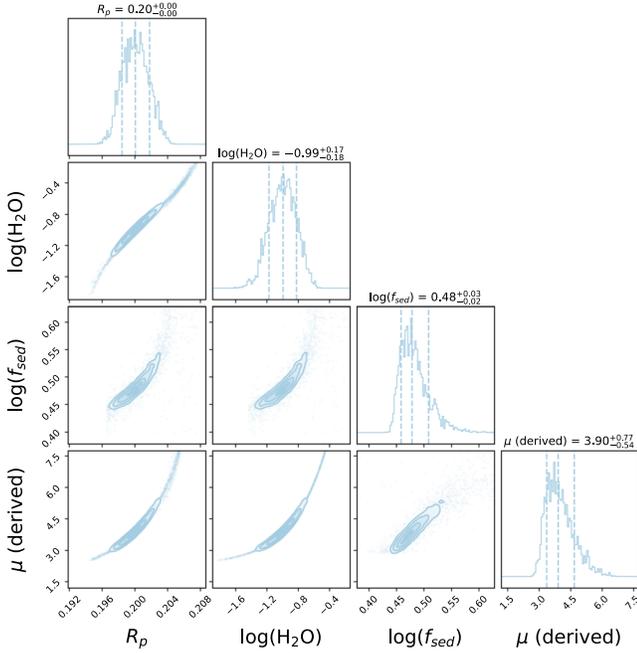


Figure 8. Example of cloud retrieval using *YunMa* integrated in *TauREx 3* (case (2) in Table 2). The sedimentation efficiency (f_{sed}), which is the main parameter controlling the cloud microphysics, is well recovered together with the vapor water mixing ratio.

minimal spectral features of the cloudy heavy atmospheres due to N_2 injection, *YunMa* is still able to retrieve these atmospheric parameters in simple cases.

4.2.5. Degeneracy between Clouds and Heavy Atmosphere

In cases (21)–(25), we conduct more complex retrievals to test the degeneracy between clouds and X_{N_2} . The corresponding transit spectra are shown in Figure 6 (purple lines). Similar to N_2 , the existence of clouds mitigates the spectral features, and the difference between these two scenarios may be difficult to distinguish from the current data quality. Case (21) retrieves these parameters except X_{N_2} fixed to 50% for comparison with case (22) to investigate the degeneracy imposed by the uncertainty of X_{N_2} . The result of case (21) shows how the atmospheric parameters, with the exception of X_{N_2} , can be retrieved in a heavy atmosphere with X_{N_2} fixed to the ground truth. When we include the uncertainty of X_{N_2} (case (22)), the GTPs for R_p , $X_{\text{H}_2\text{O}}$, f_{sed} , T_{surf} , and T_{top} still fall into the 2σ confidence range, where R_p , $X_{\text{H}_2\text{O}}$, and f_{sed} have accuracy levels $> 60\%$ and T_{surf} and $T_{\text{top}} > 45\%$. An illustration of the retrieved T – p profile is shown in Figure 9. f_{sed} is significantly less constrained in case (22) than in case (21) due to the uncertainty of X_{N_2} .

One hypothesis is that, if we do not include N_2 among the priors, the model will add clouds to compensate for the missing N_2 . In case (23), we test this hypothesis by forcing X_{N_2} to zero and then monitor the cloud parameters in the posteriors obtained. The results suggest that potential degeneracy could happen, plotted in Figure 10: when we omit N_2 among the priors, the retrieval tries to compensate for the missing radiative-inactive gas by decreasing R_p and T_{surf} , while increasing $X_{\text{H}_2\text{O}}$, and the Bayesian evidence of case (23) (3757.08) is close to that from case (22) (3757.40).

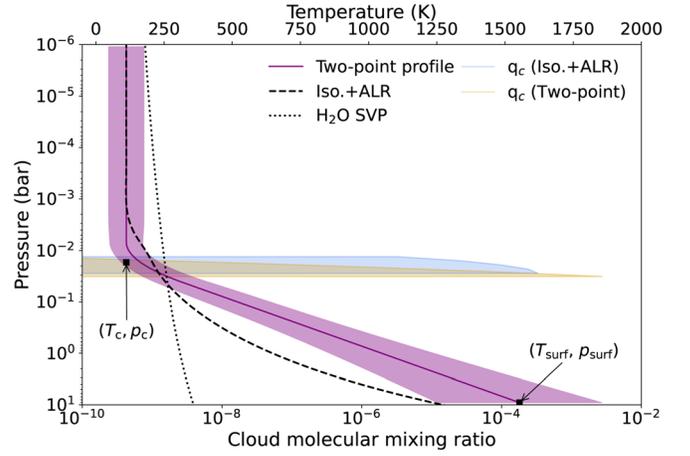


Figure 9. Retrieved T – p profile of case (22) in Table 2. The isothermal temperature (T_c), surface temperature (T_{surf}), and the pressure where the isothermal profile starts (p_c) are retrieved using *YunMa* and indicated by squares in the plot. The solid purple line indicates the two-point profile fitted from the retrieved T_c , T_{surf} , and p_c values. The shaded area indicates the standard deviation of the posterior distribution. We also plot a T – p profile with an isothermal upper atmosphere, a DALR-estimated lower atmosphere and a MALR-estimated cloud-forming region to show how the two-point profile can deviate from the adiabatic lapse rate (ALR) when similar cloud cover forms. The two-point profile is a useful approximation to estimate cloud formation while reducing the retrieval computing time.

4.2.6. Comparison of Cloud Retrieval Models

In the experiment of model comparison, case (24) simulates the forward spectra with *YunMa* cloud microphysics and then retrieves the atmospheric parameters with another simplified cloud retrieval framework. The simplified clouds are described as an opaque cloud deck across wavelengths in *TauREx 3*, which is commonly used to retrieve data from the last decades with narrow wavelength coverage, e.g., Hubble Space Telescope/WFC3. In a nonopaque case, we compare the retrieved results from the simple opaque cloud retrieval model in *TauREx 3* (case (24)) and *YunMa*. The posteriors using the two different cloud models are compared with each other in Figure 10. The former shows relatively flat posterior distributions of the cloud and atmospheric parameters than using *YunMa*, and in this case, lower accuracy and precision of the retrieved results. In case (25), we deliberately omit cloud parameters in the retrieval priors and learn if and how other parameters can compensate for those missing. Without the cloud in the prior (case (25)), the results show a lack of constraints on the T – p profile while there is a comparable performance on other parameters with case (22).

4.2.7. Retrievals of Featureless Spectra

In cases (26)–(28), we retrieve the atmospheric parameters from spectra with minimal spectroscopic features, where the spectral standard deviation (σ_{spec}) is less than 1 ppm. The spectra are featureless due to heavy atmospheres and cloud contribution ($p_c = 5 \times 10^{-2}$ bar). The transit spectrum for cases (27) and (28) is shown in Figure 6, yellow line: the spectral signal is very small compared to the observational uncertainty (10 ppm). In case (28), we unrealistically decrease the uncertainty to 1 ppm at resolving power 100 to evaluate *YunMa*'s performance with idealized data quality. As expected, the retrieval performance greatly improves when the

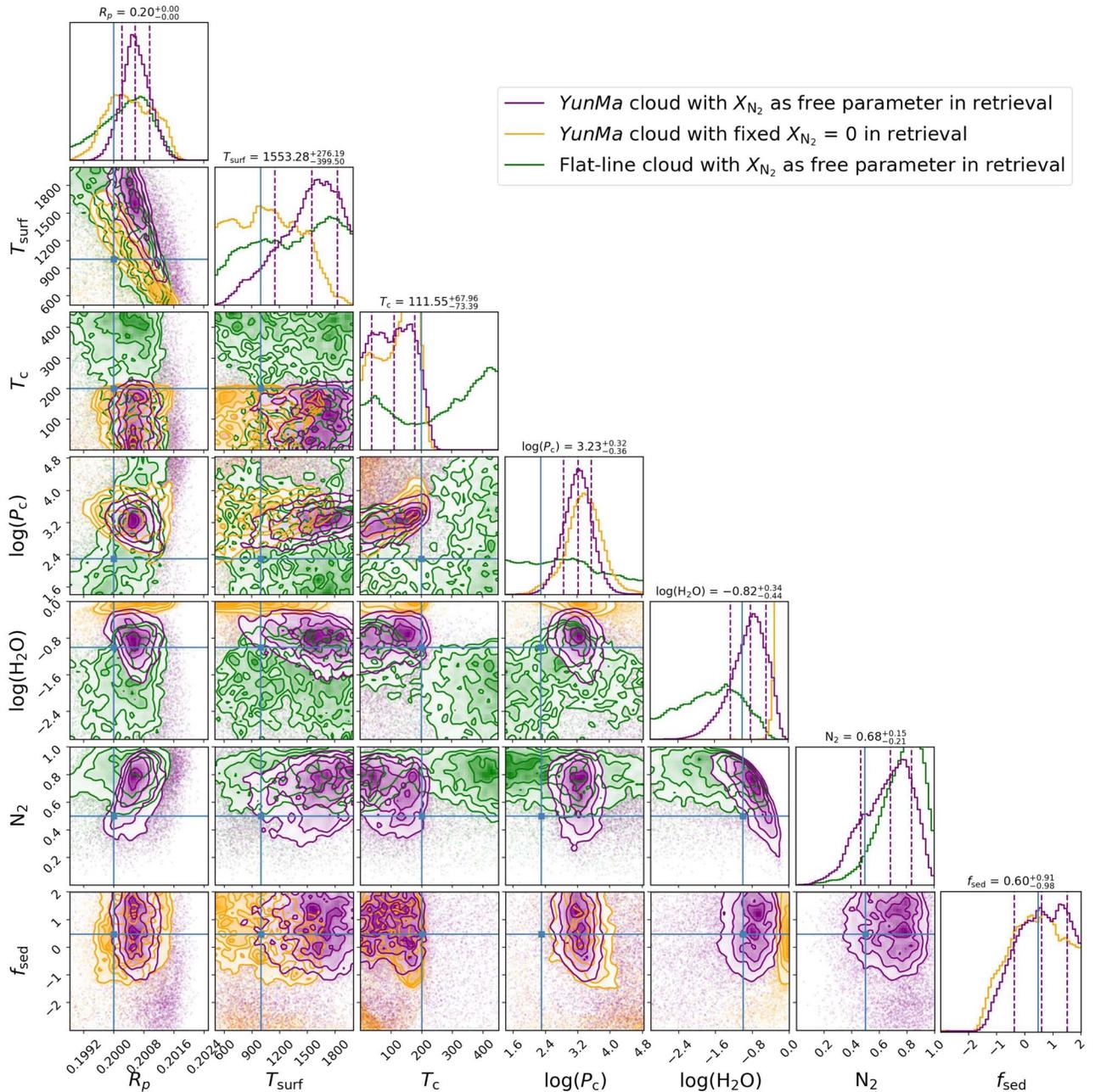


Figure 10. Retrieval posteriors for cases (22) (purple), (23) (orange), and (24) (green) in Table 2. Blue crosses indicate the ground truth parameters, and the vertical dashed lines in histograms indicate the 1σ and 2σ confidence ranges of the posterior distribution. These three retrievals use the same transit spectrum as input (thin cloud with $X_{\text{N}_2} = 0.5$ in Figure 6) but different retrieval assumptions. Case (22) (purple): H_2O and N_2 are included as priors, and the cloud formation is simulated by *YunMa*. Case (23) (orange): same as case (22) except that N_2 is not included among the priors. Case (24) (green): same as case (22) except that the cloud is simulated by a simpler model from *TauREx*, in which the atmosphere becomes opaque below the cloud deck. The only retrieved cloud parameter for the simpler model is the cloud deck pressure, which is not shown here for simplicity.

observational uncertainty unrealistically decreases to 1 ppm (Figure 11).

5. Discussion

5.1. Transit Spectroscopy Using *YunMa*

To understand the performances of *YunMa* in detail, we performed retrieval experiments for over a hundred cases, with different chemistry models, atmospheric and cloud scenarios for super-Earths and/or sub-Neptunes, hot-Jupiters and brown dwarfs. A variety of cloud species were modeled and analyzed.

This paper presents a selection of representative examples. We have validated the A-M cloud size distribution in *YunMa* against those from previous literature simulating NH_3 clouds in Jupiter’s atmosphere, KCl clouds on artificial large exoplanets and brown dwarfs, and H_2O clouds on K2-18 b. We have also tested a number of numerical settings, including fitting methods, tolerances, and retrieval samplings.

In *YunMa*, the mixing ratios of the condensable gas and the condensate are strongly correlated; therefore, when a cloud forms, the condensable species in the gas phase decreases. Also, a balance is imposed between the upward turbulent

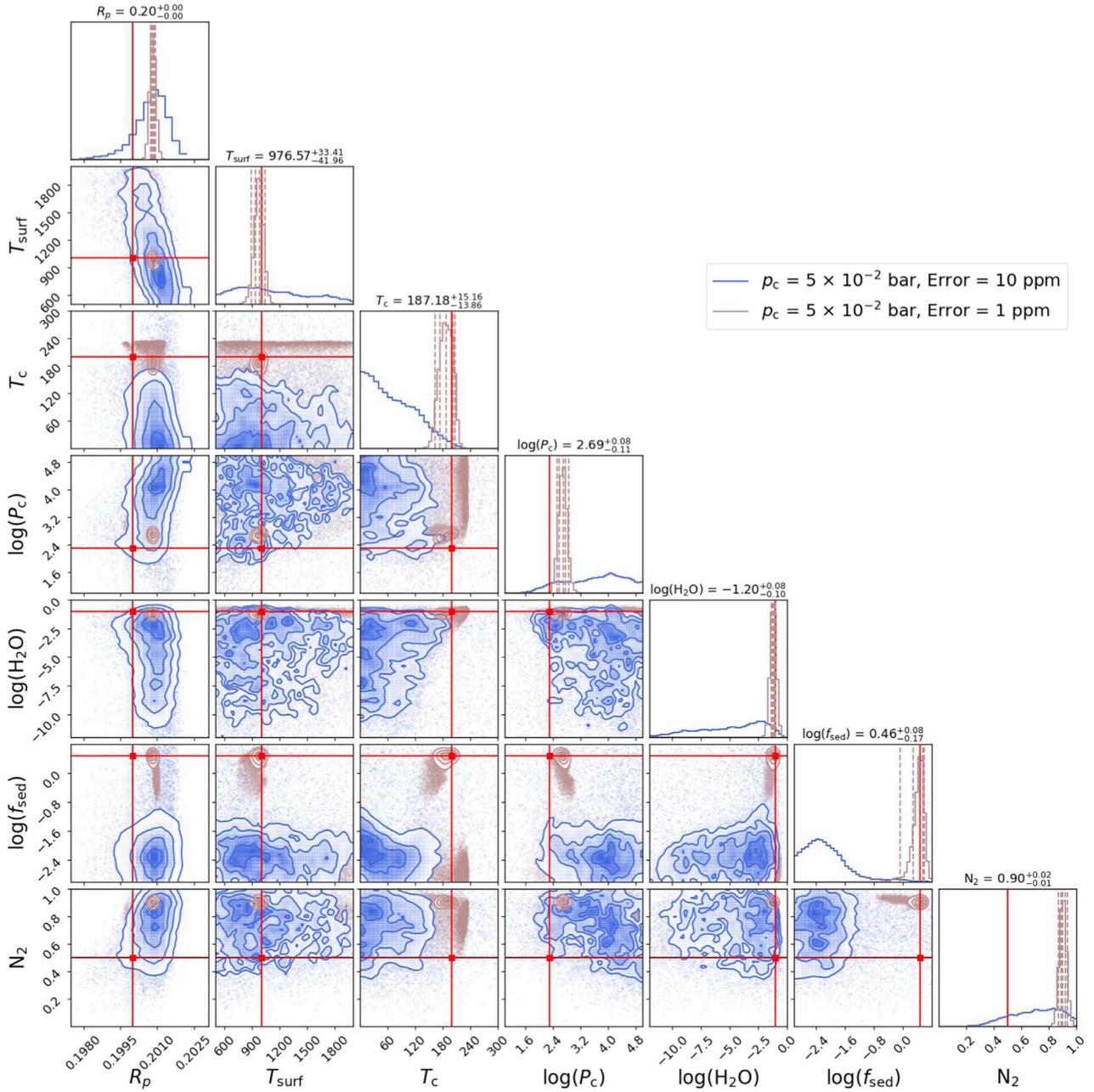


Figure 11. *YunMa* retrieval experiments with thick clouds and $X_{N_2} = 0.5$ (orange line in Figure 6). Observational uncertainty = 10 ppm (blue line, case (27) in Table 2) and = 1 ppm (brown line, case (28) in Table 2). Red lines indicate GTPs.

mixing and the downward sedimentation velocity from the A-M approach (Equation (2)). In the current *YunMa*, at each pressure level, the particle number density represents the total number density of particles with different radii. The Earth measurements shown in Figure 4 of Ackerman & Marley (2001) suggest a bimodal distribution of the particle sizes at the same pressure level. *YunMa* is able to use radius bins with their respective number densities to represent more precisely the cloud distribution in the spectral simulation. For the retrieval calculations, however, we had to simplify this information to reduce the computing time.

The cloud opacity is determined by the cloud particle size and number density; different particle sizes have absorption and scattering peaks at different wavelengths. Optically thick clouds cause transit depths with negligible or no modulations

as a function of wavelength: the atmosphere below the cloud deck is, in fact, undetectable while the atmosphere sounded above the cloud deck is more rarefied. Retrieving information about the atmospheric composition and structure is very difficult in the most extreme cases. For an atmosphere with optically thin clouds, the absorption features due to radiative-active gases (water vapor here) are detectable but less prominent compared with a clear atmosphere. The abundances of these gases can be retrieved largely from their absorption features. If these are condensable species, their abundances further constrain the cloud microphysics. Both the wavelength-dependent features and the overall transit depth help the retrieval performance.

f_{sed} is the ratio between the sedimentation velocity and the turbulent convective velocity. From the definition, a higher f_{sed}

means a shorter sedimentation timescale, as we fixed K to a constant value. A higher sedimentation efficiency leads to larger offsets to the downward draft, constraining the upward supplement of water vapor and cloud formation. On the contrary, for small f_{sed} , the sedimentation timescale is much longer compared to the diffusion timescale, so the condensation continues at a lower pressure to balance the downward sedimentation and upward turbulent mixing; therefore, the cloud region expands. f_{sed} is sensitive to the cloud particles' nucleation rate (Gao et al. 2018). It has a close relationship with the condensate particle size and can be expressed by the particle radius in the lognormal distribution power-law approximation:

$$f_{\text{sed}} = \frac{\int_0^\infty r^{3+\alpha} \frac{dn}{dr} dr}{r_w^\alpha \int_0^\infty r^3 \frac{dn}{dr} dr}, \quad (11)$$

which indicates that small f_{sed} encourages small cloud particle formation. The upward transport is stronger than the sedimentation when the cloud particles are small and vice versa, which is in line with the experimental results in Figures 7(a), (b). The spectrum is sensitive to f_{sed} when this parameter has values between 10^{-1} and 10^{-3} , as shown in Figure 7 (c), indicating the detectability of f_{sed} in this interval.

In our experiments, the particle radii are typically 1–10 μm , so they do not block the Rayleigh scattering slope caused by H_2 and H_2O at the optical wavelengths. For smaller radii (e.g., cyan line in Figure 5), they are expected to contribute more to the optical spectrum, although theoretically it would be hard to form particles at very small sizes according to the nucleation theory in cloud formation (Gao et al. 2018). For any particle radii even smaller presented in this paper, it is just for model test purposes in extreme cases, and we make no efforts to show their detailed analysis here.

Radiative-inactive gases such as N_2 , if present, can change the atmospheric scale height. The increase of scale height decreases the transit depth and dampens its spectroscopic features, as illustrated in Figure 6. As mentioned in previous sections, the presence of radiative-inactive gases cannot be detected directly through spectroscopic signatures. Opaque clouds may behave similarly to inactive gases in mitigating spectroscopic features, leading to a potential degeneracy in retrieval experiments. Illustrated in Figure 10, the comparison between cases (22) and (23) in Table 2 suggests the potential degeneracy between R_p , the baseline condensable gas abundance, T - p profile, and the N_2 abundance. In case (23), we force $X_{\text{N}_2}=0$ in the retrieval to monitor how *YunMa* compensates for the missing gases in the atmosphere. An adjustment of R_p translates the transit depth without an impact on spectroscopic features. The baseline condensable gas abundance and T - p profile are correlated to the formation of clouds. The results indicate that, when N_2 is absent in prior, the mixing ratio of water vapor—the radiative-active gas—is significantly increased to compensate for the missing molecular weight, while more clouds are formed to further reduce the spectroscopic features. The decrease of T_{surf} helps in the same way. While the potential degeneracy exists from the analysis and the results suggest the model behavior in the case of missing radiative-inactive gases, the model chose from statistics the scenario closer to the ground truth in case (22), showing the model's potential in retrieving clouds in heavy atmospheres when it is not opaque.

As mentioned before, the cloud formation and T - p profile are correlated in *YunMa*. Therefore, the presence of optically thin clouds can help to constrain the T - p profile in retrievals: for instance, p_c , T_c , and T_{surf} are well retrieved in case (22). On the contrary, the T - p profile is not well constrained if clouds are completely absent (case (25)), or the microphysics part is removed from the retrieval (case (24)).

When simulating transit observations, *YunMa* uses a 1D approach to estimate the cloud formation at terminators according to the thermal profiles present at those locations. If separate observations of the morning and evening terminators are available, these will help us to understand the impact of atmospheric dynamics and horizontal effects. Similarly, the phase curves or eclipse ingress and egress observations will be pivotal to completing the 3D picture of the planet.

5.2. Cloud Formation with the Next-generation Facilities' Data

With the high-quality transit spectra offered by the next-generation facilities, the uncertainties, wavelength coverage, and spectral resolution will be significantly improved compared to most current data, and a simple opaque cloud model is insufficient for the transit study of next-generation data according to our results; the modeling of cloud radiative transfer and microphysics such as *YunMa* is needed. In our experiments, we chose a nominal 10 ppm as the observational uncertainty and found that clouds can be well-characterized in most experiments. With the next-generation data and *YunMa*, there are still limits in retrieving the featureless spectra. A smaller uncertainty may help in retrieving cloud parameters in the most difficult cases: we therefore adopted an unrealistic 1 ppm to test their detectability in an ideal case (cases (10), (15), and (10)). With 1 ppm, the atmospheres were constrained well, although the spectra were featureless. The results of cases (2) and (13) suggest that broad wavelength coverage is paramount to characterizing clouds well: here, optical wavelengths play a critical role when combined with infrared spectral coverage.

5.3. Numerical Instabilities in Cloud Microphysics Simulations

We have selected the explicit Runge–Kutta method of order 8 (DOP853, Hairer et al. 1993) to solve Equation (2) as it delivered the most stable numerical performance for our experiments. However, numerical instabilities may occur when large relative (*rtol*) and absolute (*atol*) tolerances are chosen. Sometimes, the ODE solver cannot converge for q_t and indicates *no clouds* as a solution due to the numerical instability. Caution is also needed in estimating the cloud mixing ratio (q_c), as q_t might be 2 orders of magnitude larger than q_c , and numerical errors could be injected when q_s is subtracted from q_t . When solving the ODE with too large tolerance, q_t might converge to a certain value, but q_c would be estimated as negligible according to Equations (2) and (1). In other words, even though when the ODE solution for the gas phase seems numerically stable, it might not be precise and accurate enough. In those cases, the retrieval performances are affected, as shown in Figure 12, where multiple islands of solutions in the posteriors are visible; these are caused by numerical instabilities, and the issue is more obvious for larger tolerance values. After many tests, we have decided to use

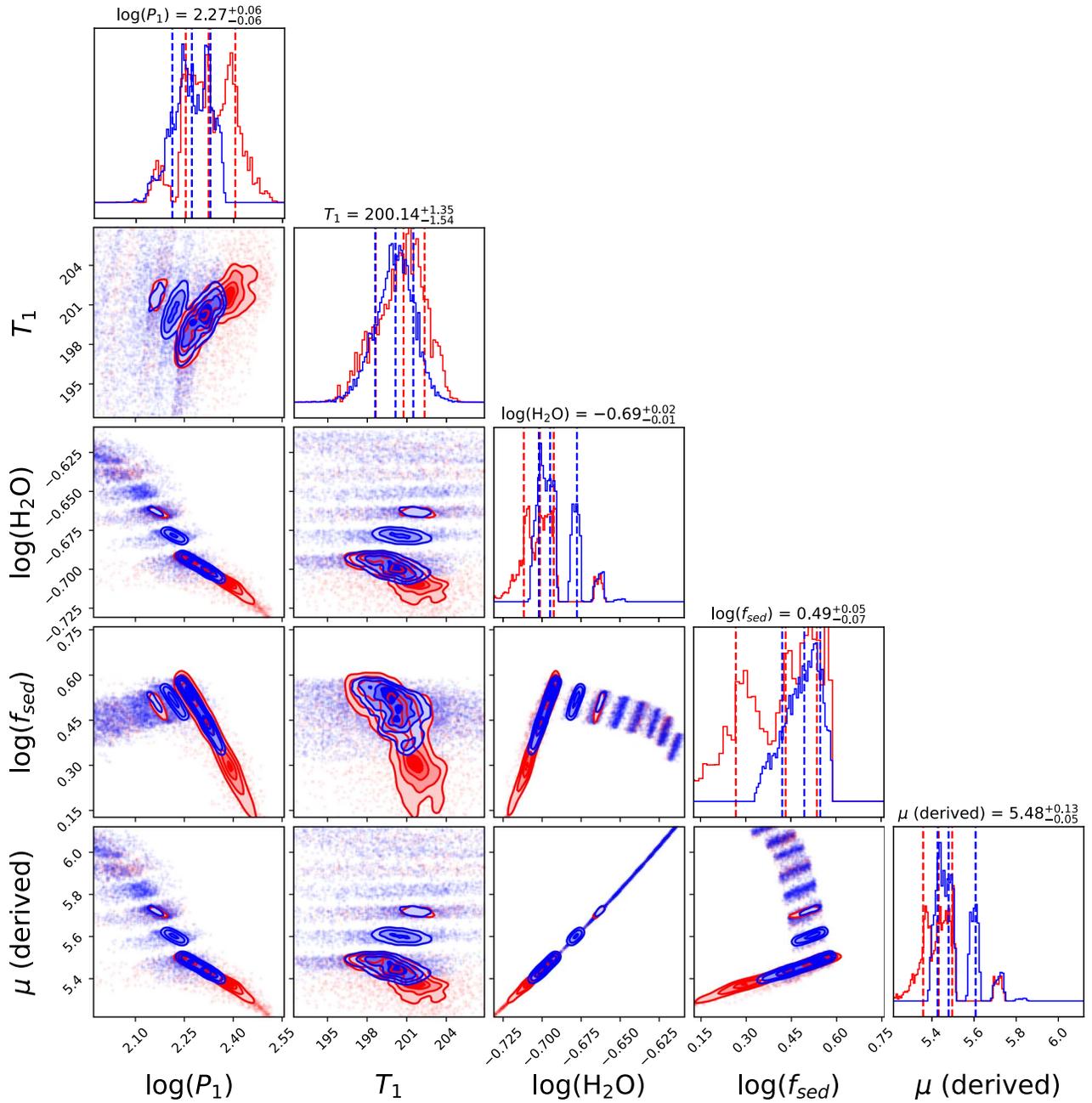


Figure 12. Example of numerical instability in cloud formation simulations when large relative ($rtol$) and absolute ($atol$) tolerances are chosen in solving Equation (2) with Runge–Kutta method. Blue plots: $rtol=1 \times 10^{-8}$, and $atol=1 \times 10^{-12}$. Red plots: $rtol=1 \times 10^{-12}$, and $atol=1 \times 10^{-15}$. In these tests, the retrieval performances are affected by numerical instability, as illustrated in this figure where multiple islands of solutions in the posteriors are clearly visible.

“DOP853” in solving the ODE with $rtol=1 \times 10^{-13}$, and $atol=10^{-16}$, which guarantee numerically stable results.

6. Conclusions

YunMa is a state-of-art cloud simulation and retrieval package optimized for the interpretation of the next generation of exoplanetary atmospheric data, as provided by, e.g., JWST, Roman, Twinkle, and Ariel. These facilities will provide an unprecedented amount of high-quality data, where the cloud formation process and cloud scattering properties can no longer be ignored.

YunMa cloud microphysics is based on the model published by Ackerman & Marley (2001), while the scattering properties

of clouds are calculated through the open-source BH-Mie code. When coupled to the *TauREx* framework (Al-Refaie et al. 2021), *YunMa* becomes a very versatile model that can simulate transit and eclipse spectra for a variety of cloudy exoplanets with different masses, atmospheric compositions, and temperatures. Most importantly, *YunMa+TauREx* can be used as a spectral retrieval framework optimized for cloudy atmospheres.

We have validated *YunMa* against previous work that adopted the A-M approach and compared our results with a 3D model simulation when consistent assumptions are adopted to produce the vertical profile. We have validated the radiative transfer calculations in *YunMa*, including cloud scattering, against PyMieScatt.

We have run over one hundred retrieval experiments with *YunMa*, with different cloud compositions (e.g., KCl, MgSiO₃, Fe clouds). This paper presents and discusses 28 cases of water clouds in the atmosphere of a temperate super-Earth, K2-18 b-like. Through these experiments, we have learnt that *YunMa* is capable of retrieving cloud formation and atmospheric parameters when clouds are not so opaque to mask all the atmospheric features at most wavelengths. More specifically, if we assume spectroscopic data covering the 0.4–14 μm range, uncertainties at a level of 10 ppm, and spectral resolution $R=100$, we can retrieve the sedimentation coefficient, the baseline condensable gases, and the T - p profile points with accuracy levels $>60\%$ in our respective experiments. This is not the case using the simple opaque cloud model, which shows more degeneracy in the posterior distributions of the atmospheric parameters retrieved.

An extension of *YunMa* to interpret phase-curve observations will be a valuable next step. 2D cloud models, which include horizontal convection, will soon be within reach of computing speed and might be considered in future versions of *YunMa*. While we are aware of the limitations of the specific cloud microphysics model embedded in *YunMa*, our results advocate for the need to include more realistic cloud models in spectral retrievals to interpret correctly the results of the next-generation facilities.

Acknowledgments

The work presented in this paper was partially supported by UKSA, grant ST/X002616/1 and ST/W00254X/1, and ExoMolHD ERC grant 883830. Y.I. is supported by JSPS KAKENHI JP22K14090. Q.C. is supported by the ESA Research Fellowship Program. The authors acknowledge the use of the University College London (UCL) Kathleen High-Performance Computing Facility (Kathleen@UCL) and associated support services in completing this work. This work utilized the Cambridge Service for Data-Driven Discovery (CSD3), part of which is operated by the University of Cambridge Research Computing on behalf of the STFC DiRAC HPC Facility (dirac.ac.uk). The DiRAC component of CSD3 was funded by BEIS capital funding via STFC capital grants ST/P002307/1 and ST/R002452/1 and STFC operations grant ST/R00689X/1. DiRAC is part of the National e-Infrastructure.

We thank Dr. Andrew S. Ackerman for the constructive communication on the A-M model and for sharing the original code, Dr. Kai-Hou Yip and Sam Wright for suggestions on the numerical methods, and Dr. Yui Kawashima for the cloud and haze model discussion. We thank the anonymous reviewer for the constructive comments, which improved the paper greatly.

Appendix A

Supplementary Equations to Estimate the Cloud Mixing Profile

In Section 2.1.1, the SVP can be estimated with the Clausius–Clapeyron equation:

$$e_s = e_0 \exp \left[\frac{\ell}{R_{\text{SV}}} \left(\frac{1}{T_0} - \frac{1}{T} \right) \right], \quad (\text{A1})$$

where ℓ is the latent heat of evaporation, R_{SV} is the specific gas constant for the vapor, T is the atmospheric temperature, and T_0 is the temperature at vapor pressure e_0 . We use the results from

laboratory measurements of these parameters for the different chemical species when available.

β is the Cunningham slip factor:

$$\beta = 1 + 1.26N_{\text{Kn}}. \quad (\text{A2})$$

The Knudsen number (N_{Kn}) is the ratio between the molecular mean free path and the droplet radius.

YunMa adopts two ways to estimate η . One is that Lavvas et al. (2008) suggested to use the following:

$$\eta = \frac{1}{3} \rho_a \bar{V} \lambda_a, \quad (\text{A3})$$

where \bar{V} is the thermal velocity of gaseous components, and λ_a is the mean free path. Another uses the definition by Rosner (2012), which is also adopted by Ackerman & Marley (2001):

$$\eta = \frac{5}{16} \frac{\sqrt{\pi m k_B T}}{\pi d^2} \frac{(k_B T / \epsilon)^{0.16}}{1.22}, \quad (\text{A4})$$

where d is the diameter of a gas particle, and ϵ is the atmospheric Lennard-Jones potential well depth. When using the A-M approach, v_f and the particle size are positively correlated using η either from Rosner (2012) or Lavvas et al. (2008).

Appendix B

Derivation of the Cloud Particle Size and Number Density

In Section 2.1.2, assuming a lognormal cloud particle size distribution, the geometric mean (r_g) is defined as follows:

$$r_g = e^{\frac{\int_0^\infty \ln r \frac{dn}{dr} dr}{\int_0^\infty \frac{dn}{dr} dr}}. \quad (\text{B1})$$

The power-law approximation allows representation of f_{sed} using the particle size distribution.

$$f_{\text{sed}} \approx \frac{\int_0^\infty r^{3+\alpha} \frac{dn}{dr} dr}{r_w^\alpha \int_0^\infty r^{\frac{3n}{dr}} dr}, \quad (\text{B2})$$

where n is the accumulated number density as defined in Section 2, and σ_g is the geometric standard deviation of the lognormal particle radius distribution. Through an integration of the lognormal distribution, Ackerman & Marley (2001) derived that

$$r_g = r_w f_{\text{sed}}^{\frac{1}{\alpha}} \exp \left(-\frac{\alpha + 6}{2} \ln^2 \sigma_g \right), \quad (\text{B3})$$

where the σ_g is the geometric standard deviation of the particle radii.

The effective mean radius (r_{eff}) is the area-weighted average radius defined by Hansen & Travis (1974) to approximately represent the scattering properties of the whole size distribution by a single parameter when the particle radius is larger than the radiation wavelength. To derive r_{eff} , we first recall that, in lognormal distribution, the t th raw moment is given by the following:

$$m_t = N_0 r_g \exp \left[\frac{(t\sigma)^2}{2} \right]. \quad (\text{B4})$$

Therefore, r_{eff} , which is the area-weighted average radius, can be estimated through the following:

$$r_{\text{eff}} = \frac{\int_0^\infty r\pi r^2 \frac{dn}{dr} dr}{\int_0^\infty \pi r^2 \frac{dn}{dr} dr} = r_w f_{\text{sed}}^{\frac{1}{3}} \exp\left(-\frac{\alpha + 1}{2} \ln^2 \sigma_g\right). \quad (\text{B5})$$

Similarly, the total number density for the particles is estimated by using the volume-weighted mean:

$$N = \frac{3\varepsilon\rho_a q_c}{4\pi\rho_p r_g^3} \exp\left(-\frac{9}{2} \ln^2 \sigma_g\right). \quad (\text{B6})$$

ORCID iDs

Sushuang Ma  <https://orcid.org/0000-0001-9010-0539>

Yuichi Ito  <https://orcid.org/0000-0002-0598-3021>

Ahmed Faris Al-Refaie  <https://orcid.org/0000-0003-2241-5330>

Quentin Changeat  <https://orcid.org/0000-0001-6516-4493>

Billy Edwards  <https://orcid.org/0000-0002-5494-3237>

Giovanna Tinetti  <https://orcid.org/0000-0001-6058-6654>

References

- Ackerman, A. S., & Marley, M. S. 2001, *ApJ*, 556, 872
- Adams, D. J., Kataria, T., Batalha, N. E., Gao, P., & Knutson, H. A. 2022, *ApJ*, 926, 157
- Al-Refaie, A. F., Changeat, Q., Venot, O., Waldmann, I. P., & Tinetti, G. 2022a, *ApJ*, 932, 123
- Al-Refaie, A. F., Changeat, Q., Waldmann, I. P., & Tinetti, G. 2021, *ApJ*, 917, 37
- Al-Refaie, A. F., Venot, O., Changeat, Q., & Edwards, B. 2022b, arXiv:2209.11203
- Baeyens, R., Decin, L., Carone, L., et al. 2021, *MNRAS*, 505, 5603
- Barstow, J. K. 2020, *MNRAS*, 497, 4183
- Batalha, N. E., Marley, M. S., Lewis, N. K., & Fortney, J. J. 2019, *ApJ*, 878, 70
- Baudino, J. L., Bézard, B., Boccaletti, A., et al. 2015, *A&A*, 582, A83
- Bean, J. L., Stevenson, K. B., Batalha, N. M., et al. 2018, *PASP*, 130, 114402
- Benneke, B. 2015, arXiv:1504.07655
- Benneke, B., Knutson, H. A., Lothringer, J., et al. 2019, *NatAs*, 3, 813
- Bohren, C. F., & Huffman, D. R. 2008, *Absorption and Scattering of Light by Small Particles* (New York: Wiley)
- Boucher, A., Darveau-Bernier, A., Pelletier, S., et al. 2021, *AJ*, 162, 233
- Broggi, M., & Line, M. R. 2019, *AJ*, 157, 114
- Burrows, A. S. 2014, *Natur*, 513, 345
- Caldas, A., Leconte, J., Selsis, F., et al. 2019, *A&A*, 623, A161
- Carlson, B. E., Rossow, W. B., & Orton, G. S. 1988, *JatS*, 45, 2066
- Changeat, Q., Al-Refaie, A. F., Edwards, B., Waldmann, I. P., & Tinetti, G. 2021, *ApJ*, 913, 73
- Changeat, Q., Edwards, B., Al-Refaie, A. F., et al. 2022a, *ApJS*, 260, 3
- Changeat, Q., Edwards, B., Al-Refaie, A. F., et al. 2022b, *ExA*, 53, 391
- Charnay, B., Bézard, B., Baudino, J. L., et al. 2018, *ApJ*, 854, 172
- Charnay, B., Blain, D., Bézard, B., et al. 2021b, *A&A*, 646, A171
- Charnay, B., Tobie, G., Lebonnois, S., & Lorenz, R. D. 2022, *A&A*, 658, A108
- Cho, J. Y.-K., Skinner, J. W., & Thrastarson, H. T. 2021, *ApJL*, 913, L32
- Chubb, K. L., Rocchetto, M., Yurchenko, S. N., et al. 2021, *A&A*, 646, A21
- Cox, A. N. 2015, *Allen's Astrophysical Quantities* (4th ed.; New York: Springer)
- Cubillos, P. E., & Blecic, J. 2021, *MNRAS*, 505, 2675
- Drummond, B., Tremblin, P., Baraffe, I., et al. 2016, *A&A*, 594, A69
- Edwards, B., Changeat, Q., Tsiaras, A., et al. 2022, arXiv:2211.00649
- Edwards, B., Rice, M., Zingales, T., et al. 2019, *ExA*, 47, 29
- Fleck, J. A. J., & Canfield, E. H. 1984, *JCoPh*, 54, 508
- Forget, F., Hourdin, F., Fournier, R., et al. 1999, *JGR*, 104, 24155
- Fortney, J. J., Barstow, J. K., & Madhusudhan, N. 2021, in *ExoFrontiers: Big Questions in Exoplanetary Science*, ed. N. Madhusudhan (Bristol: IOP Publishing), 17
- Gandhi, S., & Madhusudhan, N. 2018, *MNRAS*, 474, 271
- Gao, P., Marley, M. S., & Ackerman, A. S. 2018, *ApJ*, 855, 86
- Gao, P., Thongren, D. P., Lee, E. K. H., et al. 2020, *NatAs*, 4, 951
- Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, *SSRv*, 123, 485
- Gierasch, P., & Conrath, B. 1985, in *Recent Advances in Planetary Meteorology*, ed. G. E. Hunt (Cambridge: Cambridge Univ. Press), 121
- Goyal, J. M., Mayne, N., Sing, D. K., et al. 2018, *MNRAS*, 474, 5158
- Greene, T. P., Line, M. R., Montero, C., et al. 2016, *ApJ*, 817, 17
- Hairer, E., Nørsett, S. P., & Wanner, G. 1993, *Solving Ordinary Differential Equations. I: Nonstiff Problems* (Berlin: Springer).
- Hansen, J. E., & Travis, L. D. 1974, *SSRv*, 16, 527
- Harrington, J., Himes, M. D., Cubillos, P. E., et al. 2022, *PSJ*, 3, 80
- Helling, C. 2023a, in *Planetary Systems Now*, ed. L. M. Lara & D. Jewitt (Singapore: World Scientific), 235
- Helling, C., Iro, N., Corrales, L., et al. 2019, *A&A*, 631, A79
- Helling, C., Samra, D., Lewis, D., et al. 2023b, *A&A*, 671, A122
- Hourdin, F., Musat, I., Bony, S., et al. 2006, *CIDJ*, 27, 787
- Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, *A&A*, 553, A6
- Irwin, P., Teanby, N., De Kok, R., et al. 2008, *JQSRT*, 109, 1136
- JWST Transiting Exoplanet Community Early Release Science Team 2023, *Natur*, 614, 649
- Karman, T., Gordon, I. E., van der Avoird, A., et al. 2019, *Icar*, 328, 160
- Kawashima, Y., & Ikoma, M. 2018, *ApJ*, 853, 7
- Kitzmann, D., Heng, K., Oreshenko, M., et al. 2020, *ApJ*, 890, 174
- Komacek, T. D., Showman, A. P., & Parmentier, V. 2019, *ApJ*, 881, 152
- Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, *Natur*, 505, 69
- Lavvas, P., Coustenis, A., & Vardavas, I. 2008, *P&SS*, 56, 67
- Lee, J. M., Fletcher, L. N., & Irwin, P. G. J. 2012, *MNRAS*, 420, 170
- Lewis, J. S. 1969, *Icar*, 10, 365
- Line, M. R., Wolf, A. S., Zhang, X., et al. 2013, *ApJ*, 775, 137
- Lueber, A., Kitzmann, D., Bowler, B. P., Burgasser, A. J., & Heng, K. 2022, *ApJ*, 930, 136
- Lunine, J. I., Hubbard, W., Burrows, A., Wang, Y.-P., & Garlow, K. 1989, *ApJ*, 338, 314
- MacDonald, R. J., & Madhusudhan, N. 2017, *MNRAS*, 469, 1979
- Madhusudhan, N. 2019, *ARA&A*, 57, 617
- Madhusudhan, N., Nixon, M. C., Welbanks, L., Piette, A. A., & Booth, R. A. 2020, *ApJL*, 891, L7
- Madhusudhan, N., & Seager, S. 2009, *ApJ*, 707, 24
- Mai, C., & Line, M. R. 2019, *ApJ*, 883, 144
- Marley, M. S., Gelino, C., Stephens, D., Lunine, J. I., & Freedman, R. 1999, *ApJ*, 513, 879
- Min, M., Hovenier, J. W., & de Koter, A. 2005, *A&A*, 432, 909
- Min, M., Ormel, C. W., Chubb, K., Helling, C., & Kawashima, Y. 2020, *A&A*, 642, A28
- Mollière, P., Stolker, T., Lacour, S., et al. 2020, *A&A*, 640, A131
- Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, *A&A*, 627, A67
- Nixon, M. C., & Madhusudhan, N. 2022, *ApJ*, 935, 73
- Ormel, C. W., & Min, M. 2019, *A&A*, 622, A121
- Parmentier, V., Showman, A. P., & Lian, Y. 2013, *A&A*, 558, A91
- Pinhas, A., Madhusudhan, N., Gandhi, S., & MacDonald, R. 2019, *MNRAS*, 482, 1485
- Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018, *MNRAS*, 480, 2597
- Robbins-Blanch, N., Kataria, T., Batalha, N. E., & Adams, D. J. 2022, *ApJ*, 930, 93
- Rooney, C. M., Batalha, N. E., Gao, P., & Marley, M. S. 2022, *ApJ*, 925, 33
- Rosner, D. E. 2012, *Transport Processes in Chemically Reacting Flow Systems* (New York: Dover)
- Roudier, G. M., Swain, M. R., Gudipati, M. S., et al. 2021, *AJ*, 162, 37
- Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, *Natur*, 529, 59
- Stevenson, K. B. 2016, *ApJL*, 817, L16
- Sumlin, B. J., Heinson, W. R., & Chakrabarty, R. K. 2018, *JQSRT*, 205, 127
- Tennyson, J., & Yurchenko, S. N. 2012, *MNRAS*, 425, 21
- Tennyson, J., & Yurchenko, S. N. 2021, *A&G*, 62, 6.16
- Tinetti, G., Drossart, P., Eccleston, P., et al. 2018, *ExA*, 46, 135
- Tinetti, G., Eccleston, P., Haswell, C., et al. 2021, *Ariel: Enabling planetary science across light-years* SCI(2020)1, ESA https://www.cosmos.esa.int/documents/1783156/3267291/Ariel_RedBook_Nov2020.pdf
- Tinetti, G., Encrenaz, T., & Coustenis, A. 2013, *A&ARv*, 21, 63
- Toon, O. B., Turco, R. P., Hamill, P., Kiang, C. S., & Whitten, R. C. 1979, *JatS*, 36, 718
- Tremblin, P., Amundsen, D. S., Mourier, P., et al. 2015, *ApJL*, 804, L17
- Tsai, S.-M., Innes, H., Lichtenberg, T., et al. 2021, *ApJL*, 922, L27
- Tsiaras, A., Waldmann, I. P., Tinetti, G., Tennyson, J., & Yurchenko, S. N. 2019, *NatAs*, 3, 1086
- Tsiaras, A., Waldmann, I. P., Zingales, T., et al. 2018, *AJ*, 155, 156
- Turco, R. P., Hamill, P., Toon, O. B., Whitten, R. C., & Kiang, C. S. 1979, *JatS*, 36, 699

- Venot, O., Parmentier, V., Blecic, J., et al. 2020, [ApJ](#), 890, 176
- Wakeford, H. R., Sing, D. K., Deming, D., et al. 2018, [AJ](#), 155, 29
- Wang, F., Fujii, Y., & He, J. 2022, [ApJ](#), 931, 48
- Warren, S. G., & Brandt, R. E. 2008, [JGRD](#), 113, D14220
- Welbanks, L., & Madhusudhan, N. 2021, [ApJ](#), 913, 114
- Welbanks, L., Madhusudhan, N., Allard, N. F., et al. 2019, [ApJL](#), 887, L20
- Windsor, J. D., Robinson, T. D., Kopparapu, R. K., et al. 2023, [PSJ](#), 4, 94
- Xuan, J. W., Wang, J., Ruffio, J.-B., et al. 2022, [ApJ](#), 937, 54
- Yu, X., Moses, J. I., Fortney, J. J., & Zhang, X. 2021, [ApJ](#), 914, 38
- Zhang, M., Chachan, Y., Kempton, E. M. R., Knutson, H. A., & Chang, W. H. 2020, [ApJ](#), 899, 27
- Zhang, X., & Showman, A. P. 2018, [ApJ](#), 866, 1