Astronomy & Astrophysics

ARES

VI. Viability of one-dimensional retrieval models for transmission spectroscopy characterization of exo-atmospheres in the era of JWST and Ariel

Adam Yassin Jaziri[1,2,10], William Pluriel[2], Andrea Bocchieri[3], Emilie Panek[4], Lucas Teinturier[5,6], Anastasiia Ivanova[1], Natalia E. Rektsini[4,7], Pierre Drossart[4,5], Jean-Philippe Beaulieu[4,7], Aurélien Falco[8,9], Jérémy Leconte[10], Lorenzo V. Mugnai[11,12,3,13], and Olivia Venot[14]

[1] LATMOS/IPSL, UVSQ Université Paris-Saclay, Sorbonne Université, CNRS, Guyancourt, France
  e-mail: yassin.jaziri@latmos.ipsl.fr
[2] Observatoire Astronomique de l'Université de Genève, département d'astronomie Chemin Pegasi 51, 1290 Versoix, Switzerland
[3] Dipartimento di Fisica, La Sapienza Università di Roma, Piazzale Aldo Moro 2, 00185 Roma, Italy
[4] Institut d'Astrophysique de Paris – CNRS, Sorbonne Université, Paris, France
[5] LESIA, Observatoire de Paris, Université PSL, Sorbonne Université, Université Paris Cité, CNRS, 5 place Jules Janssen, 92195 Meudon, France
[6] Laboratoire de Météorologie Dynamique, IPSL, CNRS, Sorbonne Université, Ecole Normale Supérieure, Université PSL, Ecole Polytechnique, Institut Polytechnique de Paris, 75005 Paris, France
[7] School of Natural Sciences, University of Tasmania, Private Bag 37 Hobart, Tasmania 70001, Australia
[8] Université de Paris Cité, Institut de Physique du Globe de Paris, CNRS 75005 Paris, 1 rue Jussieu, 75005 Paris, France
[9] Laboratoire AIM, CEA, CNRS, Univ. Paris-Sud, UVSQ, Université Paris-Saclay, 91191 Gif-sur-Yvette, France
[10] Laboratoire d'astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France
[11] School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, Cardiff CF24 3AA, UK
[12] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[13] INAF – Osservatorio Astronomico di Palermo, Piazza del Parlamento 1, 90134 Palermo, Italy
[14] Université Paris Cité and Univ. Paris Est Créteil, CNRS, LISA, 75013 Paris, France

ABSTRACT

Observed exoplanet transit spectra are usually retrieved using 1D models to determine atmospheric composition. However, planetary atmospheres are 3D. With the new state-of-the-art *James Webb* Space Telescope (JWST) and future space telescopes such as Ariel (Atmospheric Remote-sensing Infrared Exoplanet Large-survey), we will be able to obtain increasingly accurate transit spectra. The 3D effects on the spectra will be visible, and we can expect biases in the 1D extractions. In order to elucidate these biases, we have built theoretical observations of transit spectra, from 3D atmospheric modeling through transit modeling to instrument modeling. For this purpose, we used a global climate model (GCM) to simulate the atmosphere, a 3D radiative transfer model to calculate theoretical transmission spectra, and adapted instrument software from JWST and Ariel to reproduce telescope noise. Next, we used a 1D radiative transfer inversion model to retrieve the known input atmosphere and disentangle any biases that might be observed. The study was done from warm planets to ultra-hot planets to assess biases as a function of average planet temperature. Three-dimensional effects are observed to be strongly nonlinear from the coldest to the hottest planets. These effects also depend on the planet's metallicity and gravity. Considering equilibrium chemistry, 3D effects are observed through very strong variations in certain features of the molecule or very small variations over the whole spectrum. We conclude that we cannot rely on the uncertainty of retrievals at all pressures, and that we must be cautious about the results of retrievals at the top of the atmosphere. However the results are still fairly close to the truth at mid-altitudes (those probed). We also need to be careful with the chemical models used for planetary atmosphere. If the chemistry of one molecule is not correctly described, this will bias all the others, and the retrieved temperature as well. Finally, although fitting a wider wavelength range and higher resolution has been shown to increase retrieval accuracy, we show that this could depend on the wavelength range chosen, due to the accuracy on modeling the different features. In any case, 1D retrievals are still correct for the detection of molecules, even in the event of an erroneous abundance retrieval.

Key words. planets and satellites: atmospheres – planets and satellites: composition – planets and satellites: detection

1. Introduction

Over the past three decades, there has been an exponential rise in the discovery of exoplanets that have dramatically expanded our understanding of exoplanets. Their observed diversity (e.g., Gaudi et al. 2021) has fueled the excitement and imagination of the scientific community and the general public and challenged long-held assumptions about planet formation and evolution.

The focus in the exoplanet field has shifted from studying the bulk and orbital parameters to understanding the true nature of exoplanets through their compositions, atmospheres, and climates (Gaudi et al. 2021; Guillot et al. 2022). The atmospheres are shaped by the stellar environment and offer a glimpse into the planetary interior, which holds evidence of the formation process of the planet. Transiting exoplanets are very well suited for atmospheric analysis as the stellar light filtering through their

atmosphere provides a wealth of information on its composition and thermodynamic state (Sing 2018; Tsiaras et al. 2018). Transmission spectroscopy has therefore emerged as the most promising technique for atmospheric characterization (Seager & Sasselov 2000; Tinetti et al. 2007; Kreidberg et al. 2014; Line et al. 2016; Welbanks et al. 2019; Madhusudhan 2019; Edwards et al. 2020; Skaf et al. 2020; Pluriel et al. 2020a; Guilluy et al. 2020; Mugnai et al. 2021a). Until very recently the main challenges in atmospheric studies using this technique were limited sensitivity and spectral range that prevented proper sampling of molecular features. In this regard, the launch of the *James Webb* Space Telescope (JWST) marks a new milestone in the exploration of exoplanets and provides us with a unique opportunity to uncover their true nature and formation-evolution histories. The upcoming space mission Atmospheric Remote-Sensing Infrared Exoplanet Large-survey (Ariel) of the European Space Agency (ESA) will provide yet another important contribution by systematically characterizing the atmospheres of entire populations of exoplanets (e.g., Charnay & Drossart 2023).

The JWST (McElwain et al. 2023) is a large highly sensitive infrared-optimized space telescope developed in a collaboration between NASA, the ESA, and the Canadian Space Agency (CSA). The JWST science goals were organized into four themes, including Planetary Systems and the Origins of Life, which aims to determine the physical and chemical properties of planetary systems and investigate the potential for the origins of life in those systems (Gardner et al. 2006; Beichman & Greene 2018). The overall design of the observatory includes three main components: the telescope and scientific instruments, the five-layer sunshield, and the spacecraft bus. The telescope has a unique 6.5 m class primary mirror comprising 18 hexagonal segments, phased together to act as a single mirror. There are four science instruments, each with several observing modes: the Near Infrared Camera (NIRCam, 0.6–5.0 µm; e.g., Beichman et al. 2012), the Near Infrared Spectrograph (NIRSpec, 0.6–5.3 µm; Jakobsen et al. 2022), the Near Infrared Imager and Slitless Spectrograph (NIRISS, 0.6–5.0 µm; Doyon et al. 2012), and the Mid-Infrared Instrument (MIRI, 5–28.5 µm; Rieke et al. 2015). The telescope and science instruments are passively cooled to below 50 K by the sunshield and thermally isolated from the spacecraft bus and solar array. The combined spectral coverage of all science instruments is from 0.6 to 28.5 µm, opening uncharted territory in atmospheric characterization, as evidenced by recent inferences of CO_2 and SO_2 in the atmosphere of the hot Jupiter WASP-39b (Bean et al. 2018; Rustamkulov et al. 2023; Feinstein et al. 2023; Ahrer et al. 2023; Tsai et al. 2023; Alderson et al. 2023).

Ariel (Tinetti et al. 2018, 2021) is the M4 ESA mission of the Cosmic Vision program, and will operate from the L2 point starting in 2029. Ariel is a dedicated mission for the spectroscopic observation of transiting exoplanets; it will conduct the first unbiased survey of a large and diverse sample of approximately 1000 planets in the optical and near-infrared (e.g., Mugnai et al. 2021b). The Ariel payload will mount an off-axis Cassegrain telescope with a 1m class primary mirror feeding a collimated beam, split by dichroics, into two separate instruments with a coincident field of view: the Fine Guidance System (FGS) and the Ariel InfraRed Spectrometer (AIRS). FGS includes three photometric channels (VISPhot, 0.5–0.6 µm; FGS1, 0.6–0.80 µm; FGS2, 0.80–1.1 µm) and NIRSpec (1.1–1.95 µm and $R \geq 15$); AIRS has two low- to medium-resolution IR channels (CH0, 1.95–3.9 µm and $R \geq 100$; CH1, 3.9–7.8 µm and $R \geq 30$). The payload module comprising the telescope and the science instruments will be passively cooled to ~55 K

with the thermal shield assembly. With this instrumentation, Ariel will provide simultaneous observations of the whole 0.5–7.8 µm spectral band, encompassing the emission peak of warm and hot exoplanets and sufficient to detect all molecular species (Encrenaz et al. 2015). The key science questions to be addressed by the Ariel mission are the following: What are the physical processes shaping planetary atmospheres?, What are exoplanets made of?, and How do planets and planetary systems form and evolve? (Turrini et al. 2018; Changeat et al. 2020a). JWST and Ariel are complementary. JWST, with its exquisite sensitivity, is already providing transformative science for the field of exoplanets and, thanks to its precision launch, will continue to do so for the next 20+ yr (McElwain et al. 2023), with about 25% time allocated to exoplanet observations. Ariel, on the other hand, during the four-year nominal mission will already observe a conspicuous fraction of the known planetary population, providing the statistics to interpret the full gamut of observed planetary bodies and place the JWST observations into a wider context.

The exquisite sensitivity and wavelength range offered by JWST, and to a lesser extent by Ariel, enable the study of subtler effects than previously possible, opening new pathways to understanding the true nature of exoplanets. A prime example is the characterization of the 3D nature of exoplanets. In this regard, the geometry of a transit allows us to probe only a limited volume around the terminator line: the region of the atmosphere at the border between the hot day-side and the cooler night-side (Brown 2001; Kreidberg 2018). Therefore, transit spectroscopy measurements average the atmospheric features from the morning and evening sides of the planet (Fortney et al. 2010; Caldas et al. 2019; Wardenier et al. 2022). Often, a uniform terminator is assumed when interpreting transit spectroscopy data, disregarding 3D atmospheric effects. This assumption can be sufficient to interpret the data in the case of cold planets that have more homogeneous atmospheres (MacDonald et al. 2020), but not for hot Jupiters and ultra-hot Jupiters. Hot and ultra-hot Jupiters present strong day-night contrast in temperature due to the shorter radiative timescale compared to the dynamical timescale (Sudarsky et al. 2000; Guillot 2010; Bell & Cowan 2018; Arcangeli et al. 2018), and also chemical heterogeneities (Changeat et al. 2019; Baeyens et al. 2021). These differences diminish as we move toward warm Jupiters (equilibrium temperature below 1400 K), where zonal circulation tends to homogenize atmospheres (Venot et al. 2020; Baeyens et al. 2022). Global climate models (GCMs) are numerical models of the full 3D structures and dynamics of atmospheres that can help us explore and gain a better understanding of atmospheric processes and characteristics that cannot be captured by a 1D plane-parallel approach (Showman et al. 2008; Leconte et al. 2013; Venot et al. 2014; Parmentier et al. 2018).

This work addresses questions on the interpretability of transmission spectra given the 3D nature of exoplanetary atmospheres. The first question is how the 3D atmospheric structure affects the transmission spectra of exoplanets, from a cold planet to an ultra-hot Jupiter. Recent works (Pluriel et al. 2020b; Lacy & Burrows 2020; Wardenier et al. 2022; Pluriel et al. 2022) concluded that, for ultra-hot Jupiters, the 3D structure plays a major role in shaping the transmission spectra. If the temperature of the atmosphere is not high enough to dissociate a molecule on both the day-side and the night-side, the amplitude of its spectral features will be larger than predicted by a 1D plane-parallel approach. In addition, Falco et al. (2022) showed how the changes in planet orientation during the transit allow us to probe the horizontal variations in the atmosphere. Pluriel et al. (2022) investigated 3D effects in the transmission spectra of hot

Jupiters, dividing them into three main groups: vertical effects, horizontal effects along the limb, and horizontal effects through the limb.

The second question we ask is whether 1D retrievals can find consistent parameters (T-P profile, abundances, C/O ratio, metallicity, and clouds). This question is closely related to the first one: whether the 3D atmospheric structure strongly affects transmission spectra, and whether the 1D retrieval models can find correct atmospheric parameters. In this regard, MacDonald et al. (2020) investigated the conundrum of why most inferred temperatures from transmission spectra are far colder than expected with respect to the equilibrium temperature. They concluded that a 1D model can fit the transmission spectra of planets with asymmetric terminators, but retrieved atmospheric parameters may not represent true terminator averages. The retrieved temperatures of planetary terminators may also be biased by hundred degrees below their real value, and this bias is most extreme in the case of ultra-hot Jupiters (but see Welbanks & Madhusudhan 2022). This study also mentions the biases in chemical abundances derived from 1D retrievals. Pluriel et al. (2020b) concluded the same: if the temperature and the chemical composition vary across the limb, which is the case for 3D structures, 1D retrievals cannot find the correct molecular abundances. This also affects the inferred C/O ratio, which is an indirect estimate based on the abundances of all C- and O-bearing molecules. Recently, Zingales et al. (2022) and Welbanks & Madhusudhan (2022) demonstrated that the choice of the retrieval model is critical for correctly retrieving the thermal structure of the atmosphere. Lastly, Pluriel et al. (2022) provided a "cheat sheet" of the minimum model assumptions needed to avoid biases in interpreting atmospheric properties. When optical absorbers are present, 1D models are adequate to describe transmission spectra for atmospheric equilibrium temperatures lower than 1400 K; in their absence, the 1D assumption can extend up to 2000 K. Above these temperatures, retrievals with 1D models return biased estimates of the parameters in the forward model.

The problem of terminator inhomogeneities addressed here can be solved by more sophisticated retrieval using 2D atmospheric models (MacDonald et al. 2020). As these inhomogeneities would be observable by JWST (Espinoza & Jones 2021), this is part of the current needs. Zingales et al. (2022) and Welbanks & Madhusudhan (2022) have studied hot and ultra-hot Jupiter to show that models beyond the 1D approach such as 2D models can retrieve consistent temperature profiles. Similarly, MacDonald & Lewis (2022) and Nixon & Madhusudhan (2022) have implemented a 3D parametric pressure-temperature profile to model day-night temperature variations with improved retrieval accuracy. This study is part of this new challenge in exoplanet research, and will specifically address the issue of errors introduced by 1D retrieval models when applied to contrasting 3D atmospheres.

To address the above-mentioned issues, we need two distinct branches of action: one that can produce simulated transmission spectra for different planets, accounting for 3D structures, and one that can perform atmospheric retrievals and compare the inferred parameters with the forward models. Our goal is to assess the extent to which our retrievals can reconstruct the true atmospheric composition. This end-to-end system allows us to investigate the interpretation of our atmospheric retrievals consistently. The first branch starts from precomputed GCMs for three planets (see Sect. 2.1):

– GJ 1214 b, a Neptune-like planet orbiting an M-type star;
– HD 189733 b, a hot Jupiter around a K-type star;
– WASP-121 b, an ultra-hot Jupiter orbiting an F-type star.

Table 1. Planet–star parameters used to simulate the transmission spectra and as input into PandExo.

Parameters	GJ 1214 b	HD 189733 b	WASP-121 b
T_p (K)	596	1209	2358
R_* (Rsun)	0.21	0.75	1.46
T_* (K)	3250	5052	6459
Z_*	0.29	–0.02	0.13
$\log(g)$	5.03	4.49	4.24
R_p (Rjup)	0.243	1.13	1.91
References	Charbonneau et al. (2009); Cloutier et al. (2021)	Stassun et al. (2017); Addison et al. (2019)	Delrez et al. (2016)

Then it uses the Pytmosph3R code to produce 3D transmission spectra (see Sect. 2.2). In this step, we use two different configurations for each planet: an equilibrium chemistry model and a model with chemical profiles constant with altitude. The final products of the first branch are spectra with attached error bars, to reproduce spectra as observed by JWST and Ariel. The expected error bars are estimated using PandExo and ArielRad, two simulators of the noise performance of JWST and Ariel, respectively (see Sect. 2.3). The second branch starts from these simulated observations and performs a Bayesian retrieval to estimate the best-fitting parameters of the model. As a retrieval tool, we use the retrieval framework TauREx 3 (Tau Retrieval for Exoplanets), briefly described in Sect. 3.1. Section 3.3 details the retrieval procedure, the chemical configurations, and the other atmospheric parameters. The retrieval procedure is the same for all spectra: the same set of atmospheric models is applied to remove our a priori knowledge and compare the obtained results correctly. We discuss the results from this comparison and their implications for interpreting transmission spectra with the 1D assumption in Sect. 4. Our conclusions are summarized in Sect. 5.

2. Transmission spectra simulations

The following three planets (GJ 1214 b, HD 189733 b, WASP-121 b) have been chosen from warm Neptune to ultra-hot Jupiter (see Table 1) in order to study the transmission spectra and retrieval biases depending on the temperature of the planets. Following Al-Refaie et al. (2022a) study, we go a step further by focusing on biases arising from 1D vertical thermal variation as well as full 3D thermal structure. We simulate JWST (NIRSpec + MIRI) and Ariel observations. We do not consider clouds to focus on general planetary properties biases.

2.1. Global climate models

GJ 1214 b has been simulated using the generic Planetary Climate Model. This model has been specifically developed for exoplanets and paleoclimate studies (Charnay et al. 2015; Leconte et al. 2013). The dynamical core solves the primitive hydrostatic equations of meteorology on an Arakawa C grid, using a finite difference scheme. Radiative transfer is solved using the correlated-k model. Radiative effects of H_2, He, H_2O, CH_4, NH_3, CO, and CO_2 are taken into account, assuming a 100× solar metallicity. The horizontal resolution is 64 × 48 and we use 45 vertical layers between 80 bar and 3 Pa, equally spaced in log pressure. The star is taken as a blackbody at 3026 K, and we assume an internal temperature of 60 K. The

dynamical time-step is 60 s and the physical/radiative time-step is 300 s. The model was integrated for 1600 days. For a more complete description of the model and the simulation, we refer the reader to Charnay et al. (2015), as the model used here is from this study.

For HD 189733 b, we make use of the Met Office Unified Model (Drummond et al. 2018). The model solves the deep atmosphere, nonhydrostatic Navier–Stokes equations on an Arakawa C grid. Radiative transfer is handled through the SOCRATES[1] code adapted for hot Jupiters (Amundsen et al. 2016). A chemical relaxation scheme is used and the radiative transfer is computed in 32 wavelength bins. The chemistry includes CH_4, CO and H_2O. The simulation is integrated for 1000 Earth days. For a more complete description of the model, see Drummond et al. (2018) as the model used here is the same as in this study.

WASP-121 b has been simulated using the SPARC/MIT global circulation model (Showman et al. 2009). The model solves the same primitive equations of meteorology as the generic Planetary Climate Model on a cubic-sphere grid. It has been widely used for various hot Jupiters (Showman et al. 2009; Kataria et al. 2015; Parmentier et al. 2016, 2018, 2021) and has also been applied to ultra-hot Jupiters (Kreidberg 2018; Arcangeli et al. 2019). For this study, we use the model published in Parmentier et al. (2018). The horizontal resolution is $C32$, equivalent to 128 cells in longitude and 64 in latitude and 53 vertical levels with pressure ranging from 200 bar to 2 μbar. Radiative transfer is handled using the two-stream approximation with 11 wavelength bins, as done in Kataria et al. (2013). The model assumes chemical equilibrium, taking into account the thermal dissociation of water and hydrogen. The chemical species taken into account are H_2O, H^-, CO, TiO, Na, Vo, K and H_2. However, H_2 recombination is neglected despite its nonnegligible impact on the thermal and dynamical structure (Tan & Komacek 2019). For a more complete description of the model and the simulation, we refer the reader to Parmentier et al. (2018).

In the three models considered here, disequilibrium chemistry is not taken into account and the models are cloud-free and haze-free.

For each planet, we build a pseudo-1D version of the 3D model. The temperature profiles of the whole grid are replaced by the same 1D profile. This profile is the temperature profile at the equator of the western terminator (coldest one) for each model. Thus, the pseudo-1D models are representative of each temperature condition from warm Neptune to ultra-hot Jupiter. The purpose of these pseudo-1D models is to control the correct behavior of the retrieval code. If we consider the 1D retrieval code, it should correctly retrieve the pseudo-1D models. Thus, we can confidently untangle the 3D effects.

2.2. Pytmosph3R

Based on the planetary atmospheres described in Sect. 2.1, we used the latest version of Pytmosph3R (Falco et al. 2022) to generate the transmission spectra. It takes into account the 3D structure of the atmosphere and uses directly the line-by-line cross sections calculated by ExoMol (Yurchenko et al. 2011, 2014; Tennyson & Yurchenko 2012; Barton et al. 2013, 2014), more specifically cross sections are taken from Chubb et al. (2021). Species abundances were established in two different ways: (i) a constant chemistry model, that is where abundances are independent of temperature and pressure, thus constant everywhere, (ii) an equilibrium chemical model.

The abundances of the constant chemistry model have been chosen close to the equilibrium ratio for a given temperature representative of each planet, and set up over all longitudes, latitudes and altitudes. We have taken into account only the main species: H_2O, CO, CH_4, CO_2, HCN, C_2H_2, NH_3, C_2H_4, and in addition TiO, VO, K, Na, SiO, FeH for HD 189733 b and WASP-121 b. The values are given in Tables B.1–B.3, input column in red. This theoretical construction, not representative of the real atmosphere, removes one degree of freedom (chemistry) to check how 1D retrieval models handle 3D atmospheric thermal structures and thus what can be the biases.

Equilibrium chemistry leads to variable chemical profiles. In the modeled atmospheres, equilibrium chemistry can be expected in the hottest and densest regions. However, we know that non-equilibrium chemistry must be accounted for, especially in the upper atmosphere (Cooper & Showman 2006; Moses et al. 2011, 2012; Venot et al. 2012, 2020; Molaverdikhani et al. 2019; Tsai et al. 2021, 2022). This has been very recently implemented in the retrieval code (FRECKLL code developed by Al-Refaie et al. 2022b) but this is computationally expensive. We used ACE chemistry (see Sect. 3.2 for details) to model the input equilibrium chemistry of GJ 1214 b and the chemical abundances used in Parmentier et al. (2016) to model the input equilibrium chemistry of HD 189733 b and WASP-121 b. This theoretical construct, relative to the constant case, focuses on the research biases that can arise from the chemical retrieval models.

2.3. Uncertainty model

2.3.1. PandExo – JWST

To simulate JWST observables (spectra Fig. 1), we utilize the PandExo 1.5 package (Batalha et al. 2017). This program is a noise simulator designed for JWST transiting observations of exoplanets. We make use of the model to simulate one transit of each planet using NIRSpec-PRISM and one transit using MIRI-LRS. For each planet, we consider a saturation limit of 80% of the full well and a fraction of time out-of-transit to in-transit of 1. For each instrument and planet, we use the optimize option for the number of groups per integration, which automatically defines the best settings to carry the observations. The planet and star-specific parameters used for producing the observables are summarized in Table 1.

PandExo neither includes the varying stellar noise nor takes into account the transit ingress and egress. However, we estimate that the produced observables are a good enough approximation to what one will observe with the JWST facility and should not change the conclusions of this work.

It is worth noting that the three systems studied here have a J-band stellar magnitude below 11.4, which is the estimated saturation limit of NIRSpec-PRISM. Thus, these systems are not observable with this instrument configuration. Yet, we still performed the computations and studied these systems, as the goal of this analysis is not to prepare JWST observations per se, but to highlight possible biases introduced by retrievals on JWST-like data sets. We did not add random noise, as explain in the following section.

2.3.2. ArielRad – Ariel

The Ariel Radiometric model. ArielRad (Mugnai et al. 2020) is the radiometric simulator of the Ariel payload, developed and maintained by the Ariel Consortium. We briefly describe ArielRad here; for a technical description, including
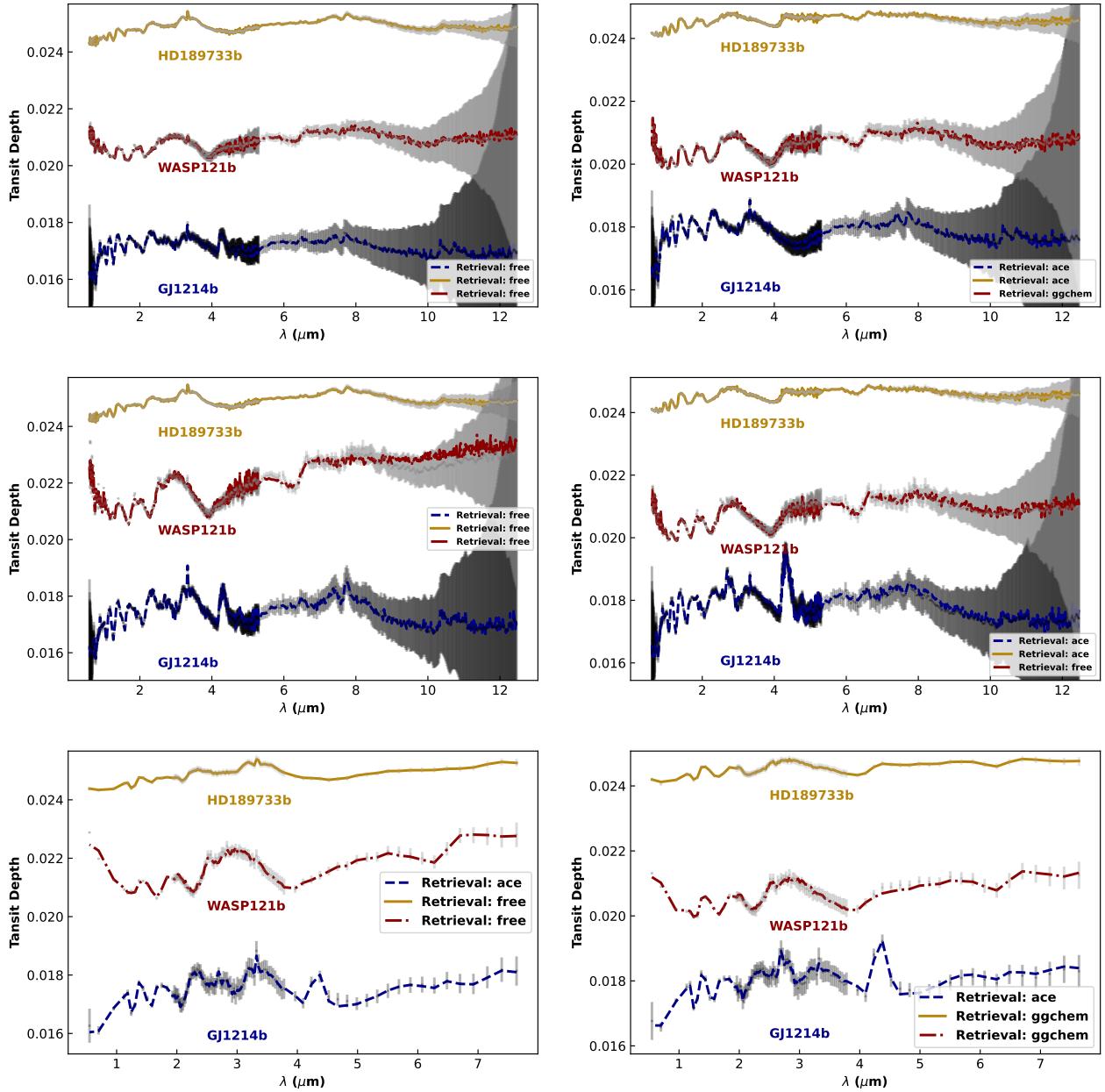
---

[1] https://code.metoffice.gov.uk/trac/socrates

Fig. 1. Simulation of observations for HD 189733 b, WASP-121 b, and GJ 1214 b (gray) with the corresponding best retrieval configuration (solid yellow, red dash-dotted, and blue dashed lines, respectively). Left panels: constant input chemistry. Right: equilibrium input chemistry. Top and middle: JWST simulated observation for a 1D and 3D atmospheric assumption, respectively. Bottom: Ariel simulated observation for a 3D atmospheric assumption.

the detailed noise model, the reader is encouraged to read the original article.

Given the description of the payload and a list of candidate exoplanets, ArielRad outputs the expected experimental uncertainty on their measured atmospheric transmission or emission spectra. The simulation propagates the stellar light through the payload, accounting for each transmission or dispersion by interposing optical components until reaching the focal planes. Then, ArielRad evaluates the noise contributions (with margins) from stationary processes: stellar photon noise, detector noise, dark current, zodiacal background, and instrument emission. Jitter noise is computed externally by ExoSim (Sarkar et al. 2021), the end-to-end time-domain simulator of Ariel observations, and included in the final noise budget.

Then, ArielRad returns the uncertainty estimates on a single transit or eclipse observation. Ariel defines an observation to last 2.5 times the time between the first and the last contact between the planetary and the stellar disks to obtain a sufficiently long baseline integration for the light curve fit and the transit depth estimation (Mugnai et al. 2020). Because the astronomical measurement is the contrast ratio with the signal from the stellar host, ArielRad uses the contrast ratio to compute the observed spectrum's expected signal-to-noise ratio (S/N).

The Ariel mission adopts a four-tier observation strategy in which, after each observation, the resulting spectrum from each spectrometer is binned in data analysis according to specific requirements to optimize the S/N and the mission scientific return (Edwards et al. 2019). ArielRad, knowing the binning and

Table 2. Versions of the codes used to generate the Ariel spectra.

Code	Version
ArielRad	2.4.25
ExoRad	2.1.111
ArielRad-Payloads	0.0.16

spectral resolution implemented in the different tiers, computes the S/N in the spectral bin according to the tier of interest. Then, it estimates the number of observations required for each planet to reach the tier's required S/N.

From the number of observations, ArielRad obtains the final noise estimate for a planet by rescaling the noise on a single observation. These uncertainties can be attached to simulated forward models of transmission or emission spectra, binned down to the tier spectral resolution to obtain the simulated observed spectra.

Computing the error bars. We utilize the general procedure described above to calculate the Ariel observation uncertainties for the three planetary targets of interest. We use an updated version of the code, which is a wrapper of ExoRad, the instrument-independent version of the radiometric simulator publicly available on GitHub[2], and ArielRad-Payloads, the repository of configuration files for the payload maintained by the Ariel Consortium. For reproducibility, we report the code versions in Table 2.

For each planet, we assume a mean molecular weight of 2.3 a.m.u. to simulate H_2-He dominated atmospheres. We use this parameter to calculate the atmospheric scale height and, consequently, the contrast ratio of the transit. We acknowledge that assuming an H_2-He dominated atmosphere conflicts with the GJ 1214 b model detailed above, which uses a much higher metallicity. Assuming a lower mean molecular weight leads us to an overoptimistic estimate of the number of transits for the modeled spectrum. Therefore, the S/N (and retrievals) will reflect a worst-case scenario.

We utilize the Ariel strategy for collecting data during an observation; therefore, the ratio of observing time in and out of transit is 1/1.5. Then, for each planet, we estimate the noise per spectral bin and from there, the S/N for a transit observation in Tier 3. In this Tier, the raw spectral data from each spectrometer are binned at $R = 20, 100, 30$ in FGS-NIRSpec (1.1–1.95 μm), AIRS-Ch0 (1.95–3.9 μm), and AIRS-Ch1 (3.9–7.8 μm), respectively. We find that 1, 3, and 4 observations are needed to achieve the Ariel Tier 3 required S/N for HD 189733 b, GJ 1214 b, and WASP-121 b, respectively. Then, we rescale the noise by the square root of the corresponding number of observations, assuming each observation has a Gaussian noise distribution. Finally, we attach the rescaled noise estimates to the respective transmission spectra, binned down at the Tier 3 wavelength grid. It should be noted that we do not scatter the spectra according to random noise corresponding to the estimated error bars. This was done throughout the paper to avoid introducing a susceptibility to individual random noise realizations, which would defeat the purpose of characterizing retrieval biases and finding intrinsic correlations between the atmospheric parameters. While using unscattered spectra may result in unrealistically precise constraints, if the spectra contain sufficiently redundant

information, discrepancies vs. using scattered spectra should not be too large (Feng et al. 2018; Changeat et al. 2020a).

Alternatively, given the noise for a single transit observation and the atmospheric spectrum, we could calculate a more realistic S/N that does not rely on the assumed atmospheric scale height (Mugnai et al. 2021b). However, the S/N would depend on the assumed atmospheric spectrum, changing the number of observations on a single target. However, (i) an observability study is outside the scope of this paper, and (ii) in the following we show that even for GJ 1214 b (lower number of transits than realistic given the mean molecular weight assumed) we can investigate the main effects of interest.

All the spectra calculated with the methodology described in this section are shown in Fig. 1. This includes 6 input configurations for each planet (listed Table 3).

3. Transmission spectra retrievals

3.1. TauREx

As a retrieval tool, we used TauREx 3 (Tau Retrieval for Exoplanets)[3] a fully Bayesian inverse atmospheric retrieval framework (Al-Refaie et al. 2021). TauREx 3 consists of two main frameworks: the Forward Model framework and the Retrieval framework. The goal of the Retrieval framework is to fit a Forward Model to an observation. A Forward Model framework is necessary to provide information about the planet, the host star, temperature-pressure profiles as well as chemistry and contributions (e.g., collision-induced absorption (CIA), limited to H_2-He and H_2-H_2 in the current study, Rayleigh, gray clouds). TauREx 3 adopts the layer-by-layer approach for the temperature profile, which can be parameterized in different ways, such as isothermal, a radiative two-stream approximation, a custom profile loaded from a file, or a multi-point temperature profile. The vertical pressure profile is equally spaced in log-pressure, between a P_{max} and a P_{min} value specified by the user along with a number of layers N_l ($N_l = 200$ at current study). The cloud model provided by TauREx 3 is discretized along $P(z)$, allowing the user to define an opacity value in square meters for layers between the pressure at the top and the pressure at the bottom of the cloud deck.

TauREx 3 supports equilibrium chemistry using the ACE chemical code (Agúndez et al. 2012, 2020), FastChem (Stock et al. 2018), GGchem (Woitke et al. 2018), and the Free chemistry model (Al-Refaie et al. 2022a). In our study, we explore different combinations of chemistry and contribution parameters. We performed retrievals with and without clouds, using ACE, FastChem, GGchem, or Free chemistry. To perform retrievals, TauREx 3 can use several sampling techniques PyMultiNest and MultiNest, PolyChord, or dyPolyChord. For our study, we used the nested sampling retrieval algorithm Multinest with its Python version PyMultiNest (Feroz et al. 2009). The Multinest algorithm samples the parameter space by subdividing it into a set of ellipsoids according to the likelihood. This set of ellipsoids can overlap, but in cases where there are several local maxima in the parameter space, the result will be a set of multiple solutions (for more details, see Feroz & Hobson 2008 and Feroz et al. 2019). TauREx 3 is a full Bayesian retrieval framework, which returns the best-fit transmission model spectrum along with all parameter posterior distributions and the Bayesian evidence. For model comparisons, we use the Bayesian evidence as defined by (Trotta 2008; Waldmann et al. 2015) to compute the logarithmic

---

Table 3. Spectra input configurations.

Instrument	JWST			Ariel		
Dimension	1D			3D		
Chemistry	Constant	Equilibrium	Constant	Equilibrium	Constant	Equilibrium

Bayes factor,

$$\Delta \log E = \log \frac{E_{model_A}}{E_{model_B}} = \log E_{model_A} - \log E_{model_B}, \qquad (1)$$

where E_{model_A} and E_{model_B} provide the evidence of two competing models. According to Benneke & Seager (2013), by translating these Bayes factors into a statistical significance (Kass & Raftery 1995), $\Delta \log E \leq 2$ can be considered as a weak case of models distinguishability, while $2 < \Delta \log E < 5$ corresponds to moderate and $\Delta \log E \geq 5$ to strong cases of distinguishability. Since the Bayesian evidence of the different retrieval models depend on the number of free parameters, and that the free chemistry model has significantly more free parameters compared the equilibrium chemistry models, a Bayesian factor that favors the free chemistry model means a more significant improvement in goodness of fit to observations.

Finally, a molecule is detected if it has a signature greater than 3σ, which is retrieved.

3.2. Chemical model

We consider in this study two types of chemistry that we are going to detail here: Free chemistry and equilibrium chemistry.

Free chemistry takes into account each chemical species with constant abundances throughout the layers of atmosphere. The Free chemistry models are considering the following species: H_2O, CO, CH_4, CO_2, HCN, NH_3, FeH, SiO, Na, K, TiO and VO. This configuration gives the model a certain degree of freedom, as it imposes no physical or chemical constraints on what will be retrieved. In this way, it is possible to retrieve abundances corresponding to non-equilibrium chemistry, or various other species distributions since the species are not correlated with each other. However, it could also retrieve nonrealistic chemical abundances. It should be remembered that this is a 1D model and we are therefore limited when faced with strong vertical variation (which could be compensated for by the two-layer method of Changeat et al. 2019).

Chemical equilibrium is based on temperature and pressure conditions, this assumption is a classic hypothesis when considering exoplanetary atmospheres (Seager & Sasselov 2000; Burrows et al. 2007, 2008; Fortney et al. 2008; Madhusudhan et al. 2011; Kataria et al. 2014; Al-Refaie et al. 2021). A system is at a thermodynamic equilibrium state when there is thermal, mechanical, and chemical equilibrium at the same time. This equilibrium is characterized by the minimum of a thermodynamic potential, such as the Gibbs free energy. It happens in exoplanets' atmospheres when the dynamical timescales can be considered longer than the chemical reaction timescales and when we suppose negligible the irradiation by a dissociating or ionizing source (photochemistry or cosmic rays induced processes). Thus, chemical abundances vary with altitude according to the retrieved TP profile, as we assume they are at chemical equilibrium. As the chemical profiles are not forced to be vertically constant, this approach should be more accurate for

real atmospheres than the Free chemistry approach. For very hot planets this approximation is close to reality, on the other hand, for cooler planets, vertical mixing and photodissociation have an effect on the chemical composition and the atmospheres are no longer at a thermodynamic equilibrium state. This disequilibrium chemical composition must then be taken into account with a more complex kinetic model (Cooper & Showman 2006; Moses et al. 2011, 2012; Venot et al. 2012, 2020; Molaverdikhani et al. 2019; Tsai et al. 2021, 2022; Al-Refaie et al. 2022b). If the observed planet's atmosphere exhibits these non-equilibrium mechanisms, or longitudinal–latitudinal heterogeneities, the retrieved parameters will be erroneous. In such cases, it may not be possible to find a consistent fit, or the retrieval may find an adequate fit that corresponds to erroneous parameters.

Although several different algorithms have been made to calculate the chemical composition at equilibrium, we focused on three algorithms for this study: ACE (Agúndez et al. 2012, 2020), GGchem (Woitke et al. 2018) and FastChem (Stock et al. 2018). The basic principle of these three models is the same, they start from an initial composition made up of initial abundances for each molecule taken into account then they iterate until a convergent state. However, each model has a slightly different procedure, whether in the molecules taken into account as input, the iteration method, or the network of chemical reactions. We study in more detail these differences in this part.

ACE minimizes the total Gibbs free energy by applying the algorithm first introduced by White et al. (1958). ACE is based on an algorithm implemented in the NASA/CEA program and presented in detail in Gordon & McBride (1994). For a closed system of N chemical compounds at a certain temperature and pressure, in the absence of disturbance (transport, UV radiation, etc.), the equilibrium chemical composition can be calculated theoretically, thanks to standard-state chemical potential expressed as a function of the standard-state enthalpy and entropy of the species. These thermodynamic quantities can be calculated using NASA polynomial coefficients (see, e.g., McBride et al. 2002) in databases such as NASA/CEA (McBride et al. 2002) or the Third Millennium Thermochemical Database (Goos et al. 2016). The chemical species used include 105 neutral species composed of C, H, O, and N, more specifically species up to 2 carbon atoms and the main nitrogen species (NH_3, HCN, N_2, NO_x). It has been validated for temperatures as low as 300 K. We refer to Venot et al. (2012) for more details on the ACE code thermodynamic coefficients and calculation of thermochemical equilibrium.

Both FastChem and GGchem use a second type of method for determining the chemical composition at the equilibrium state. These two programs use the law of mass action and equilibrium constants, with some subtleties (for FastChem equilibrium constants are based on Gibbs free energy while for GGchem they are based on partition functions). This amounts to solving a system of N algebraic equations with N unknowns, which correspond to the conservation equations for N elements. The partial pressure of each molecule is defined as the partial pressure of the constituent atoms by the atomization equilibrium constant.

This system of equations can then be solved by any root-finding algorithms like the Newton–Raphson method for example (Russell 1934; Brinkley 1947; Tsuji 1973).

The thermodynamical data used in FastChem are mainly from the NIST-JANAF database detailed in Chase (1998). The list of species used has been modified to take into account molecules that may be of interest in astrophysics with data from Tsuji (1973), Barin & PLatzki (1995), Burcat & Ruscic (2005) and Goos et al. (2016). The total of species used amounts to 396 neutral and 114 charge species and the code has been validated for parameters down to 100 K and up to 1000 bar. We refer to Stock et al. (2018) for more details on the list of species used. We note that the FastChem code has recently been updated with FastChem 2 (Stock et al. 2022) which is more efficient but not yet available in combination with TauREx.

Compared to the other two codes we are using, GGchem takes condensation into account. In fact, the formation of liquids and solids in the atmosphere will have an effect on the composition at thermodynamic equilibrium. Condensed species can consume certain elements leaving a significant difference in composition between before condensation and after condensation (Woitke & Helling 2004; Juncher et al. 2017). Condensation will have an effect, especially at temperatures below 2000 K so this mechanism will mainly affect GJ 1214 b and HD 189733 b in our work. The data included in this code includes 552 molecules and 257 condensates, including 38 liquids, and GGchem has been proven robust down to 100 K. All elements from hydrogen to zirconium are included, as well as the option to add tungsten and charges. We refer to Woitke et al. (2018) for more details on the list of species. We note that the list of active molecules that can appear as a feature in the spectra are H_2O, CO, CH_4, CO_2, HCN, NH_3, FeH, SiO, Na, K, TiO and VO. It depends on the opacity files loaded in the TauREx program and is independent of the type of chemical model chosen.

3.3. Retrieval procedure

All spectra configurations are retrieved with the same set of retrieval models. The set of retrieval models covers all the simulated configurations. We expect that each simulated spectra will be best retrieved by the corresponding retrieval model.

Each retrieval model assumes a four points temperature-pressure (TP) profile (T_{top}, $T_{surface}$, T_1, and T_2) with the corresponding pressure level P_1 and P_2 free to converge between $P_{surface}$ and P_{top}. It has already been shown (Rocchetto et al. 2016; Pluriel et al. 2022) that retrieving an isothermal temperature profile generates biases for hot planets. We also retrieve the radius of the planet. We duplicate our retrievals by adding a gray cloud level as a parameter. Finally, we took into account two chemical configurations in the retrievals, which we call Free and equilibrium chemistry (see Sect. 3.2 for more details).

For the Free models, we retrieve one value of abundance for each considered species : H_2O, CO, CH_4, CO_2, HCN, NH_3, FeH, SiO, Na, K, TiO and VO. We note that for GJ 1214 b we did not retrieve Na, K, TiO, and VO because these species cannot be in gaseous form under these temperature and pressure conditions. Equilibrium chemistry is calculated using three different models included in TauREx: ACE, FastChem, and GGchem (details in Sect. 3.2). Metallicity (Z) and C/O ratio, from which abundance profiles are derived, are retrieved. ACE, FastChem and GGchem have already been shown in Al-Refaie et al. (2022a) to be equivalent using the same molecules and without condensation for GGchem. Thus, we consider all the molecules of each model as well as condensation for GGchem.

Table 4. Free parameters and priors for the retrievals.

Parameters	Bounds		
	GJ 1214 b	HD 189733 b	WASP-121 b
T_{top} (K)	100–2000	500–2500	450–3750
T_{bot} (K)	100–2000	500–2500	450–3750
T_1 (K)	100–2000	500–2500	450–3750
T_2 (K)	100–2000	500–2500	450–3750
$\log_{10}(P_1)$ (Pa)	2–6	2–6	2–6
$\log_{10}(P_2)$ (Pa)	–1 to 6	–1 to 6	–1 to 6
$\log_{10}(P_{clouds})$ (Pa)	–2 to 6	–2 to 6	–2 to 6
Radius (R_{jup})	0.1–0.5	0.5–2.0	1.0–2.5
Chemistry			
Equilibrium			
log(Z)	–1 to 3	–1 to 3	–1 to 3
C/O	0.01–2	0.01–2	0.01–2
Constant			
$\log_{10}(H_2O)$	–12 to –1	–12 to –1	–12 to –1
$\log_{10}(CO)$	–12 to –1	–12 to –1	–12 to –1
$\log_{10}(CH_4)$	–12 to –1	–12 to –1	–12 to –1
$\log_{10}(CO_2)$	–12 to –1	–12 to –1	–12 to –1
$\log_{10}(HCN)$	–12 to –1	–12 to –1	–12 to –1
$\log_{10}(NH_3)$	–12 to –1	–12 to –1	–12 to –1
$\log_{10}(FeH)$	–12 to –1	–12 to –1	–12 to –1
$\log_{10}(SiO)$	–12 to –1	–12 to –1	–12 to –1
$\log_{10}(Na)$	–12 to –1	–12 to –1	–12 to –1
$\log_{10}(K)$	–12 to –1	–12 to –1	–12 to –1
$\log_{10}(TiO)$	–12 to –1	–12 to –1	–12 to –1
$\log_{10}(VO)$	–12 to –1	–12 to –1	–12 to –1

Therefore, we end up with 4 retrieval models (Free, ACE, FastChem, GGchem) for each of the 3 planets (GJ 1214 b, HD 189733 b, WASP-121 b) considering each of the 6 input configurations (listed Table 3). This makes 4×3×6 = 72 retrievals (×2 with clouds). For better readability of the large number of results, we do not show GGchem results for GJ 1214 b, where the temperature is not high enough to be affected by the additional condensation considered by GGchem, and we do not show ACE results for WASP-121 b, where the missing species in ACE mean that the model is not representative of this type of planet (see Sect. 3.2 for more details on these chemical models). Retrievals are compared to each other considering their relative Bayes Factor as described in Sect. 3.1, Eq. (1), following the same idea developed in Tsiaras et al. (2018) with a baseline. However, we define here the Bayes factor (ΔlogE) as the difference with the logarithmic evidence of the worse model (lower one). This allows us to compare the different models between each other as it is done in Panek et al. (2023).

Table 4 shows free parameters and priors for the retrievals. We used a uniform sampling in log space for chemical abundances in the Free chemistry model, the metallicity, the pressure P_1 and P_2, and the pressure of the clouds, and a uniform sampling in linear space for the temperatures, the radius, and the C/O ratio.

4. Results and discussion

4.1. From 1D to 3D

In this section, we present how transmission spectra are affected when considering 1D or 3D GCM models, with or without equilibrium chemistry. We have developed this study from warm to

ultra-hot planets, but with some specific differences on metallicity and planet gravity at the theoretical surface. This highlights the fact that it is not just the effective temperature trend that brings biases, but that other parameters are also involved.

Constant chemistry. We first focus on the simulated spectra assuming constant chemistry simulations which correspond to the left panels of Fig. 2.

We consider in this case that the abundances are constant everywhere in the atmosphere no matter the temperature and pressure conditions. Thanks to this assumption, the effects on the spectra are only due to the thermal structure of the atmosphere. For our coldest case GJ 1214 b, we see strong differences comparing 1D and 3D transmission spectra. Indeed, the 3D GCM model of GJ 1214 b (Charnay et al. 2015) shows large day-night asymmetries with an extended hot day side combined with a 30° eastward shifted hot spot. The metallicity of 100 is likely to result in a higher day-night temperature contrast compared to a solar metallicity, as the chemical composition of the atmosphere can have a significant impact on the efficiency of day-night energy redistribution (Kataria et al. 2014). Due to this shift, the light coming from the star thus probes through the hot day side with a larger scale height compared to the limb (which is used to compute the 1D temperature profile of the pseudo-1D model). This implies hundreds of ppm differences in the transit depth in particular in the major absorbers, such as carbon dioxide and methane. However, on the water-dominated bands in the far infrared and in the visible, we see very small differences between 1D and 3D transmission spectra, on the order of 50 ppm. The altitude where the atmosphere is opaque on these bands is indeed deep in the troposphere (around 100 mbar). Thus, even when we take into account the inflated day side in 3D, the effective radius observed is very similar to that in 1D, because the region where the atmosphere is opaque remains at the same place at the limb. Therefore, the impact of the 3D structure of the atmosphere will depend on the wavelength and the composition.

Similar results are shown concerning the hottest study case WASP-121 b. 3D GCM models show that for highly irradiated atmospheres in tidal locking, the radiative timescale becomes substantially smaller than the dynamic timescale implying that almost no heat is transported from the day to the night side. It results in an extremely large day-side scale height compared to the night side because of the large day-night temperature contrast. Indeed, Keating et al. (2019) showed that regardless of the day-side temperature, the night-side temperature of short-period gas giants is relatively uniform, around ~1100 K. This very inflated day side is thus mainly probed during the transit which explains why 3D transmission spectra are by thousands of ppm larger than in 1D. Unlike the previous case GJ 1214 b, all absorption bands are shifted. The atmosphere is so inflated for this ultra-hot Jupiter that the altitude at which the atmosphere is opaque is much higher due to the much greater scale height.

Interestingly, the results are very different in our intermediate case HD 189733 b. Here, despite a strong day-night temperature contrast, we observe less than a 50 ppm difference between 1D and 3D transmission spectra. This means that the west limb represents well the observable and that such atmospheres are more homogeneous than colder or hotter atmospheres. This is due to the high surface gravity of the planet, which mitigates the scale height differences between the different sides of the planet. Thus, using only the limb is a fair representation of the observable.

These 3D effects are independent of the differences between the east-west limbs, even though they are both related to temperature variations. Figure A.1 shows the transmission spectra of the two limbs independently compared to the whole spectra. This shows that we have similar features but with a different scale height, since the western limb is cooler than the eastern limb (see Fig. C.2). The differences are on the order of 100–600 ppm. This is also observed by Espinoza & Jones (2021), who show that these differences can be observed by the JWST.

Equilibrium chemistry. We now look at the equilibrium chemistry simulations which are shown in the right panels of Fig. 2. Now, the simulated spectra are impacted by both the effects of the chemical and the thermal heterogeneities which is more realistic.

For GJ 1214 b, the global picture has completely changed compared to constant chemistry. The largest difference (1000–2000 ppm) concerns CO_2 bands at 2.5 and 4.5 μm. The temperature profile of the 1D model is not hot enough to obtain abundant CO_2 at the equilibrium state, whereas the day side of the planet in the 3D model reaches a temperature where CO_2 is abundant enough to show strong signatures. Furthermore, as explained above, we probe a nonnegligible part of the planet's day-side, hence the presence of broad bands of CO_2. We also see differences in the water bands (between 1 and 2 μm) which were not present in the constant chemistry model. The reason for these differences is the longitudinal variation in water abundance, which is lower on the day side of the planet. Light from the star then probes deeper regions corresponding to lower transit depth. Between 5 and 9 μm, as well as around 3.5 μm, in the region of the methane bands, the 1D and 3D spectra show fewer discrepancies than the rest of the spectra. Indeed, looking at the methane abundances in Fig. D.7, we see that its abundance is drastically reduced in the day side above 100 mbar which is deeper than where we probe. It results that in these bands, we are not affected by the hot day side and we are mainly probing at the limb which is equivalent to the 1D spectrum.

For WASP-121 b, the 1D and 3D spectra show few differences in the whole wavelength range using equilibrium chemistry compared to constant chemistry. As shown in Fig. D.9, the abundances of almost every species drastically diminished on the day side, mainly due to thermal dissociation (Parmentier et al. 2018), with the exception of carbon monoxide which is only divided by two due to its dilution in an H-dominated day side instead of a H_2-dominated atmosphere. This implies that, on the water bands, the spectrum is not affected by the day side of the atmosphere because water is almost not present there. That is why a 1D model manages to fit the spectrum as shown in Pluriel et al. (2020a). However, we can see in the residuals of Fig. 2 that in some bands, the fit is clearly less good than in the other bands. This is particularly true for the CO bands around 2.5 and 4.5 μm, as well as for the TiO and VO bands in the visible. As we explained, CO is present on the day side of the atmosphere, where extreme temperatures induce a scale height far greater than the scale height of the limbs. Consequently, the 1D model does not represent this behavior well, resulting in a large difference (around 300–400 ppm) at these regions of the spectrum.

The differences between 1D and 3D for HD 189733 b with equilibrium chemistry are very low, as with constant chemistry. Indeed, we see in Fig. D.8 that the main species such as H_2O, CH_4, CO_2, and NH_3 do not display strong longitudinal variations. In addition, we have shown before that due to similar scale height, there is no significant impact comparing 1D and 3D spectra with constant chemistry. As a consequence, a 1D model at the limb is equivalent to the 3D model meaning that for such warm atmospheres, we are probing near the limb.
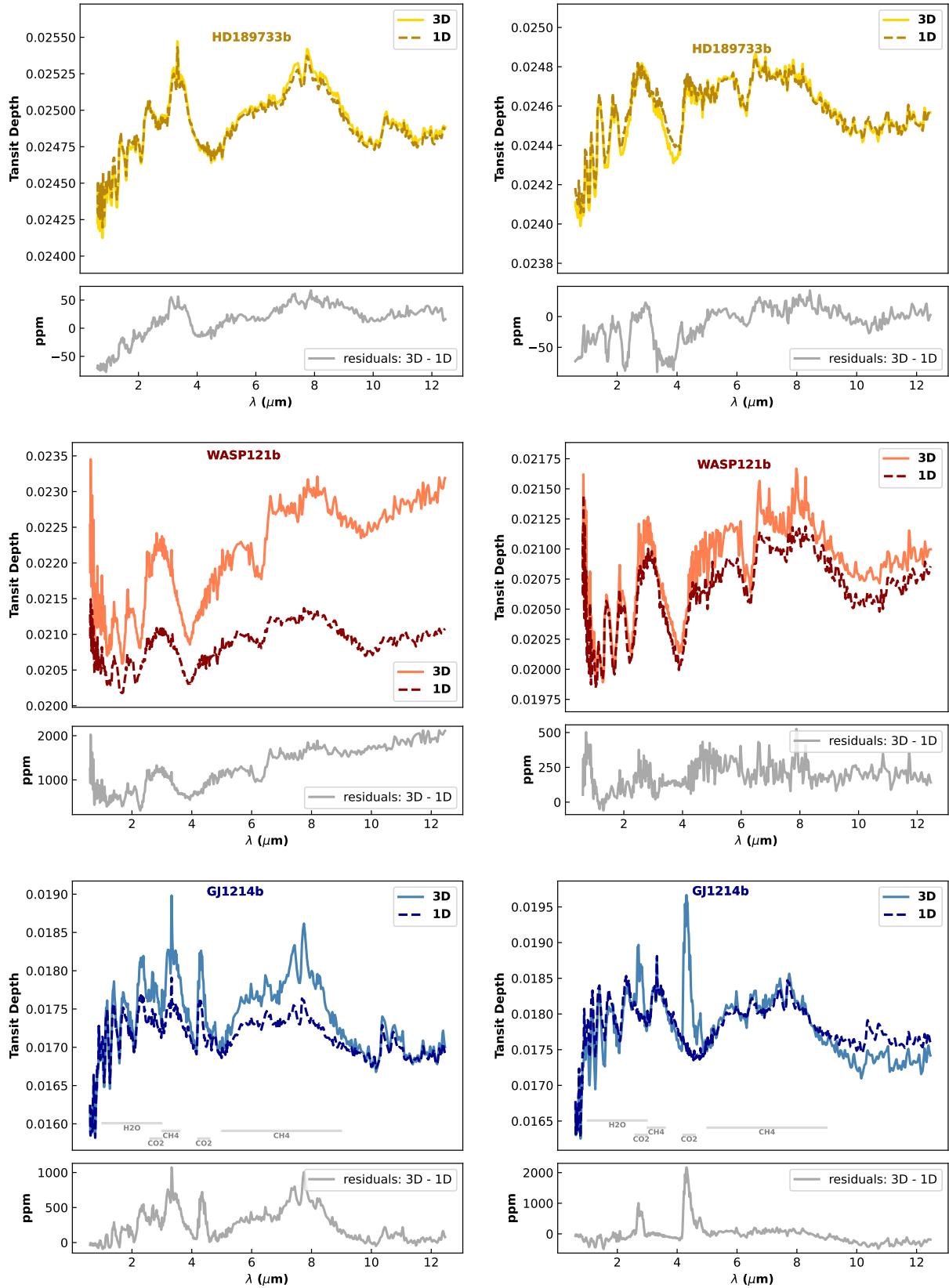
Fig. 2. Transmission spectra simulated with Pytmosph3R ([Falco et al. 2022](#)) for HD 189733 b (top), WASP-121 b (middle), and GJ 1214 b (bottom). Each panel compares two transmission spectra based on a 1D and a 3D atmosphere (dashed and solid lines, respectively). Left panels: constant input chemistry. Right: equilibrium input chemistry. The differences between the 3D and the 1D spectra are plotted below each panel in gray.

These 3D effects are independent of the differences between the east-west limbs, as explain for the constant chemistry.

4.2. Cloud effect

Even if we have good reasons to think that cloud decks are present in many exoplanets (Parmentier et al. 2016; Tsiaras et al. 2018), each atmospheric model used in this study, 1D or 3D, for the three planets, is cloudless. It would be interesting to add clouds in the simulations, in particular, because they affect the short wavelengths observed by JWST and Ariel. However, the aim of this study is to see the impact of chemical and thermal 3D effects on the transmission spectra and how to deal with these 3D effects in the context of atmospheric characterization using 1D retrieval models. For this reason, we chose to not overcomplicate our models. Clouds would require a dedicated paper. Clouds are nevertheless part of the retrieval parameters in the TauREx model. This can break up possible degeneracies (Pluriel et al. 2022; Changeat et al. 2020b), and we verified that the model works correctly by not retrieving a cloud layer when we knew that none was implemented. For each retrieval performed, the Bayes factor has always privileged a cloudless model, and in the retrievals assuming clouds, the cloud deck was always pushed near to the surface pressure thus without impact on the spectra. We thus decided not to present retrievals including clouds in this paper as they bring no more information compared to cloudless retrievals.

4.3. What if atmospheres are 1D?

We used the theoretical 1D atmosphere (see Sect. 2.1 for construction) to check the correct behavior of the 1D retrieval code. We remind that the 3 planets have been chosen to study the retrieval biases depending on the effective temperature of the planets, from warm Neptune to ultra-hot Jupiter. In addition, the chemical construction with a constant profile or equilibrium chemistry has been considered to unravel the retrieval biases from temperature (constant chemistry) and chemistry (equilibrium chemistry). To do so, Free retrievals (constant chemistry) and equilibrium retrievals are both performed for each configuration (more details Sect. 3.3).

Figure 3 and Tables B.1–B.3 show that for all configurations the best retrieval is consistent with the input model configuration. This means that Free retrievals fit better input constant chemistry and equilibrium retrievals fit better input equilibrium chemistry. However, the best retrieval is not always more significant than the others and the equilibrium retrieval models are not all adapted for the different configurations. Figures C.1, D.1–D.3 show that the temperature and species profiles are mostly not well retrieved below the probed altitudes (deeper than $\sim 10^2$ Pa). This part of the atmosphere does not contribute to the features of the spectra, which explains why they are not well retrieved. However, the retrieved values could be well constrained while the input values is often outside the uncertainty. Thus, we cannot trust the retrieved profiles of the lower atmosphere.

Temperate-Warm planet: GJ 1214 b. The best retrieval is consistent with the input model configuration and significantly better than the other models with ΔlogE$\geq$ 19 for constant input chemistry and ΔlogE$\geq$ 7 for equilibrium input chemistry (Fig. 3 and Table B.1).

– Constant input chemistry: For lower pressure than $\sim 10^2$ Pa (high altitudes), Free retrieval shows a significantly better retrieved temperature profile within 60 K compared to the input profile (Fig. C.1). The main absorber CH_4 is retrieved by the Free retrieval at –0.11 dex while H_2O, CO_2 and NH_3 giving secondary features are retrieved within –0.15 dex (Fig. D.1). This allows the best model to be significantly in agreement with the input model even if not all input profiles are included in the uncertainty.

– Equilibrium input chemistry: For lower pressure than $\sim 10^2$ Pa, all retrievals give the same temperature profile within 90 K of the input model (Fig. C.1). All retrievals also give the same main absorber profiles (CH_4 and H_2O) below $\sim 10^2$ Pa, where the input equilibrium chemistry is constant at these temperature and pressure conditions (Fig. D.4). However, ACE is better at retrieving NH_3 profile, which varies by 4 orders of magnitude and contributes at 3 µm, which makes ACE the best model in agreement with input ACE chemical composition. We can highlight that Free and ACE retrievals prefer to retrieve at least +0.31 dex CH_4 and down to –0.17 dex H_2O than input values to better fit the spectrum, with input values out of the uncertainty. FastChem displays some discrepancies with ACE, as pointed out in Al-Refaie et al. (2022a) when using an isothermal profile. ACE fits better than FastChem which is consistent with input ACE composition. However, none of the equilibrium models can reproduce the input composition. The metallicity is retrieved by ACE within 19% and the C/O ratio only at 96% deviation to the input value. While FastChem, which is not considered the best retrieval, has a closer value to input for C/O ratio (76%) and metallicity (9%) (see Table B.1 and Fig. E.1). It shows that even at these low temperatures we already have difficulties to perfectly retrieve the input model.

To summarize for GJ 1214 b: the retrievals work well on the temperature, but not without some biases on the chemistry. Constant chemistry and 1D temperature profiles are retrieved by Free retrieval within –0.15 dex. However, even at such low temperatures, equilibrium input chemistry is best retrieved by equilibrium models because of species with strong vertical variations (here NH_3). In addition, TauREx not only has difficulty in accurately retrieving the C/O ratio and the metallicity (respectively 96% and 19% deviation for the best model ACE), but also the values closest to the input are not retrieved by the more significant model ACE, despite the fact that the best retrieval ACE is consistent with the input ACE equilibrium chemistry.

Warm-hot planet: HD 189733 b. The best retrieval is consistent with the input model configuration and significantly better than the other models with ΔlogE$\geq$ 2536 for constant input chemistry and ΔlogE$\geq$ 200 for equilibrium input chemistry (Fig. 3 and Table B.2).

– Constant input chemistry: For lower pressure than $\sim 10^2$ Pa (high altitudes), Free retrieval shows a significantly better fit of the temperature profile within 150 K compared to the input profile (Fig. C.1). However, the abundance of the main absorbers CH_4, H_2O, NH_3, TiO, and VO, are overestimated by +0.23 to +0.36 dex. This remains a better fit compared to equilibrium chemistry retrievals that show strong vertical variation in these temperature-pressure conditions (Fig. D.2).

– Equilibrium input chemistry: For lower pressure than $\sim 10^2$ Pa, Free and FastChem retrievals overestimate the temperature profile by more than 250 K while ACE and GGchem stay mostly within 250 K deviation from the input values. However, these two chemical models underestimate the temperature in the upper atmosphere by more than 500 K (Fig. C.1). In this peculiar configuration, with warm-hot temperature, molecules such as TiO, VO, and K are close to their condensation temperature. Compare to the isotherm configuration of Al-Refaie et al. (2022a), condensation occurs at mid- and high altitudes where the temperature is lower than in the deep atmosphere.
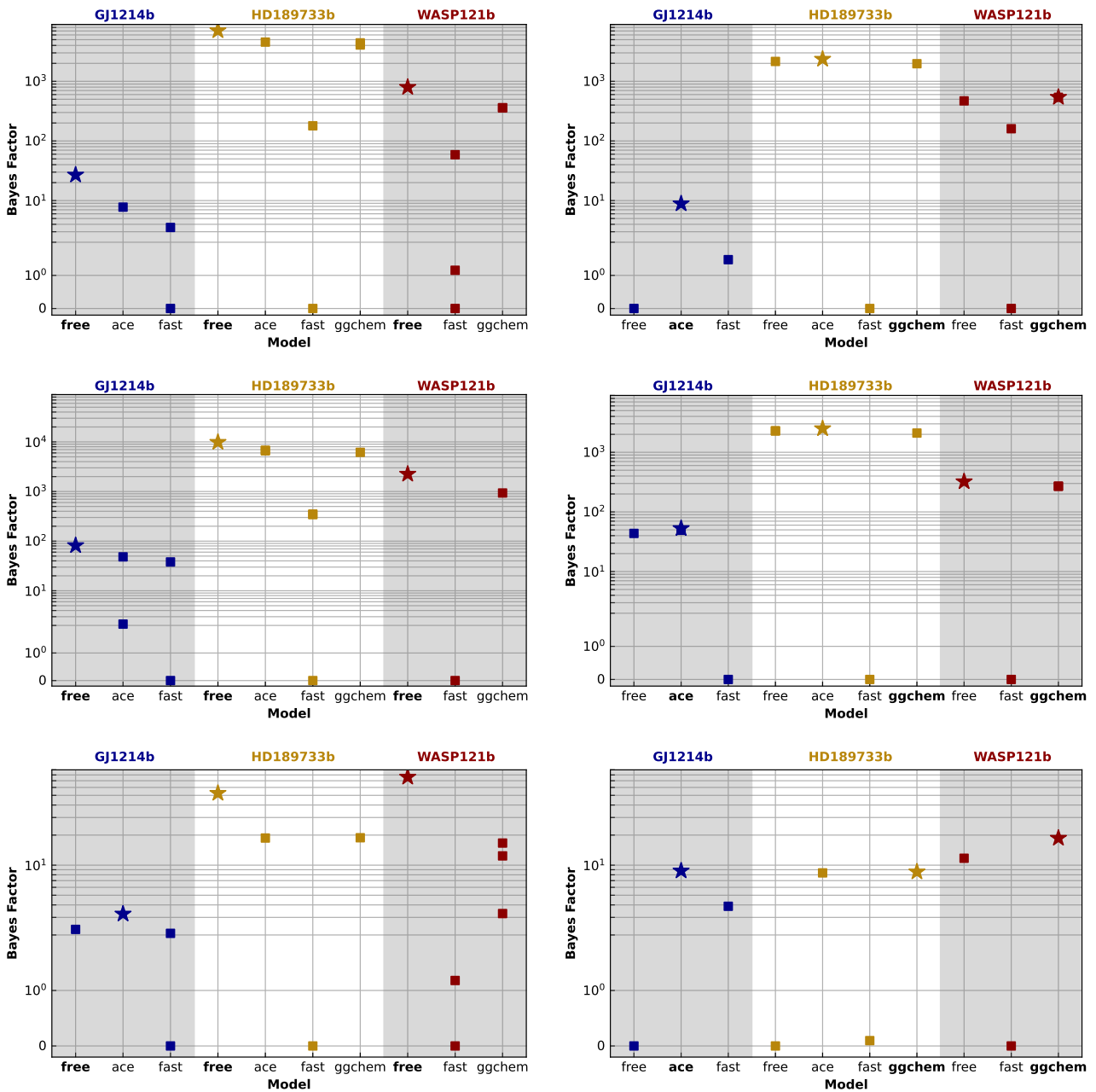
Fig. 3. Bayes factors for each retrieval and all their solutions (Free, ACE, FastChem, and GGchem). By definition, 0 is the solution with the lowest Bayesian evidence as it is the reference of the comparison. Left panels: constant input chemistry. Right: equilibrium input chemistry. Top and middle: JWST simulated observation for a 1D and 3D atmospheric assumption, respectively. Bottom: Ariel simulated observation for a 3D atmospheric assumption. The star represents the highest Bayes factor. The expected best retrieval is highlighted in bold.

Thus, a strong bias occurs with FastChem where it considers these molecules but not their condensates. Therefore, FastChem cannot retrieve properly this configuration where TiO, VO, and K have strong features in the visible while they should not be because of their condensation, Fig. C.1 shows a wrong temperature profile, Table B.2 shows wrong retrieved parameters and Fig. D.5 shows wrong species profiles. Thus, GGchem, which considers condensation of these species, should in theory solve this issue but it surprisingly does not. It shows a better fit for the temperature profile and species abundances with regards to FastChem. But still, it does not manage to give a consistent abundance of K, overestimated by a factor of 2, and shows a strong feature in the visible which bias all the spectra,

since all chemical abundances are correlated. This is explained by the discrepancy between the K abundance calculated by GGchem and the input abundance calculated by Parmentier et al. (2016). We are around 1000 K, at the K condensation limit. This implies a strong variation in K abundance. Woitke et al. (2018) (GGchem) and Parmentier et al. (2016, 2018) use different approximations and assumptions for condensation, which leads to uncertainty, more important at the condensation limit. Al-Refaie et al. (2022a) found a good agreement between ACE, FastChem, and GGchem because the same molecules are considered in the three models. However, we show here chemical discrepancies between the two models, illustrating that using imperfect chemical models with regards to what is actually

observed leads to biased interpretations. ACE, which considers only C, H, O, and N atoms, ends up giving the best-fit model by getting rid of TiO, VO, and K condensation issues. The retrieved thermal profile is similar to the one retrieved by GGchem but the main absorber H_2O is now perfectly constrained under 0.01 dex (see Fig. D.5). All these biases can be seen in Table B.2 and Fig. E.2 on the C/O ratio and the metallicity. FastChem model is far from input values while ACE has a 24% deviation. Furthermore, even if the Free is not the more significant model, it is closest to input values.

To summarize for HD 189733 b: the retrievals work well on the temperature for constant input chemistry (within 150 K) but for equilibrium input chemistry they begin to show difficulties to retrieve the top of the atmosphere (underestimated by 500 K), as well as the bottom of the atmosphere. Only the pressures corresponding to the highest atmospheric contribution (between ~10^2 and ~10^0 Pa) are well retrieved (within 250 K). In addition, we encounter even more chemical abundance bias compared to GJ 1214 b. Constant chemistry profiles are best retrieved by the Free retrieval model (which is consistent) with an overestimation of +0.23 to +0.36 dex. Due to the strong vertical variation, the equilibrium configuration is better retrieved by an equilibrium model (ACE), which is not the expected one (GGchem), but which fits very well the main absorber H_2O under 0.01 dex. However, the study performed on this planet shows all the limitations of the 3 chemical equilibrium models for retrievals. At this temperature, the role of condensates (such as TiO, VO, and K here) is essential with their feature in the visible. The discrepancies between the K chemical modeling of GGchem and the K chemical modeling of Parmentier et al. (2016), bias GGchem retrieval. This is a good example of what can be encountered when fitting observations with a model that does not perfectly reproduce the true chemistry. The C/O ratio and the metallicity are well retrieved by ACE (best model) but not as well as Free. Our results show that we can retrieve values closest to input with a model statistically not favored.

Ultra-hot planet: WASP-121 b. The best retrieval is consistent with the input model configuration and significantly better than the other models with $\Delta logE \geq 434$ for constant input chemistry and $\Delta logE \geq 71$ for equilibrium input chemistry (see Fig. 3 and Table B.3). We observe a higher overall uncertainty in the retrieved temperature and chemical species profiles compared to the other two planets. The very hot temperature coupled with a sharp day-to-night gradient brings a complexity that is more difficult to retrieve with a simpler model.

– Constant input chemistry: in the upper atmosphere, the retrieved thermal profiles are either overestimated or underestimated by ~1000 K (Fig. C.1). This is far from the input temperature profile (~ 50% deviation). However, close to a pressure corresponding to the highest atmospheric contribution, between ~10^2 and ~10^0 Pa, retrieved profiles remain within 500 K (less than 25% deviation), while the best model (Free) is not significantly better. The abundance of the main absorbers H_2O, CH_4 and TiO, FeH in the visible, are overestimated at +0.22 to +0.29 dex for the Free retrieval (best model). Despite these high values, this is still a better fit than equilibrium chemistry. The strong vertical variation in equilibrium chemistry under these temperature and pressure conditions cannot match the constant input chemistry (see Fig. D.3).

– Equilibrium input chemistry: we observe here the same behavior as constant input chemistry for the retrieved temperature profiles (see Fig. C.1). There is less than a 10% deviation between ~10^2 and ~10^0 Pa for the best retrieval GGchem, but

the retrieved profiles are mostly erroneous outside this range (reaching more than 50% deviation). Considering the strong temperature gradient between ~10^2 and ~10^0 Pa, we observe the same behavior as HD 189733 b regarding the condensation of species such as TiO, VO, and K. This makes GGchem a better model than FastChem in this case, as shown in Figs. 3, D.6. Although GGchem gives the best fit, none of the retrieval models fit all the main absorbers H_2O, TiO, CO_2, CO, and VO better than the others. The Free retrieval is better for H_2O and CO_2, while GGchem is better for the other absorbers (despite a strong deviation for both models, reaching +0.50 dex). In addition, the best model, GGchem, retrieves an erroneous C/O ratio with a deviation of 80% and an erroneous metallicity with a deviation of 530%. In contrast, the Free model ends up giving values close to the input within 13% (see Table B.3 and Fig. E.3). Thus, all the models retrieved different parts of the input, but none of them obtained the entire structure.

To summarize for WASP-121 b: the retrievals perform poorly outside the pressures corresponding to the highest atmospheric contribution (between ~10^2 and ~10^0 Pa). Due to strong large-scale vertical variation in temperature and species, 1D temperature profiles for constant and equilibrium chemistry are best retrieved at only 25% between ~10^2 and ~10^0 Pa. Constant input chemistry is best retrieved by Free retrieval, overestimated by +0.22 to +0.29 dex, and equilibrium input chemistry is best retrieved by GGchem retrieval, but with a deviation reaching ~+0.50 dex. The retrieval models are consistent with the input configurations, but not without bias on absolute values. GGchem retrieves the wrong C/O ratio and metallicity, while the Free model is within 13%. The strong vertical gradient on a large scale brings a complexity that is difficult to retrieve correctly with a simpler model, which also translates into greater uncertainty in the retrievals.

Summary. The theoretical 1D analysis validates the consistency of the 1D retrievals, but not without a few biases. This shows that below the probed altitudes, in addition to the retrieved parameters, we cannot trust the uncertainties given by the models, which are largely underestimated. This is also the case at the top of the atmosphere, as we move toward warmer planets.

Furthermore, equilibrium chemical models are not equivalent and give significantly different results:
– ACE cannot retrieve ultra-hot planets whereas it might be a better approximation for cooler planets.
–FastChem is biased toward warm, even hot planets, where species are close to or below condensation temperature. This is never the best retrieval model.
–GGchem, which should be a complete model, will still be in competition with a simplified model for cold and warm planets. Apparently the best option for hot planets, but Free retrieval can give better retrieved values.

We need to be careful with equilibrium models. Our study shows that if one part of the chemistry modeling is wrong, all the chemical abundances will be biased since everything is correlated. This agrees with the conclusion of Al-Refaie et al. (2022a) using an isotherm configuration.

4.4. When atmospheres are 3D

4.4.1. Constant chemistry

Given the 3D thermal structure, we used theoretical atmospheric models with constant chemistry (see Sect. 2.2) to disentangle the 3D effects of temperature without being biased by chemistry. Section 4.3 confirms the overall correct behavior of the retrieval

code or highlights any biases we may encounter. So, we can be confident in this approach to focus on thermal 3D effects. To this end, Free retrievals (constant-with-depth chemistry) and equilibrium retrievals are performed for each configuration (more details in Sect. 3.3). The aim is also to compare the biases of the JWST and Ariel instruments.

Figure 3 shows that the Free retrieval finds the best solution compared to equilibrium chemistry (except for GJ 1214 b in the Ariel configuration). This is consistent with the input constant chemistry. For GJ 1214 b in the Ariel configuration, ACE retrieval has a better Bayes factor but within a range of 1.5 variation compared to the others. Therefore, all models are statistically equivalent and none is preferred. We observed a stronger deviation of the Bayes factor for the JWST configuration compared to the Ariel one. This could simply be due to the higher resolution on a larger wavelength range for the JWST configuration or this may be linked to some particular wavelength band such as the lack of data points in the visible for the Ariel configuration where strong features for hot planets (such as TiO, VO, and K) are located.

Temperate-warm planet: GJ 1214 b. All retrieval models for the Ariel spectrum are within a Bayes factor deviation of 1.5 which makes all models equivalent (see Table B.1). However, the JWST spectrum is better retrieved by the Free retrieval with $\Delta\log E = 33$ compared to the second-best model. The higher resolution of the JWST spectrum, compared to the Ariel spectrum, gives more constraints on H_2O and CH_4 features at low wavelength (between 1 and 2 μm), but probably also at higher wavelength (above 4 μm). Yet, this does not translate to closer-to-truth retrieved profiles, although the retrieved uncertainties are smaller. Figures C.2 and C.3 show consistent inputs and retrieved temperature profiles for GJ 1214 b. At low altitudes (deeper than $\sim 10^2$ Pa) there is high uncertainty because these altitudes are not probed. Higher in the atmosphere (probed altitudes), the temperature day-night variation is within 300 K and all models retrieve the limb profiles. H_2O and CH_4 are responsible for the main features of the spectra. Both retrieved values of the Free retrieval are around –0.10 dex lower than the input. For the JWST spectrum, the uncertainty is lower but the input value is not within the uncertainty. While, for the Ariel spectrum, the CH_4 input value is within the uncertainty. Thus, even with more data points, the retrieved chemical abundances is not closer to the input and the uncertainties cannot be trusted.

Warm-hot planet: HD 189733 b. The best retrieval (Free) is consistent with the input model configuration and significantly better than the other models with $\Delta\log E \geq 3048$ for the JWST spectrum and $\Delta\log E \geq 34$ for the Ariel spectrum (see Fig. 3 and Table B.2). Discrepancies between retrieval models are the same as those explained in Sect. 4.3. Figures C.2 and C.3 show that the biases on the temperature profiles are the same as in Sect. 4.3. Temperatures retrieved below $\sim 10^2$ Pa correspond mainly to those of the limb. For the Ariel spectrum, the temperature is slightly warmer, but the solution for the JWST spectrum remains within the Ariel uncertainties, which are significantly higher. Equilibrium chemistry cannot reproduce the constant input chemistry, while Free retrieval gives consistent results but not without significant deviation (between +0.18 to +0.35 dex). Even with a variation of less than 500 K between day and night side, only temperature variation can bias the retrieval of species abundances. While VO is largely retrieved in the JWST spectrum, it is not in the Ariel spectrum. Retrieved uncertainties on the Ariel spectrum are larger than on the JWST spectrum. However (and to the exception of VO) the same molecules are

present with both observatories. The lack of abundant VO with Ariel is due to the coarser spectral resolution in visible light which is still sufficient to detect TiO.

Ultra-hot planet: WASP-121 b. The best retrieval (Free) is consistent with the input model configuration and significantly better than the other models with $\Delta\log E \geq 1310$ for the JWST spectrum and $\Delta\log E \geq 60$ for the Ariel spectrum (see Fig. 3 and Table B.3). Discrepancies between retrieval models are the same as those explained in Sect. 4.3. Figures C.2 and C.3 show that the Free retrieval (best model) finds a temperature at the top of the atmosphere higher than the input model, and a temperature at the bottom of the atmosphere lower than the input model. The temperature transition occurs in the atmosphere where species absorption contributes the most, around $\sim 10^2$ Pa. The temperature gradient is steep, crossing all possible temperatures from day to night side. The chemical abundances of the main absorbers H_2O, TiO, CO are retrieved between –0.25 and –0.67 dex for the JWST spectrum and between –0.28 and –0.77 dex for the Ariel spectrum, with the exception of TiO with a deviation of +0.03 dex from the input value. However, TiO feature does not match the input spectrum for both JWST and Ariel configurations. The retrieved spectra are outside the uncertainties, by several sigma. The much higher JWST resolution in the visible compared with Ariel surprisingly does not provide better constraints on TiO. While we already encounter difficulties in retrieving input values with a 1D atmosphere, the huge temperature gradient between day and night brings even more biases. The end result is a temperature from both the day and night sides that does not allow the retrieval models to find input spectra and species profiles. There is not even thermal inversion retrieved by the best model.

Summary. The conclusion in Sect. 4.3 remains the same in 3D. Secondly, using for retrievals a 4-point temperature profile gives good results for the cooler planets but not for hotter ones, which need at least 2D retrievals, as has already been pointed out in more detail by Pluriel et al. (2022). The higher resolution of the JWST spectrum, particularly in the visible, reduces uncertainties but does not provide a better fit. In addition, the input values will be out of uncertainty, making them unreliable. This is probably due not only to the high resolution, but also to the good input signal-to-noise ratio, which can be improved by binning down the spectra. Thus, lower resolution would still result in extremely small error bars and overconfident retrieval results. The lower atmosphere is still poorly retrieved, especially as we move toward hotter planets. Nevertheless, it is still possible to find the presence of the input species. Finally, the low resolution in the visible wavelength range of the Ariel spectrum has missed the presence of the visible absorber VO but never TiO with a retrieval similar to that of the JWST spectrum.

4.4.2. Equilibrium chemistry

Here the 3D thermal structure, as well as equilibrium chemistry, are considered as input. Considering the conclusion of Sects. 4.3 and 4.4.1, this will highlight biases due to the variability of chemical abundances in the atmospheres of warm to ultra-hot planets. Section 4.4.1 has already shown the temperature biases from the 3D structure. As previously done, Free retrievals (constant-with-depth chemistry) and equilibrium retrievals are both performed for each configuration (more details in Sect. 3.3).

Figure 3 shows that the equilibrium retrievals always find the best solution compared to Free chemistry (except for the WASP-121 b JWST configuration). This is consistent with the input equilibrium chemistry. The retrieval models of the Ariel

Table 5. Overview of the ability of the best retrievals to find the input values.

	JWST												Ariel					
	1D						3D											
	Constant			Equilibrium			Constant			Equilibrium			Constant			Equilibrium		
	GJ	HD	WASP	GJ	HD	WASP	GJ	HD	WASP	GJ	HD	WASP	GJ	HD	WASP	GJ	HD	WASP
Species detection	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
C/O				×	~	×				✓	~	×				✓	✓	✓
Metallicity (Z)				✓	✓	×				✓	~	×				✓	✓	×
Top of the atmosphere [$\sim 10^0$–$\sim 10^{-4}$ Pa]																		
Temperature profile	✓	✓	×	✓	×	×	Limb	Night	×	Day	×	Night	Night	Day	×	Day	Day	Day
Chemical profile	✓	~	~	✓	~	×	✓	~	~	Day	×	✓	✓	~	~	Day	Day	×
Middle of the atmosphere [$\sim 10^2$–$\sim 10^0$ Pa]																		
Temperature profile	✓	✓	~	✓	✓	✓	Limb	Day	Limb	Limb	Limb	Limb	Limb	Day	Limb	Limb	Limb	Limb
Chemical profile	✓	~	~	✓	✓	~	✓	~	~	Limb	Limb	×	✓	~	~	Limb	Limb	×
Bottom of the atmosphere [$\sim 10^6$–$\sim 10^2$ Pa]																		
Temperature profile	~	×	×	×	✓	×	✓	×	×	✓	Night	Night	✓	×	×	✓	×	×
Chemical profile	✓	~	~	×	✓	~	✓	~	~	×	~	×	✓	~	~	✓	×	×

Notes. For species detection, C/O ratio, metallicity (Z), and chemical profiles, it focuses only on the main absorbers. The temperature and chemical profiles are split depending on the region on the atmosphere; around the highest atmospheric contribution (between $\sim 10^2$ Pa and $\sim 10^0$ Pa) the atmosphere is globally well retrieved, contrary to the bottom of the atmosphere (between $\sim 10^6$ Pa and $\sim 10^2$ Pa). See Sects. 4.3 and 4.4 for more details on the specific biases.

spectrum have less deviation from each other than the JWST spectrum. This is due to the lower spectral resolution across all wavelength bands, but particularly in the visible bands. Between visible and infrared the chemical species contributing to the spectral signatures are different, TiO, VO, and K against H_2O, CH_4, CO_2, CO, and NH_3. If we try to retrieve both parts at the same time, both will be biased, as the signatures may come from different parts of the atmosphere, at different temperatures. While a larger spectral range could benefit to interpret more accurately the observations (Benneke & Seager 2013; Welbanks & Madhusudhan 2019), we show that depending on the chosen range and the model used this could bias the retrieval compared to input data.

Temperate-warm planet: GJ 1214 b. The ACE model provides the best fit to both JWST and Ariel spectra with ΔlogE$\geq$ 10 and ΔlogE$\geq$ 5 respectively (see Table B.1), which is consistent with the ACE input chemistry modeling. Figure 3 also shows that the JWST spectrum is secondly best retrieved by the Free model, while the Ariel spectrum is secondly best retrieved by the FastChem model. FastChem's retrievals poorly fit the CO and the visible bands, because of the chemical modeling differences with the input ACE chemistry modeling. Retrievals on the Ariel spectrum circumvent this issue thanks to its low spectral resolution in visible light. The main difference from the previous input configuration is that the equilibrium chemistry results in a strong dichotomy of CO_2 between the day and night side (see Fig. D.7). As a result, the retrieved temperature profiles correspond to the day side (see Figs. C.2 and C.3). Table B.1 and Fig. E.4 show that the temperature bias on the day side still keeps a good agreement on the C/O ratio and the metallicity retrieved, within 20%.

Warm-hot planet: HD 189733 b. The ACE model retrieves the JWST spectrum much better than the other models with ΔlogE$\geq$ 208, as already explained in Sect. 4.3. However, the Ariel spectrum is equivalently retrieved by ACE and GGchem (Fig. 3 and Table B.2), again due to the lack of constraint in the

visible wavelength bands where discrepancies between chemical models appear. ACE is as significant as GGchem, but Table B.2 and Fig. E.5 show that GGchem better retrieves the C/O ratio and metallicity. Figure C.3 shows that the temperature at the top of the atmosphere is unconstrained by the huge uncertainty. This part of the atmosphere, therefore, makes no significant contribution to the features of the spectra. In contrast, Fig. C.2 shows that increasing resolution adds an erroneous constraint on the temperature of the top of the atmosphere. Only the temperature around pressures corresponding to the highest atmospheric contribution (around $\sim 10^2$ Pa), is consistent between equilibrium models and the input temperature profiles. The temperature retrieved at these pressures is that of the limb. Figure D.8 shows a good agreement between the retrieved species profiles.

Ultra-hot planet: WASP-121 b. The GGchem model retrieves the Ariel spectrum better than the other models with ΔlogE$\geq$ 7, as already explained in Sect. 4.3. However, the JWST spectrum is better retrieved by the Free model with ΔlogE$\geq$ 52 (see Fig. 3 and Table B.3). Table B.3 and Fig. E.6 show that GGchem retrieves for Ariel configuration the C/O ratio very well at 5% but not the metallicity (75% deviation), while for JWST configuration it is higher than 44% considering all retrieval models. The models are not suited to the high spectral resolution of JWST, which imposes severe constraints on thermal contrast and hence on chemical distribution. This shows that such a contrasted atmosphere cannot be retrieved by a 1D model with correlated chemistry. However, the higher degree of freedom of the Free retrieval allows a better match. The temperature profiles retrieved between the Free model and the GGchem model are similar (within 500 K below 10^4 Pa, see Figs. C.2 and C.3). The conclusions on temperature biases are the same as for HD 189733 b. But Fig. D.9 shows that species abundances are more difficult to retrieve.

Summary. In addition to the previous biases from Sects. 4.3 and 4.4.1, the biases coming from the chemistry show that, even on a warm planet, it would make sense to fit the different

molecular features separately to disentangle the temperature variation that brings chemical variability. Otherwise, using a 1D retrieval model will bias all different spectral contributions. Furthermore, only the pressure where the contribution is highest should be considered as a significantly good retrieval of the observation. The rest should be treated with caution. All models remain good at detecting input molecules.

5. Conclusions

We present in Table 5 an overview of the main results obtained in this study. A limitation to the approach described in this paper has been the use of simplified equilibrium chemistry models: recent re-analysis work on transit retrievals from HST has shown the necessity to go toward non-equilibrium chemical models at least for temperate planets (Panek et al. 2023). This additional complexity is still beyond current 3D modeling for retrievals, but will certainly in the future be an important aspect to develop. The 3D effects that are presented above will be an improvement in future retrievals of the JWST observations, such as transit spectroscopy at JWST/NIRSpec resolution on WASP-39 b which presently use 1D models (Alderson et al. 2023). Phase curve observations as observed by HST, *Spitzer*, and JWST on WASP-43 b (Stevenson et al. 2014, 2017; Murphy et al. 2023; Bell et al. 2023) give access to preliminary constraints on the 3D composition, cloud coverage, and temperature structure of the planet's atmosphere thanks to JWST/MIRI/LRS sensitivity. A limitation of the parameter retrieval from the observations is today reached by the complexity of the models: combining GCM, radiative transfer codes, thermochemistry codes, and non-equilibrium chemistry is a formidable task that involves a multidisciplinary effort from various communities of molecular spectroscopists, chemists, meteorologists, and astronomers. Even with the limitations described above, this paper provides caveats to the approaches and will allow future investigators to address these questions properly.

References

Addison, B., Wright, D. J., Wittenmyer, R. A., et al. 2019, PASP, 131, 115003
Agúndez, M., Venot, O., Iro, N., et al. 2012, A&A, 548, A73
Agúndez, M., Martínez, J. I., de Andres, P. L., Cernicharo, J., & Martín-Gago, J. A. 2020, A&A, 637, A59
Ahrer, E.-M., Stevenson, K. B., Mansfield, M., et al. 2023, Nature, 614, 653
Alderson, L., Wakeford, H. R., Alam, M. K., et al. 2023, Nature, 614, 664
Al-Refaie, A. F., Changeat, Q., Waldmann, I. P., & Tinetti, G. 2021, ApJ, 917, 37
Al-Refaie, A. F., Changeat, Q., Venot, O., Waldmann, I. P., & Tinetti, G. 2022a, ApJ, 932, 123
Al-Refaie, A. F., Venot, O., Changeat, Q., & Edwards, B. 2022b, arXiv e-prints [arXiv:2209.11203]
Amundsen, D. S., Mayne, N. J., Baraffe, I., et al. 2016, A&A, 595, A36
Arcangeli, J., Désert, J.-M., Line, M. R., et al. 2018, ApJ, 855, L30
Arcangeli, J., Désert, J.-M., Parmentier, V., et al. 2019, A&A, 625, A136
Baeyens, R., Decin, L., Carone, L., et al. 2021, MNRAS, 505, 5603
Baeyens, R., Konings, T., Venot, O., Carone, L., & Decin, L. 2022, MNRAS, 512, 4877
Barin, I., & PLatzki, G. 1995, Compilation of Thermochemical Data (John Wiley & Sons, Ltd), 33
Barton, E. J., Yurchenko, S. N., & Tennyson, J. 2013, MNRAS, 434, 1469
Barton, E. J., Chiu, C., Golpayegani, S., et al. 2014, MNRAS, 442, 1821
Batalha, N. E., Mandell, A., Pontoppidan, K., et al. 2017, PASP, 129, 064501
Bean, J. L., Stevenson, K. B., Batalha, N. M., et al. 2018, PASP, 130, 114402
Beichman, C. A., & Greene, T. P. 2018, arXiv e-prints [arXiv:1803.03730]
Beichman, C. A., Rieke, M., Eisenstein, D., et al. 2012, SPIE Conf. Ser., 8442, 84422N
Bell, T. J., & Cowan, N. B. 2018, ApJ, 857, L20
Bell, T., Crouzet, N., Kreidberg, L., Moses, J., & JWST Transiting Exoplanet Community ERS Team 2023, in Am. Astron. Soc. Meeting Abstracts, 55, 124.06
Benneke, B., & Seager, S. 2013, ApJ, 778, 153
Brinkley, Stuart R., J. 1947, J. Chem. Phys., 15, 107
Brown, T. M. 2001, ApJ, 553, 1006
Burcat, A., & Ruscic, B. 2005, Argonne National Laboratory, USA, Tech. Rep., https://doi.org/10.2172/925269
Burrows, A., Hubeny, I., Budaj, J., Knutson, H. A., & Charbonneau, D. 2007, ApJ, 668, L171
Burrows, A., Budaj, J., & Hubeny, I. 2008, ApJ, 678, 1436
Caldas, A., Leconte, J., Selsis, F., et al. 2019, A&A, 623, A161
Changeat, Q., Edwards, B., Waldmann, I. P., & Tinetti, G. 2019, ApJ, 886, 39
Changeat, Q., Al-Refaie, A., Mugnai, L. V., et al. 2020a, AJ, 160, 80
Changeat, Q., Keyte, L., Waldmann, I. P., & Tinetti, G. 2020b, ApJ, 896, 107
Charbonneau, D., Berta, Z. K., Irwin, J., et al. 2009, Nature, 462, 891
Charnay, B., & Drossart, P. 2023, Comptes Rendus Phys., 24, 1
Charnay, B., Meadows, V., Misra, A., Leconte, J., & Arney, G. 2015, ApJ, 813, L1
Chase, M. W. 1998, NIST-JANAF Thermochemical Tables, 4th edn. (American Institute of Physics, -1)
Chubb, K. L., Rocchetto, M., Yurchenko, S. N., et al. 2021, A&A, 646, A21
Cloutier, R., Charbonneau, D., Deming, D., Bonfils, X., & Astudillo-Defru, N. 2021, AJ, 162, 174
Cooper, C. S., & Showman, A. P. 2006, ApJ, 649, 1048
Delrez, L., Santerne, A., Almenara, J. M., et al. 2016, MNRAS, 458, 4025
Doyon, R., Hutchings, J. B., Beaulieu, M., et al. 2012, SPIE Conf. Ser., 8442, 84422R
Drummond, B., Mayne, N. J., Manners, J., et al. 2018, ApJ, 869, 28
Edwards, B., Mugnai, L., Tinetti, G., Pascale, E., & Sarkar, S. 2019, AJ, 157, 242
Edwards, B., Changeat, Q., Baeyens, R., et al. 2020, AJ, 160, 8
Encrenaz, T., Tinetti, G., Tessenyi, M., et al. 2015, Exp. Astron., 40, 523
Espinoza, N., & Jones, K. 2021, AJ, 162, 165
Falco, A., Zingales, T., Pluriel, W., & Leconte, J. 2022, A&A, 658, A41
Feinstein, A. D., Radica, M., Welbanks, L., et al. 2023, Nature, 614, 670
Feng, Y. K., Robinson, T. D., Fortney, J. J., et al. 2018, AJ, 155, 200
Feroz, F., & Hobson, M. P. 2008, MNRAS, 384, 449
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2019, Open J. Astrophys., 2, 10
Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, ApJ, 678, 1419
Fortney, J. J., Shabram, M., Showman, A. P., et al. 2010, ApJ, 709, 1396
Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, Space Sci. Rev., 123, 485
Gaudi, B. S., Meyer, M., & Christiansen, J. 2021, in ExoFrontiers, 2514-3433 (IOP Publishing), 2–1
Goos, E., Burcat, A., & Ruscic, B. 2016, Extended Third Millenium Ideal Gas Thermochemical Database with Updates from Active Thermochemical Tables (Technion IIT, Aerospace Engineering, and Argonne National Laboratory)
Gordon, S., & McBride, B. 1994, Computer Program for Calculation of Complex Chemical Equilibrium Compositions and Applications. Part 1: Analysis (NASA)
Guillot, T. 2010, A&A, 520, A27
Guillot, T., Fletcher, L. N., Helled, R., et al. 2022, arXiv e-prints [arXiv:2205.04100]
Guilluy, G., Gressier, A., Wright, S., et al. 2020, ApJ, 161, 19
Jakobsen, P., Ferruit, P., Alves de Oliveira, C., et al. 2022, A&A, 661, A80

Juncher, D., Jørgensen, U. G., & Helling, C. 2017, A&A, 608, A70
Kass, R. E., & Raftery, A. E. 1995, J. Am. Stat. Assoc., 90, 773
Kataria, T., Showman, A. P., Lewis, N. K., et al. 2013, ApJ, 767, 76
Kataria, T., Showman, A. P., Fortney, J. J., Marley, M. S., & Freedman, R. S. 2014, ApJ, 785, 92
Kataria, T., Showman, A. P., Fortney, J. J., et al. 2015, ApJ, 801, 86
Keating, D., Cowan, N. B., & Dang, L. 2019, Nat. Astron., 3, 1092
Kreidberg, L. 2018, in Handbook of Exoplanets, eds. H. J. Deeg, & J. A. Belmonte (Springer International Publishing AG), 100
Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, Nature, 505, 69
Lacy, B. I., & Burrows, A. 2020, ApJ, 905, 131
Leconte, J., Forget, F., Charnay, B., et al. 2013, A&A, 554, A69
Lee, J.-M., Heng, K., & Irwin, P. G. J. 2013, ApJ, 778, 97
Line, M. R., Stevenson, K. B., Bean, J., et al. 2016, AJ, 152, 203
MacDonald, R. J., & Lewis, N. K. 2022, ApJ, 929, 20
MacDonald, R. J., Goyal, J. M., & Lewis, N. K. 2020, ApJ, 893, L43
Madhusudhan, N. 2019, ARA&A, 57, 617
Madhusudhan, N., Burrows, A., & Currie, T. 2011, ApJ, 737, 34
McBride, B. J., Zehe, M. J., & Gordon, S.. 2002, NASA Glenn Coefficients for Calculating Thermodynamic Properties of Individual Species, NASA Technical Report TP-2002-211556, NASA, work of the US Gov.
McElwain, M. W., Feinberg, L. D., Perrin, M. D., et al. 2023, PASP, 135, 058001
Molaverdikhani, K., Henning, T., & Mollière, P. 2019, ApJ, 883, 194
Moses, J. I., Visscher, C., Fortney, J. J., et al. 2011, ApJ, 737, 15
Moses, J. I., Madhusudhan, N., Visscher, C., & Freedman, R. S. 2012, ApJ, 763, 25
Mugnai, L. V., Pascale, E., Edwards, B., Papageorgiou, A., & Sarkar, S. 2020, Exp. Astron., 50, 303
Mugnai, L. V., Modirrousta-Galian, D., Edwards, B., et al. 2021a, AJ, 161, 284
Mugnai, L. V., Al-Refaie, A., Bocchieri, A., et al. 2021b, AJ, 162, 288
Murphy, M. M., Beatty, T. G., Roman, M. T., et al. 2023, AJ, 165, 107
Nixon, M. C., & Madhusudhan, N. 2022, ApJ, 935, 73
Panek, E., Beaulieu, J.-P., Drossart, P., et al. 2023, A&A, 677, A51
Parmentier, V., Fortney, J. J., Showman, A. P., Morley, C., & Marley, M. S. 2016, ApJ, 828, 22
Parmentier, V., Line, M. R., Bean, J. L., et al. 2018, A&A, 617, A110
Parmentier, V., Showman, A. P., & Fortney, J. J. 2021, MNRAS, 501, 78
Pluriel, W., Whiteford, N., Edwards, B., et al. 2020a, AJ, 160, 112
Pluriel, W., Zingales, T., Leconte, J., & Parmentier, V. 2020b, A&A, 636, A66
Pluriel, W., Leconte, J., Parmentier, V., et al. 2022, A&A, 658, A42
Rieke, G. H., Wright, G. S., Böker, T., et al. 2015, PASP, 127, 584
Rocchetto, M., Waldmann, I. P., Venot, O., Lagage, P. O., & Tinetti, G. 2016, ApJ, 833, 120
Russell, H. N. 1934, ApJ, 79, 317

Rustamkulov, Z., Sing, D. K., Mukherjee, S., et al. 2023, Nature, 614, 659
Sarkar, S., Pascale, E., Papageorgiou, A., Johnson, L. J., & Waldmann, I. 2021, Exp. Astron., 51, 287
Seager, S., & Sasselov, D. D. 2000, ApJ, 537, 916
Showman, A. P., Cooper, C. S., Fortney, J. J., & Marley, M. S. 2008, ApJ, 682, 559
Showman, A. P., Fortney, J. J., Lian, Y., et al. 2009, ApJ, 699, 564
Sing, D. K. 2018, arXiv e-prints [arXiv:1804.07357]
Skaf, N., Bieger, M. F., Edwards, B., et al. 2020, AJ, 160, 109
Stassun, K. G., Collins, K. A., & Gaudi, B. S. 2017, AJ, 153, 136
Stevenson, K. B., Désert, J.-M., Line, M. R., et al. 2014, Science, 346, 838
Stevenson, K. B., Line, M. R., Bean, J. L., et al. 2017, AJ, 153, 68
Stock, J. W., Kitzmann, D., Patzer, A. B. C., & Sedlmayr, E. 2018, MNRAS, 479, 865
Stock, J. W., Kitzmann, D., & Patzer, A. B. C. 2022, MNRAS, 517, 4070
Sudarsky, D., Burrows, A., & Pinto, P. 2000, ApJ, 538, 885
Tan, X., & Komacek, T. D. 2019, ApJ, 886, 26
Tennyson, J., & Yurchenko, S. N. 2012, MNRAS, 425, 21
Tinetti, G., Vidal-Madjar, A., Liang, M.-C., et al. 2007, Nature, 448, 169
Tinetti, G., Drossart, P., Eccleston, P., et al. 2018, Exp. Astron., 46, 135
Tinetti, G., Eccleston, P., Haswell, C., et al. 2021, arXiv e-prints [arXiv:2104.04824]
Trotta, R. 2008, Contemp. Phys., 49, 71
Tsai, S.-M., Lee, Malik, M., et al. 2021, ApJ, 923, 264
Tsai, S.-M., Lee, E. K. H., & Pierrehumbert, R. 2022, A&A, 664, A82
Tsai, S.-M., Lee, E. K. H., Powell, D., et al. 2023, Nature, 617, 483
Tsiaras, A., Waldmann, I. P., Zingales, T., et al. 2018, AJ, 155, 156
Tsuji, T. 1973, A&A, 23, 411
Turrini, D., Miguel, Y., Zingales, T., et al. 2018, Exp. Astron., 46, 45
Venot, O., Hébrard, E., & Agúndez, M. e. a. 2012, A&A, 546, A43
Venot, O., Agúndez, M., Selsis, F., Tessenyi, M., & Iro, N. 2014, A&A, 562, A51
Venot, O., Parmentier, V., Blecic, J., et al. 2020, ApJ, 890, 176
Waldmann, I. P., Tinetti, G., Rocchetto, M., et al. 2015, ApJ, 802, 107
Wardenier, J. P., Parmentier, V., & Lee, E. K. H. 2022, MNRAS, 510, 620
Welbanks, L., & Madhusudhan, N. 2019, AJ, 157, 206
Welbanks, L., & Madhusudhan, N. 2022, ApJ, 933, 79
Welbanks, L., Madhusudhan, N., Allard, N. F., et al. 2019, ApJ, 887, L20
White, W. B., Johnson, S. M., & Dantzig, G. B. 1958, J. Chem. Phys., 28, 751
Woitke, P., & Helling, C. 2004, A&A, 414, 335
Woitke, P., Helling, C., Hunter, G. H., et al. 2018, A&A, 614, A1
Yurchenko, S. N., Barber, R. J., & Tennyson, J. 2011, MNRAS, 413, 1828
Yurchenko, S. N., Tennyson, J., Bailey, J., Hollis, M. D. J., & Tinetti, G. 2014, PNAS, 111, 9379
Zingales, T., Falco, A., Pluriel, W., & Leconte, J. 2022, A&A, 667, A13

Appendix A: Transmission spectra

Transmission spectra simulated with Pytmosph3R ([Falco et al. 2022](#)) for HD 189733 b (top), WASP-121 b (middle) and GJ 1214 b (bottom). We compare here the differences between the east limb, the west limb, and the all planet spectra.
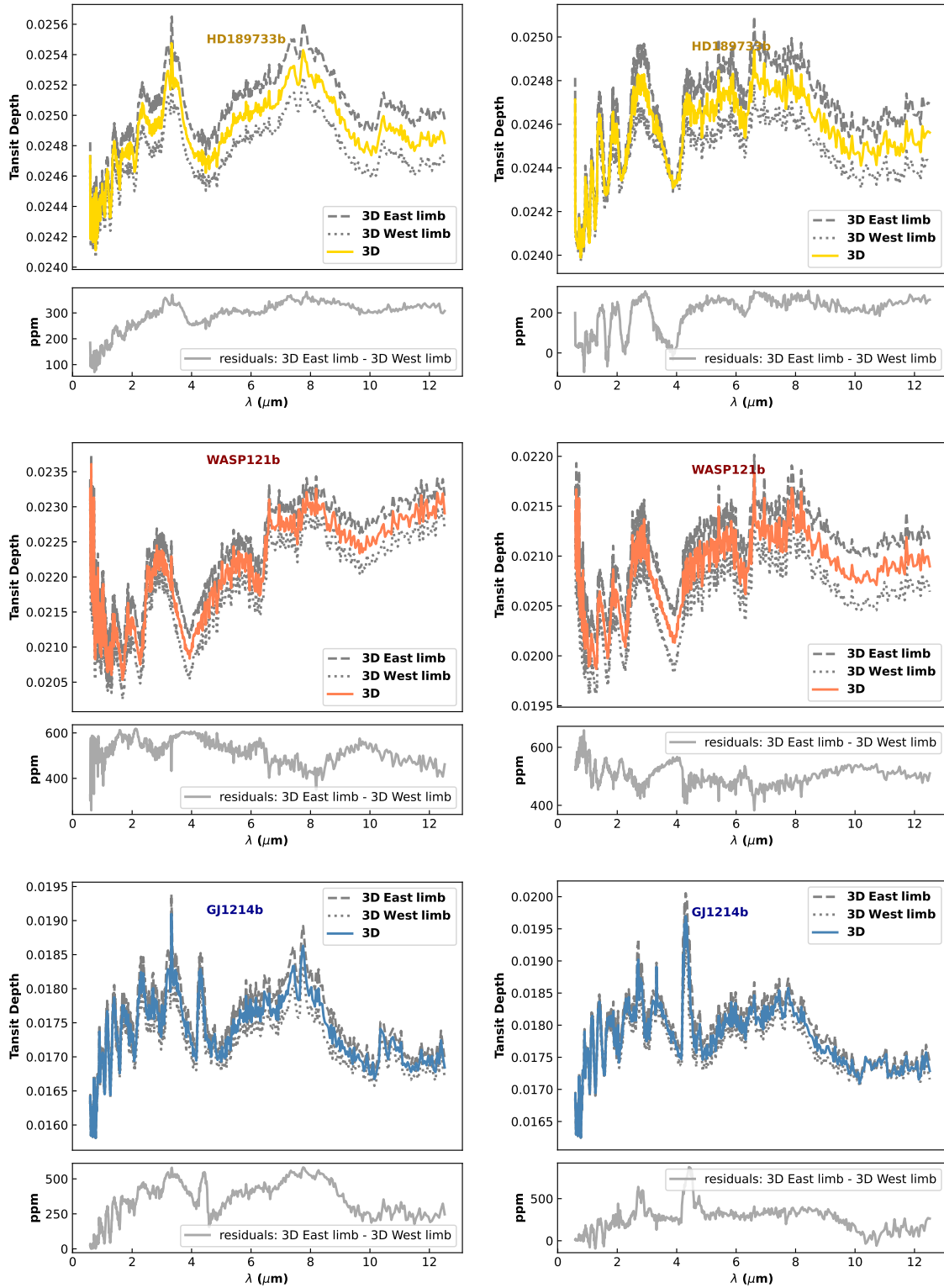

Fig. A.1. Transmission spectra simulated with Pytmosph3R ([Falco et al. 2022](#)) for HD 189733 b (top), WASP-121 b (middle), and GJ 1214 b (bottom). Each panel compares three transmission spectra based on a 3D atmosphere for the east limb, the west limb, and the whole planet, (dashed gray, dotted gray, and solid lines, respectively). Left panels: Constant input chemistry. Right: Equilibrium input chemistry. The differences between the east limb and the west limb spectra are plotted below each panel in gray.

Appendix B: Retrievals results

Tables of values retrieved by all retrieval models for all input configurations considering the best solution for each retrieval are shown here. For Free retrievals, we calculate the metallicity using TauREx 3 which includes it in the available derived parameters, while we compute the C/O ratio using formula 2 from Lee et al. (2013) and the marginalized posterior distributions of CO, CH_4, CO_2, and H_2O from each retrieval (best solution).

Table B.1. Retrieval results of GJ 1214 b. The best retrieval of each configuration is highlighted in bold.

Parameters	Input	JWST 1D constant Free	ACE	Fast	JWST 1D equilibrium Free	ACE	Fast	JWST 3D constant Free	ACE	Fast	JWST 3D equilibrium Free	ACE	Fast	Ariel constant Free	ACE	Fast	Ariel equilibrium Free	ACE	Fast
T_{surf} [K]		970	348	330	770	1037	1097	1092	1303	950	1055	1165	1422	918	1084	1162	1064	1086	1128
T_1 [K]		639	745	875	1035	1345	1341	756	305	306	407	391	317	684	610	325	621	673	533
T_2 [K]		215	236	189	331	378	334	426	560	602	306	580	674	443	569	597	331	470	579
T_{top} [K]		198	628	627	185	212	334	536	306	335	1234	619	532	524	350	367	983	814	833
$\log_{10}(P_1)$ [Pa]		4.67	2.25	2.63	3.95	4.38	4.24	4.98	2.15	2.15	2.44	2.00	2.09	4.72	4.55	2.44	4.59	4.61	4.06
$\log_{10}(P_2)$ [Pa]		0.131	1.58	2.08	2.24	3.13	2.91	3.26	1.20	1.27	-0.648	1.31	1.37	2.60	2.61	1.06	0.32	2.52	1.68
$\log_{10}(H_2O)$	**-1.00**	-1.15			-1.17			-1.09			-1.24			-1.13			-1.26		
$\log_{10}(CO)$	**-2.00**	-2.24			-8.62			-2.11			-4.86			-5.59			-7.48		
$\log_{10}(CH_4)$	**-1.52**	-1.63			-1.21			-1.64			-2.27			-1.61			-2.27		
$\log_{10}(CO_2)$	**-2.00**	-2.10			-9.34			-2.06			-1.57			-1.82			-1.46		
$\log_{10}(HCN)$	**-6.22**	-8.16			-8.35			-8.37			-8.21			-6.93			-7.29		
$\log_{10}(NH_3)$	**-3.30**	-3.45			-3.19			-3.45			-9.23			-3.57			-8.52		
Radius [R_{jup}]	**0.243**	0.242	0.248	0.247	0.241	0.240	0.238	0.242	0.241	0.244	0.245	0.240	0.240	0.242	0.240	0.241	0.247	0.245	0.247
C/O	**0.458**		0.459	0.442	0.90	0.898	0.807		0.255	0.145	0.29	0.441	0.313		0.309	0.169	0.33	0.461	0.371
$\log_{10}(Z)$	**2.00**		2.14	2.20	1.53	1.91	1.96		2.02	1.82	1.58	2.01	1.83		2.06	1.79	1.61	2.06	1.93
logE	**-**	4642	4623	4619	4631	4640	4633	4619	4585	4575	4547	4557	4504	736	737	736	731	740	735

Table B.2. Retrieval results of HD 189733 b. The best retrieval of each configuration is highlighted in bold.

Parameters	Input	JWST 1D constant Free	ACE	Fast	GGchem	JWST 1D equilibrium Free	ACE	Fast	GGchem	JWST 3D constant Free	ACE	Fast	GGchem	JWST 3D equilibrium Free	ACE	Fast	GGchem	Ariel constant Free	ACE	Fast	GGchem	Ariel equilibrium Free	ACE	Fast	GGchem
T_{surf} [K]		1352	1554	508	503	626	1760	836	1610	1313	503	980	899	705	1612	729	1371	1400	1610	1038	1451	1641	1509	1103	1433
T_1 [K]		731	517	969	1643	1382	834	694	732	745	1898	1864	1541	1509	879	735	761	1194	1531	2265	1431	1557	1098	667	1158
T_2 [K]		1140	507	536	677	659	1267	1251	1322	1308	720	688	664	1069	1367	1175	1426	1239	555	587	578	1048	1175	1179	1182
T_{top} [K]		931	760	501	501	2299	535	1264	516	1020	501	502	501	1411	548	1748	522	1311	681	1653	655	1440	1509	1569	1242
$\log_{10}(P_1)$ [Pa]		4.31	0.228	3.53	3.06	2.65	2.68	2.00	2.28	4.18	3.12	2.69	2.90	3.63	2.87	2.01	2.35	4.46	4.12	3.00	4.20	4.49	4.18	3.29	4.58
$\log_{10}(P_2)$ [Pa]		2.52	-2.63	0.552	2.63	2.42	2.39	1.59	2.12	2.13	2.85	2.47	2.20	3.27	2.60	1.91	2.22	1.61	1.03	2.56	1.22	1.33	1.51	1.21	1.64
$\log_{10}(H_2O)$	**-3.15**	-2.87				-3.19				-2.90				-3.44				-3.05				-3.71			
$\log_{10}(CO)$	**-4.00**	-7.86				-3.12				-7.92				-3.30				-7.92				-3.92			
$\log_{10}(CH_4)$	**-3.40**	-3.04				-5.34				-3.05				-6.17				-3.28				-7.77			
$\log_{10}(CO_2)$	**-7.70**	-9.16				-5.89				-9.15				-6.58				-9.37				-7.05			
$\log_{10}(HCN)$	**-7.00**	-9.19				-9.32				-9.09				-7.99				-8.41				-9.11			
$\log_{10}(NH_3)$	**-4.52**	-4.29				-6.53				-4.34				-6.30				-5.49				-8.89			
$\log_{10}(FeH)$	**-8.05**	-7.93				-11.8				-7.91				-10.8				-9.43				-10.2			
$\log_{10}(SiO)$	**-4.70**	-8.55				-9.52				-8.44				-9.37				-7.59				-8.22			
$\log_{10}(Na)$	**-5.52**	-8.00				-8.73				-11.7				-9.18				-7.06				-6.80			
$\log_{10}(K)$	**-6.70**	-11.7				-11.9				-7.75				-11.9				-7.36				-8.28			
$\log_{10}(TiO)$	**-10.0**	-9.75				-11.9				-12.0				-11.9				-9.63				-11.4			
$\log_{10}(VO)$	**-9.00**	-8.68				-11.9				-8.70				-11.9				-9.97				-11.2			
Radius [R_{jup}]	**1.13**	1.13	1.13	1.13	1.13	1.16	1.12	1.13	1.12	1.12	1.13	1.13	1.13	1.12	1.12	1.13	1.12	1.12	1.13	1.12	1.13	1.12	1.12	1.12	1.12
C/O	**0.550**		0.962	0.726	0.797	0.54	0.677	0.0377	0.418		0.921	2.00	0.874	0.58	0.614	0.0368	0.308		0.954	1.87	0.869	0.38	0.676	0.143	0.544
$\log_{10}(Z)$	**0.00**		0.00139	0.417	-0.047	0.02	0.0864	1.22	0.359		-0.444	-0.989	0.263	-0.18	0.144	1.34	0.541		0.0480	-0.858	-0.163	-0.44	0.150	1.03	0.0782
logE	**-**	5818	3282	-1089	3146	5545	5747	3390	5365	5792	2744	-3708	2102	5541	5749	3253	5354	843	809	791	809	840	848	840	848

Table B.3. Retrieval results of WASP-121 b. The best retrieval of each configuration is highlighted in bold.

Parameters	Input	JWST												Ariel					
		1D						3D											
		constant			equilibrium			constant			equilibrium			constant			equilibrium		
		Free	Fast	GGchem	Free	Fast	GGchem	Free	Fast	GGchem	Free	Fast	GGchem	Free	Fast	GGchem	Free	Fast	GGchem
T_{surf} [K]		2045	3206	2122	2422	2590	1830	880	3342	475	2736	928	3277	1805	1273	3473	2372	1604	1558
T_1 [K]		1519	641	1652	3446	586	2828	908	709	465	1282	521	1081	614	622	779	1265	836	1151
T_2 [K]		975	2290	776	978	2658	766	3745	2704	3016	2195	2162	2394	3671	2799	2554	2048	2186	2143
T_{top} [K]		2877	596	2714	2737	468	3596	3727	784	466	705	1269	710	3003	558	837	2373	844	2664
$\log_{10}(P_1)$ [Pa]		4.78	2.01	2.53	3.63	2.02	3.10	2.05	2.51	3.52	2.40	3.81	2.10	2.18	2.84	2.46	3.78	4.30	3.11
$\log_{10}(P_2)$ [Pa]		3.22	0.31	2.43	3.21	-0.616	2.37	1.68	2.40	3.47	1.56	3.28	1.57	1.33	2.67	1.98	1.30	3.45	0.989
$\log_{10}(H_2O)$	**-3.30**	-3.08			-3.16			-3.62			-4.12			-3.58			-4.05		
$\log_{10}(CO)$	**-3.30**	-3.24			-3.18			-3.97			-3.14			-4.07			-3.16		
$\log_{10}(CH_4)$	**-5.00**	-4.71			-9.31			-9.60			-9.81			-9.24			-9.50		
$\log_{10}(CO_2)$	**-7.40**	-7.98			-6.73			-9.36			-7.15			-9.37			-7.69		
$\log_{10}(HCN)$	**-6.40**	-8.52			-9.29			-9.23			-9.86			-8.87			-9.35		
$\log_{10}(NH_3)$	**-5.22**	-4.99			-6.89			-9.09			-9.64			-9.23			-9.29		
$\log_{10}(FeH)$	**-5.70**	-5.46			-9.55			-6.28			-9.42			-6.82			-10.6		
$\log_{10}(SiO)$	**-4.22**	-4.22			-5.48			-6.12			-5.53			-8.11			-7.41		
$\log_{10}(Na)$	**-5.52**	-1.80			-8.96			-5.28			-7.22			-3.64			-3.66		
$\log_{10}(K)$	**-6.70**	-9.58			-10.9			-11.2			-10.6			-10.1			-9.39		
$\log_{10}(TiO)$	**-7.00**	-6.75			-7.41			-7.25			-7.93			-6.97			-7.62		
$\log_{10}(VO)$	**-9.00**	-8.90			-11.2			-11.5			-9.36			-11.0			-10.9		
Radius [R_{jup}]	**1.91**	1.95	1.92	1.94	1.89	1.93	1.89	1.98	1.95	2.03	1.92	1.97	1.91	1.98	1.99	1.94	1.94	1.96	1.96
C/O	**0.550**		0.772	0.897	0.48	0.0365	0.114		0.862	0.883	0.90	0.792	0.266		0.924	0.669	0.88	0.883	0.523
$\log_{10}(Z)$	**0.00**		-0.222	0.132	0.02	1.23	0.800		-0.214	-0.804	-0.14	-0.708	-0.992		0.169	-0.308	0.20	-0.678	-0.614
logE	-	4938	4201	4504	4851	4541	4922	4396	2152	3086	4872	4550	4820	641	566	581	761	749	768

Appendix C: Temperature profiles

The input temperature profiles at the equator using the 1D and 3D thermal structure, overplotted with the temperature profiles of all the retrieval models are shown here. For the homogeneous 1D thermal structure model, longitude profiles are on top of each other since it has been build homogeneously. For the 3D models, limb profiles (90° and 270° longitude) are in bold with dot markers.
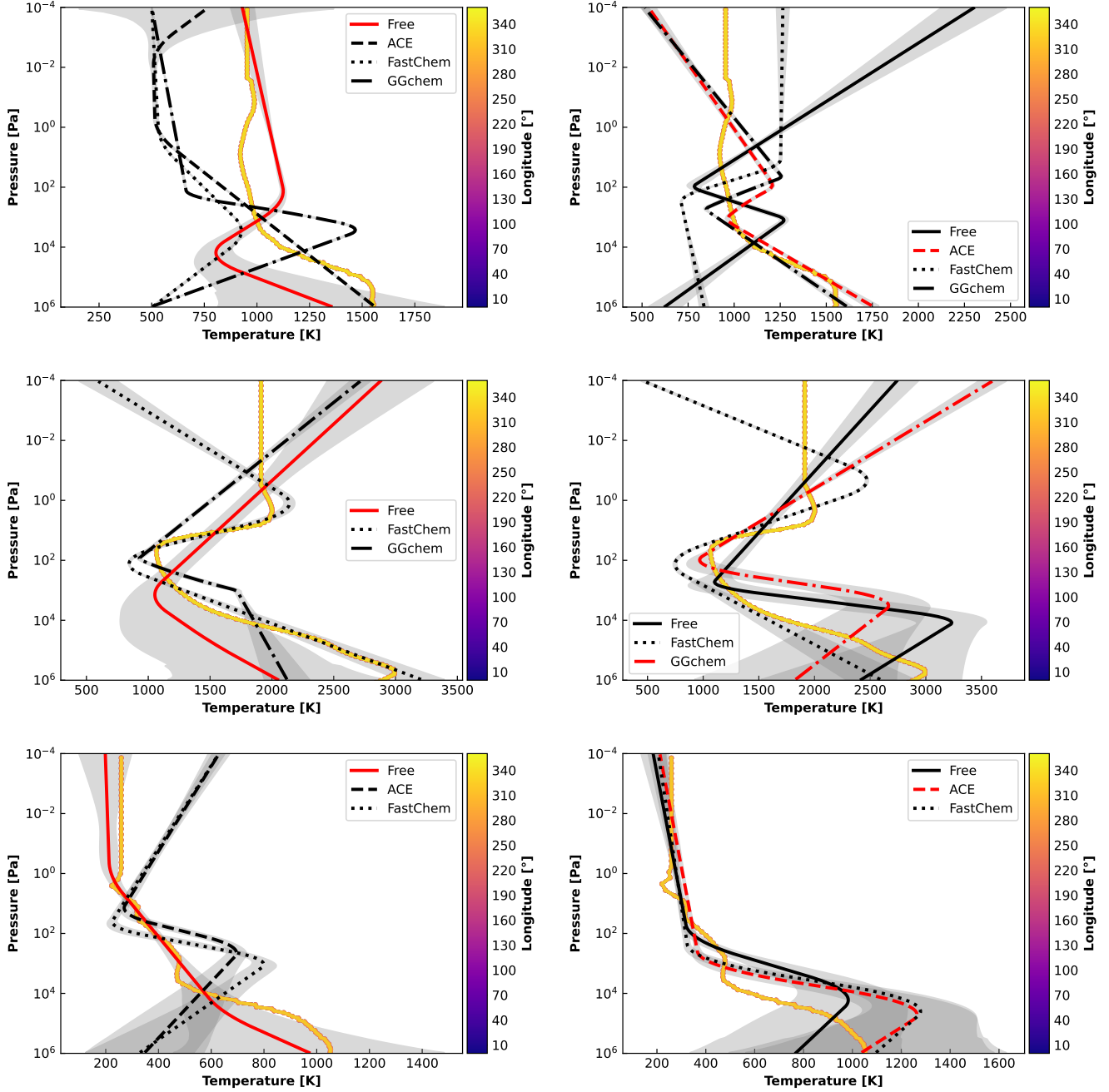

Fig. C.1. JWST simulations. Top to Bottom: Temperature pressure profiles at the equator assuming 1D atmospheres for HD 189733 b, WASP-121 b, and GJ 1214 b. Substellar point at 180° longitude. Left panels: Constant input chemistry. Right: Equilibrium input chemistry. Overplotted are the best TP-profiles of Free, ACE, FastChem, and GGchem retrievals (black solid, dashed, dotted, and dash-dotted lines, respectively). The retrieval having the highest Bayes factor is plotted in red.
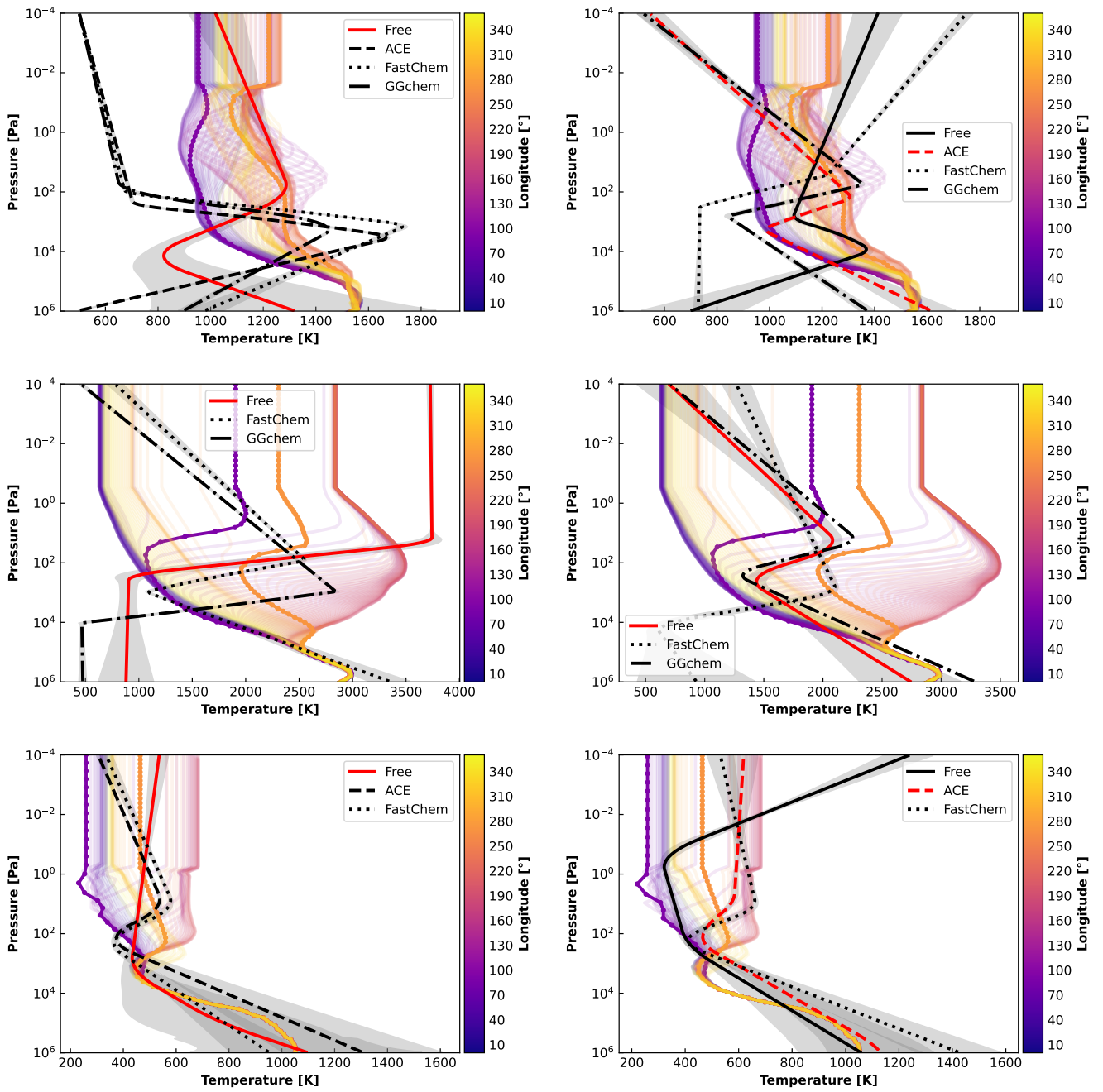
Fig. C.2. Same as Figure C.1, but using 3D thermal structure for the atmospheres.
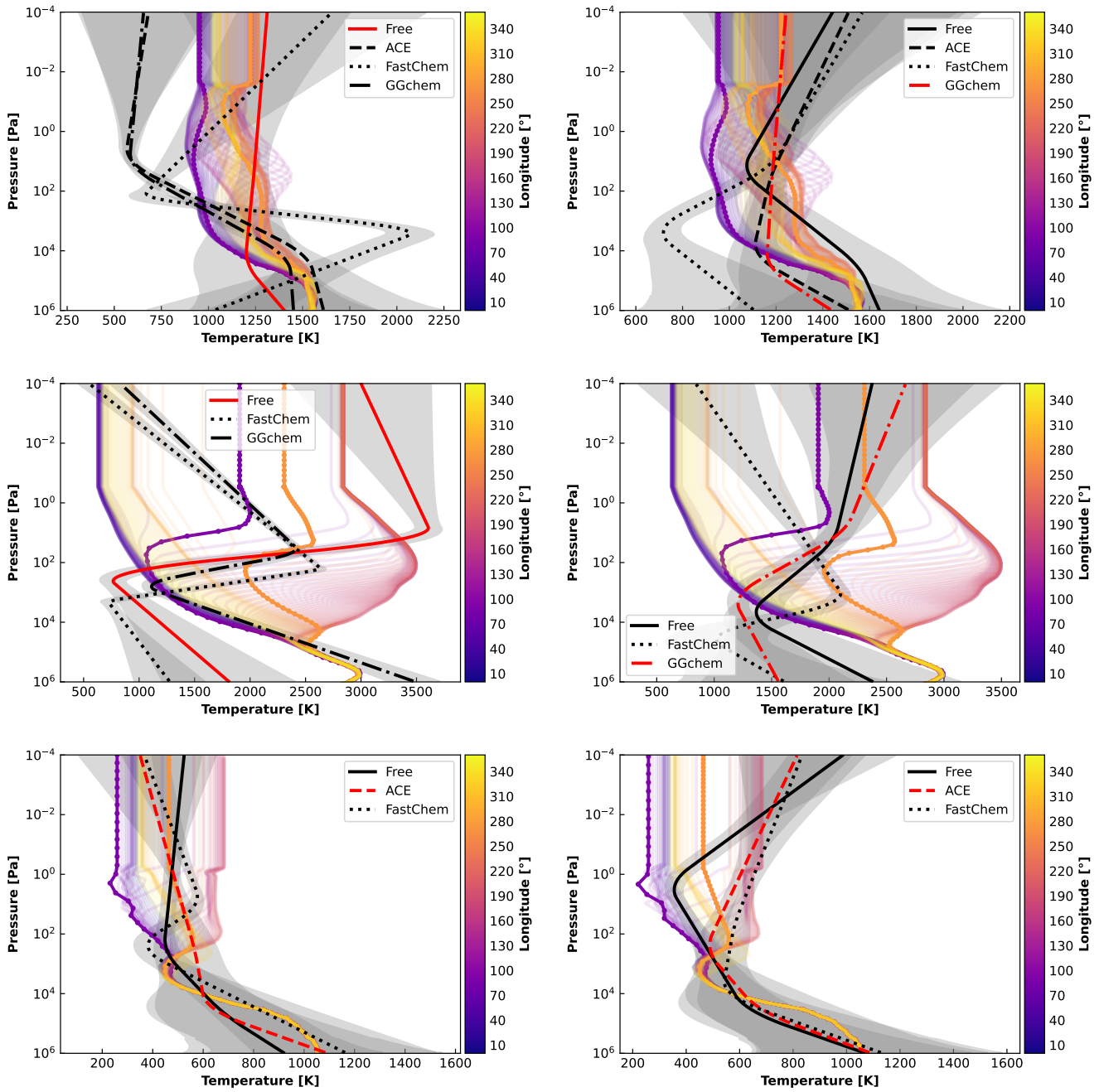
Fig. C.3. Same as Figure C.2, but for Ariel simulations.

Appendix D: Abundances profiles

The input abundances profiles at the equator using the 1D and 3D thermal structure, overplotted with the abundances profiles of all the retrieval models are shown here. For the homogeneous 1D thermal structure model, longitude profiles are on top of each other since it has been build homogeneously.
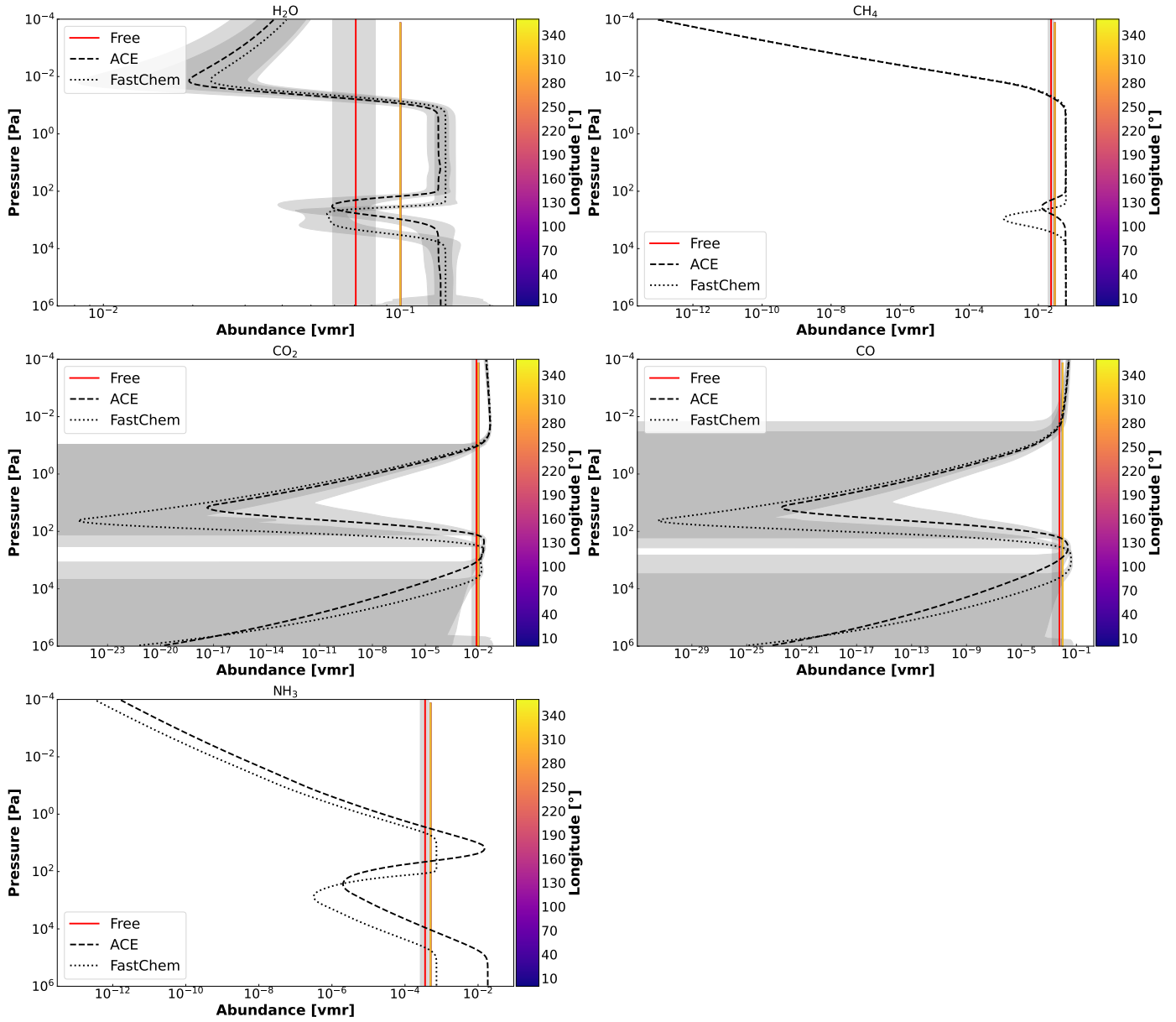
Fig. D.1. JWST simulation with constant chemistry assuming 1D atmosphere for the thermal structure. Shown are the main absorbers volume mixing ratio profiles at the equator on GJ 1214 b. Substellar point at 180° longitude. Overplotted are the best species profiles of Free, ACE, FastChem and GGchem retrievals (black solid, dashed, dotted, and dash-dotted lines, respectively). The retrieval having the highest Bayes factor is plotted in red.
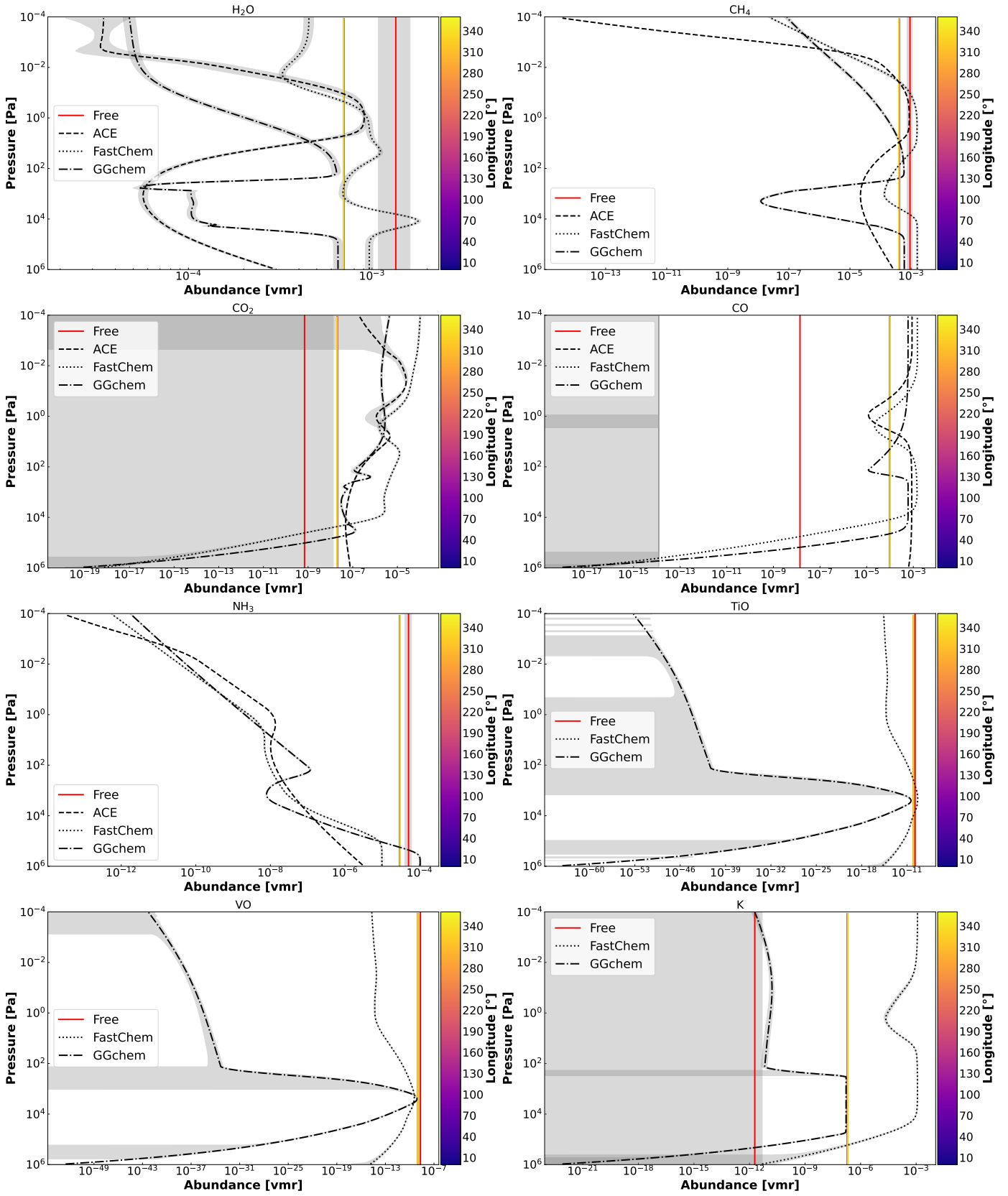
Fig. D.2. Same as Figure D.1, but for HD 189733 b.
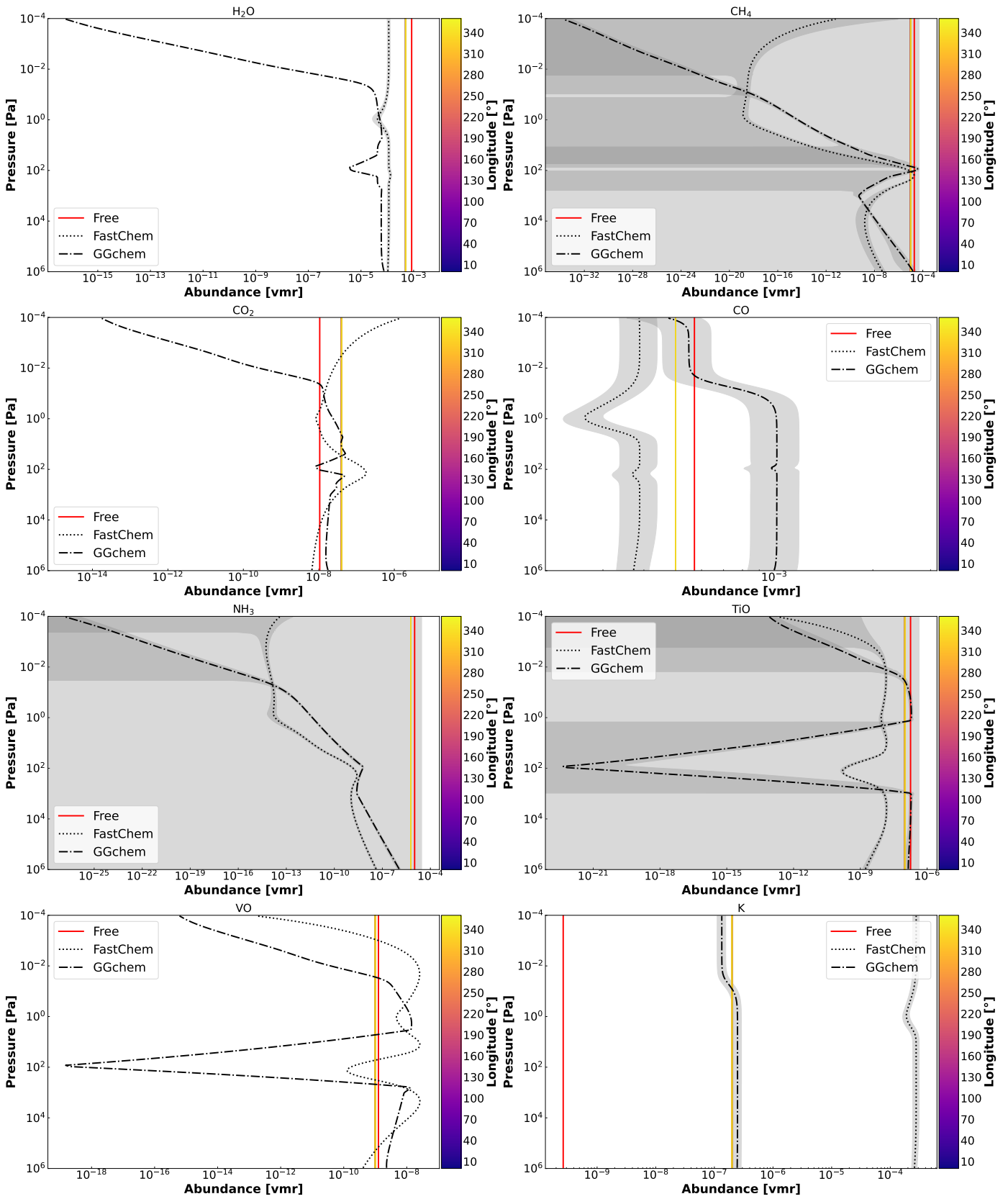
Fig. D.3. Same as Figure D.1, but for WASP-121 b.
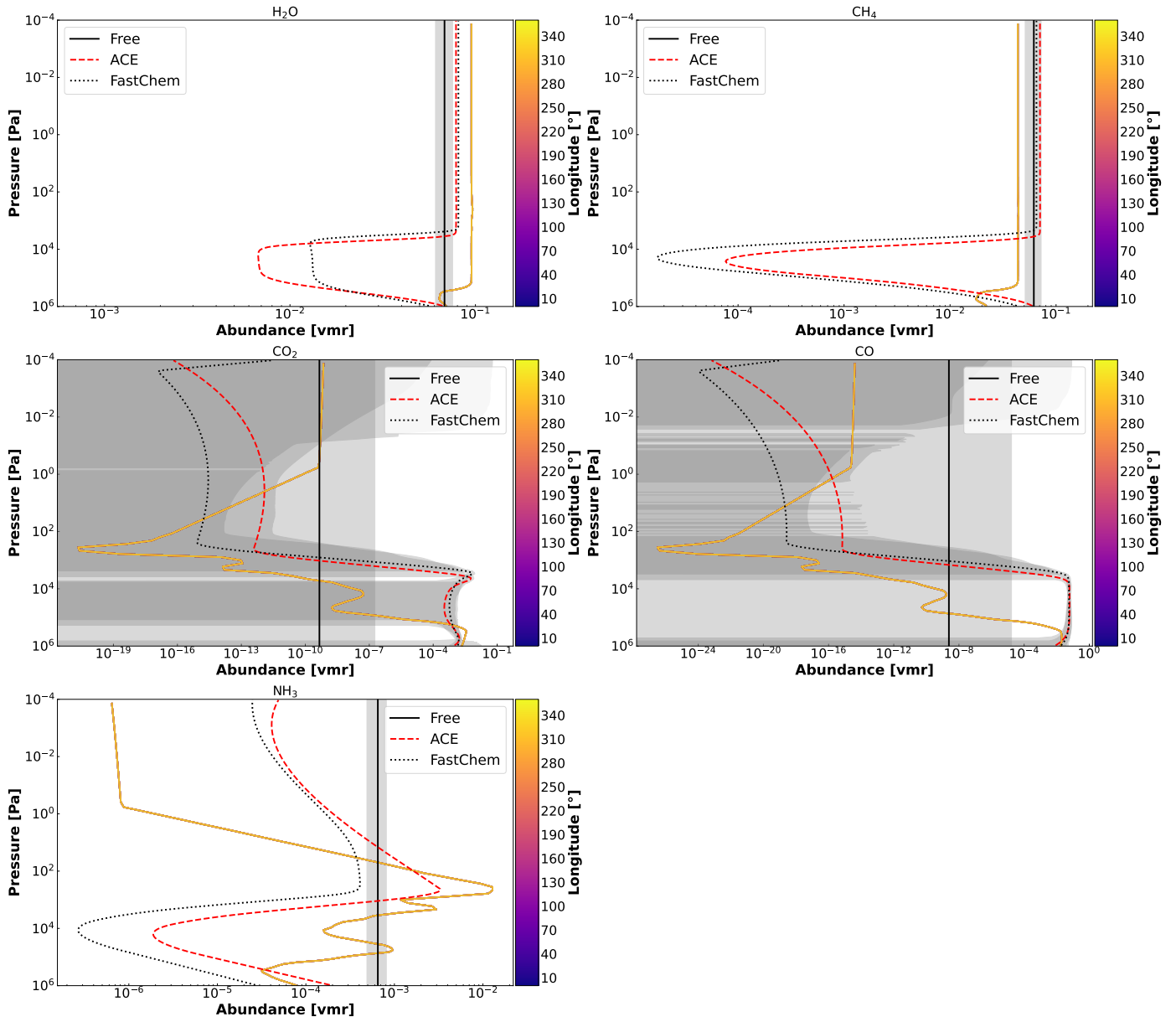
Fig. D.4. Same as Figure D.1, but for GJ 1214 b with input equilibrium chemistry.
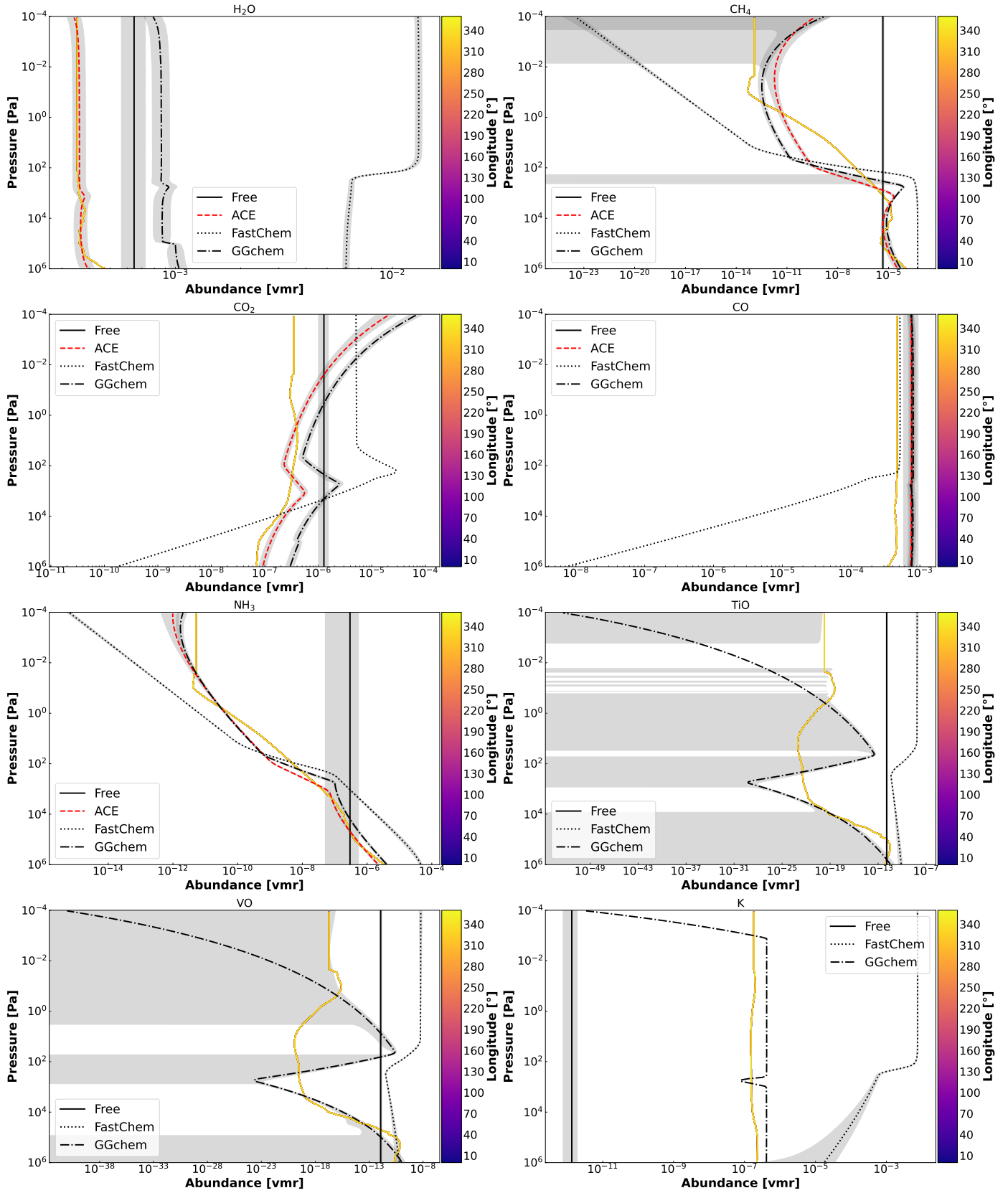
Fig. D.5. Same as Figure D.1, but for HD 189733 b with input equilibrium chemistry.
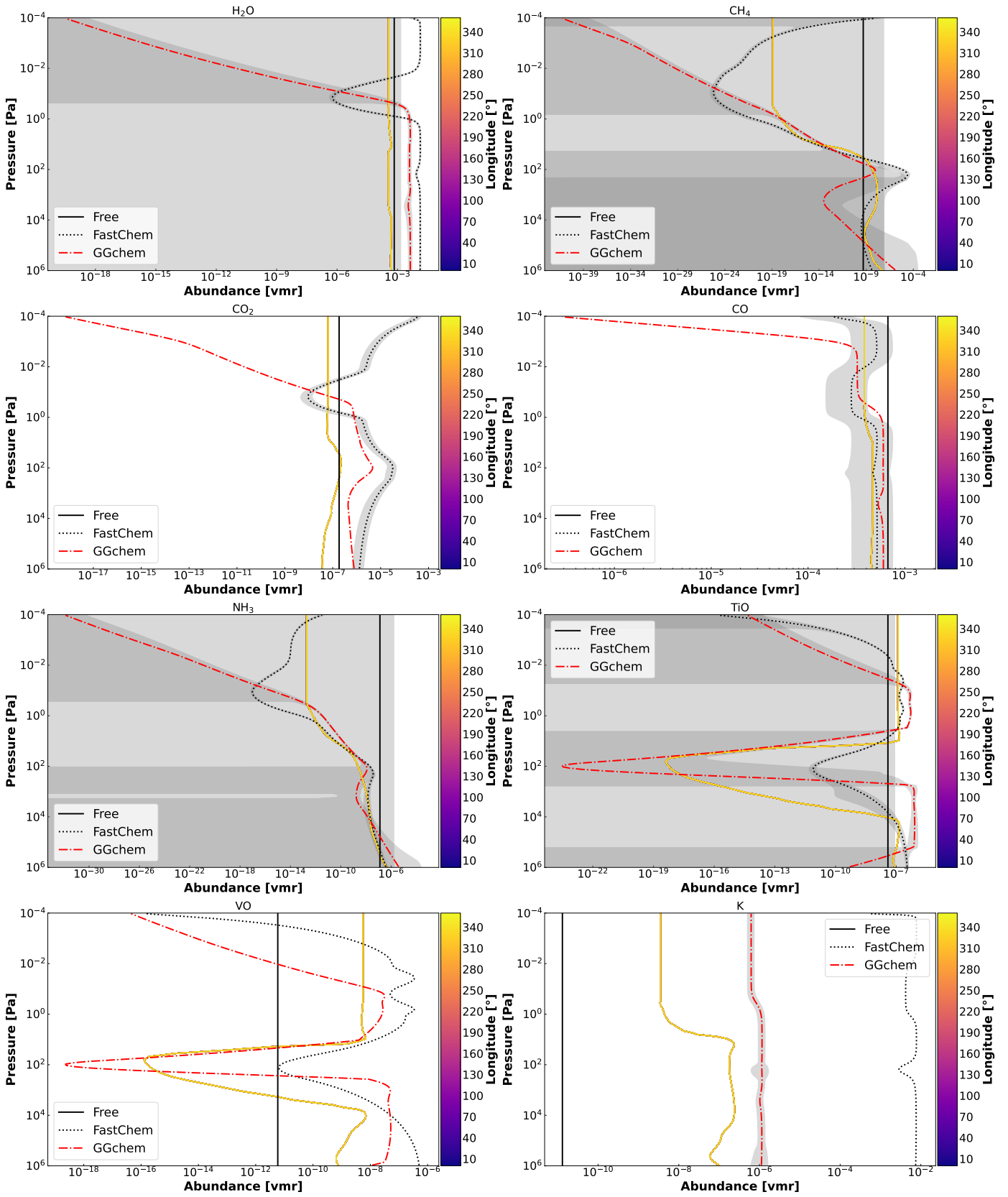
Fig. D.6. Same as Figure D.1, but for WASP-121 b with input equilibrium chemistry.
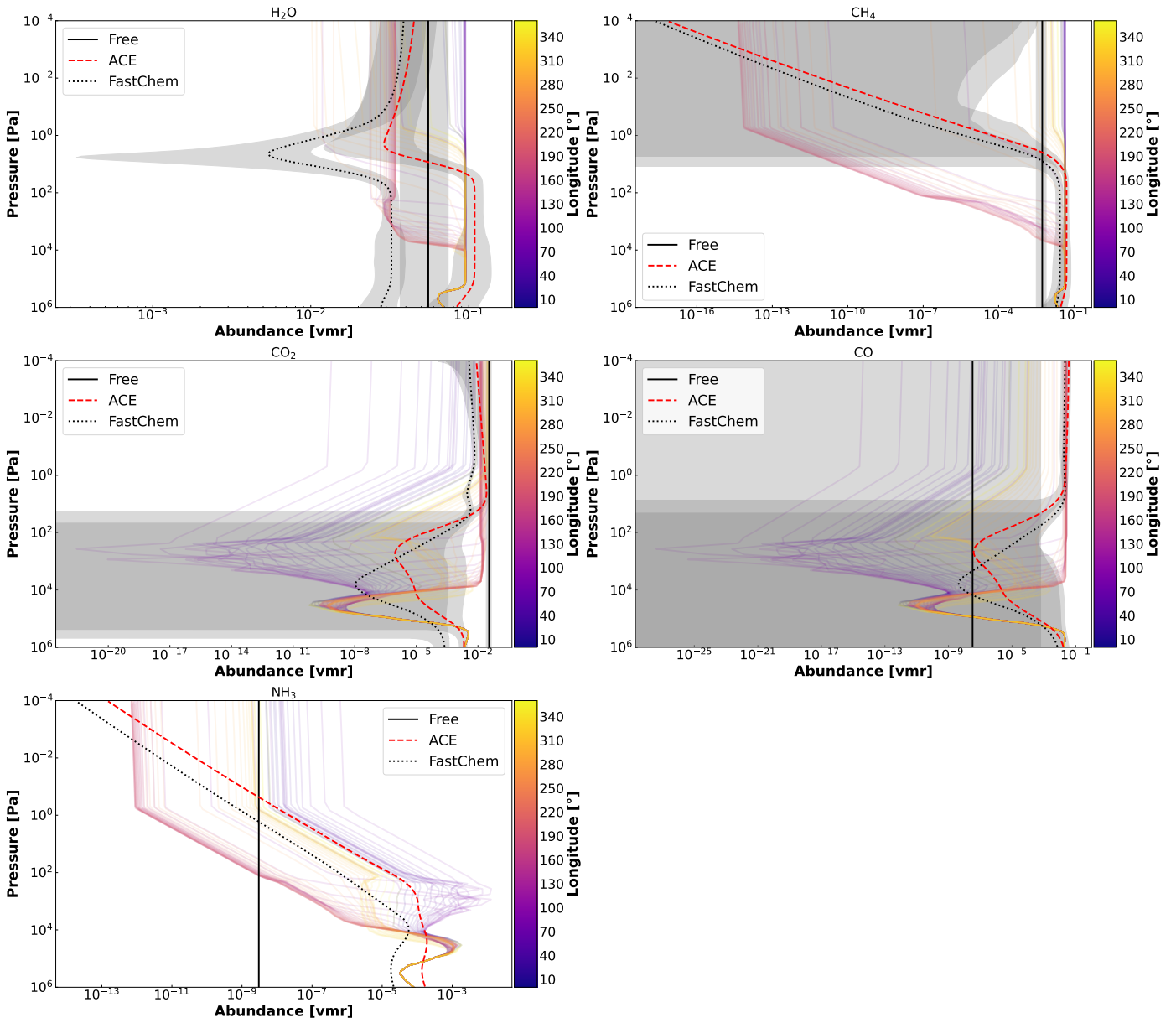
Fig. D.7. Same as Figure D.1, but for GJ 1214 b with input equilibrium chemistry and with Ariel simulation assuming a 3D atmosphere for the thermal structure.
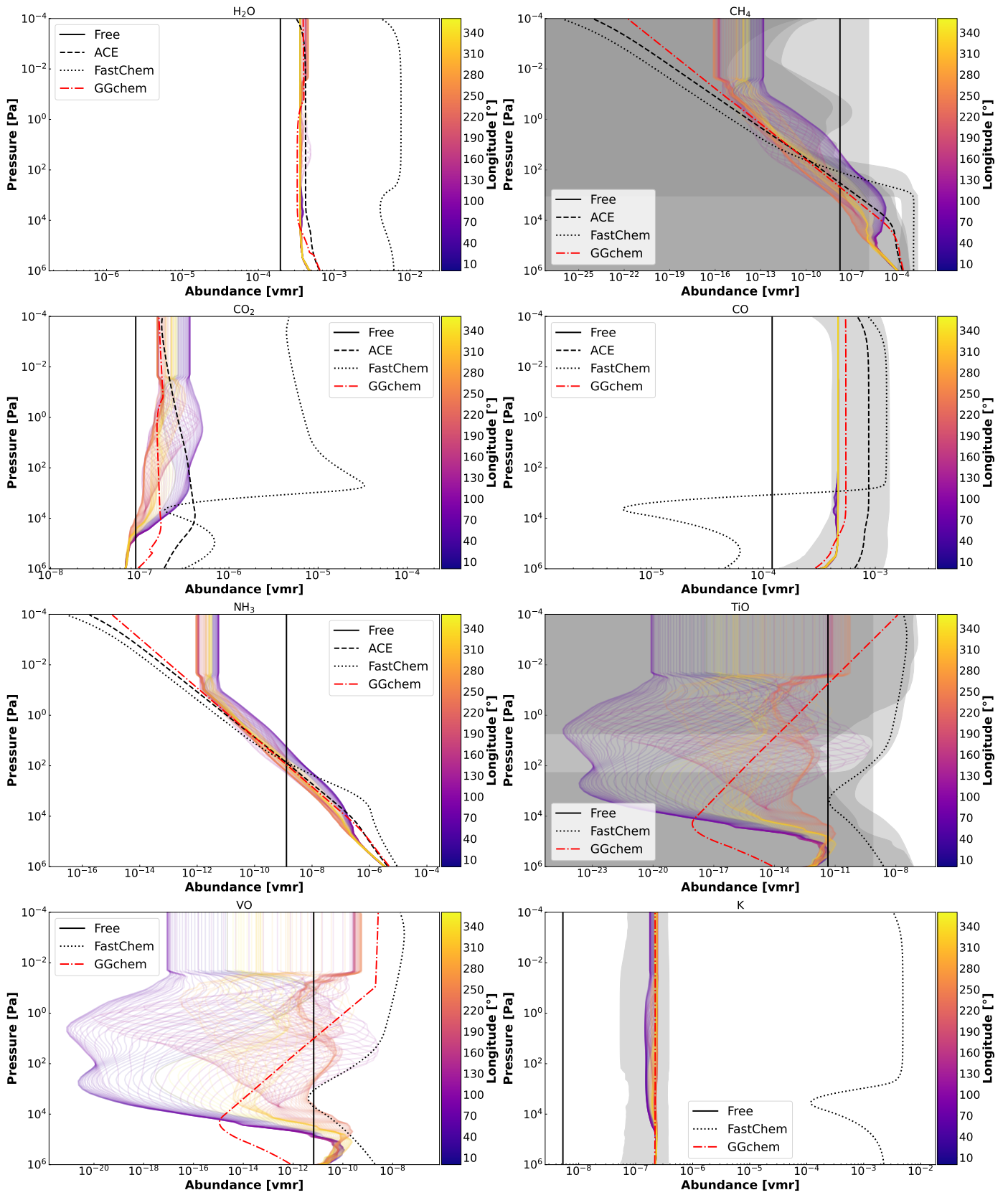
Fig. D.8. Same as Figure D.1, but for HD 189733 b with input equilibrium chemistry and with Ariel simulation assuming a 3D atmosphere for the thermal structure.
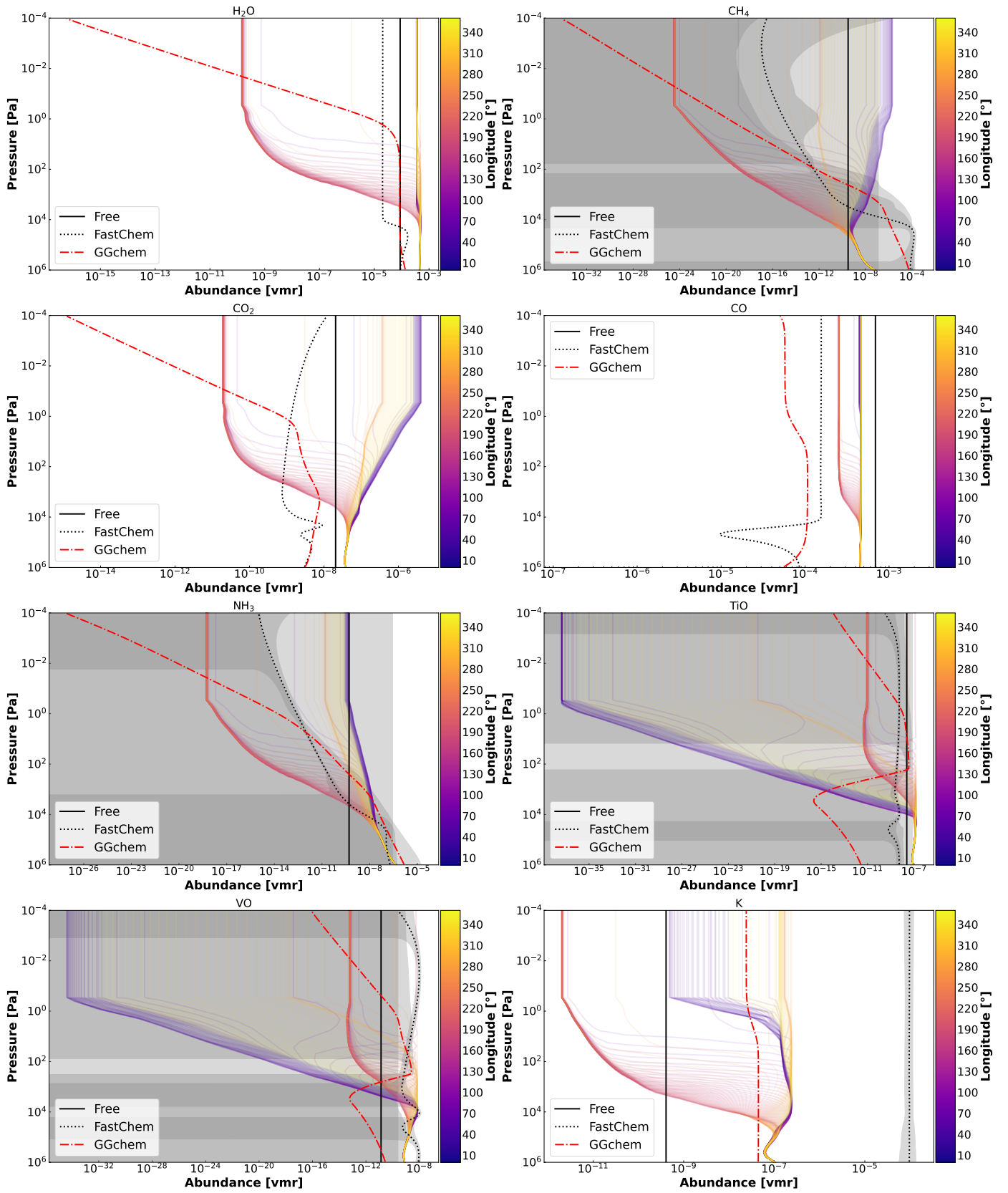
Fig. D.9. Same as Figure D.1, but for WASP-121 b with input equilibrium chemistry and with Ariel simulation assuming a 3D atmosphere for the thermal structure.

Appendix E: Corner plots

The corner plot of the retrievals, overplotted with the input C/O ratio, metallicity (Z) and radius of the planet (R_p) for equilibrium chemistry are shown here.
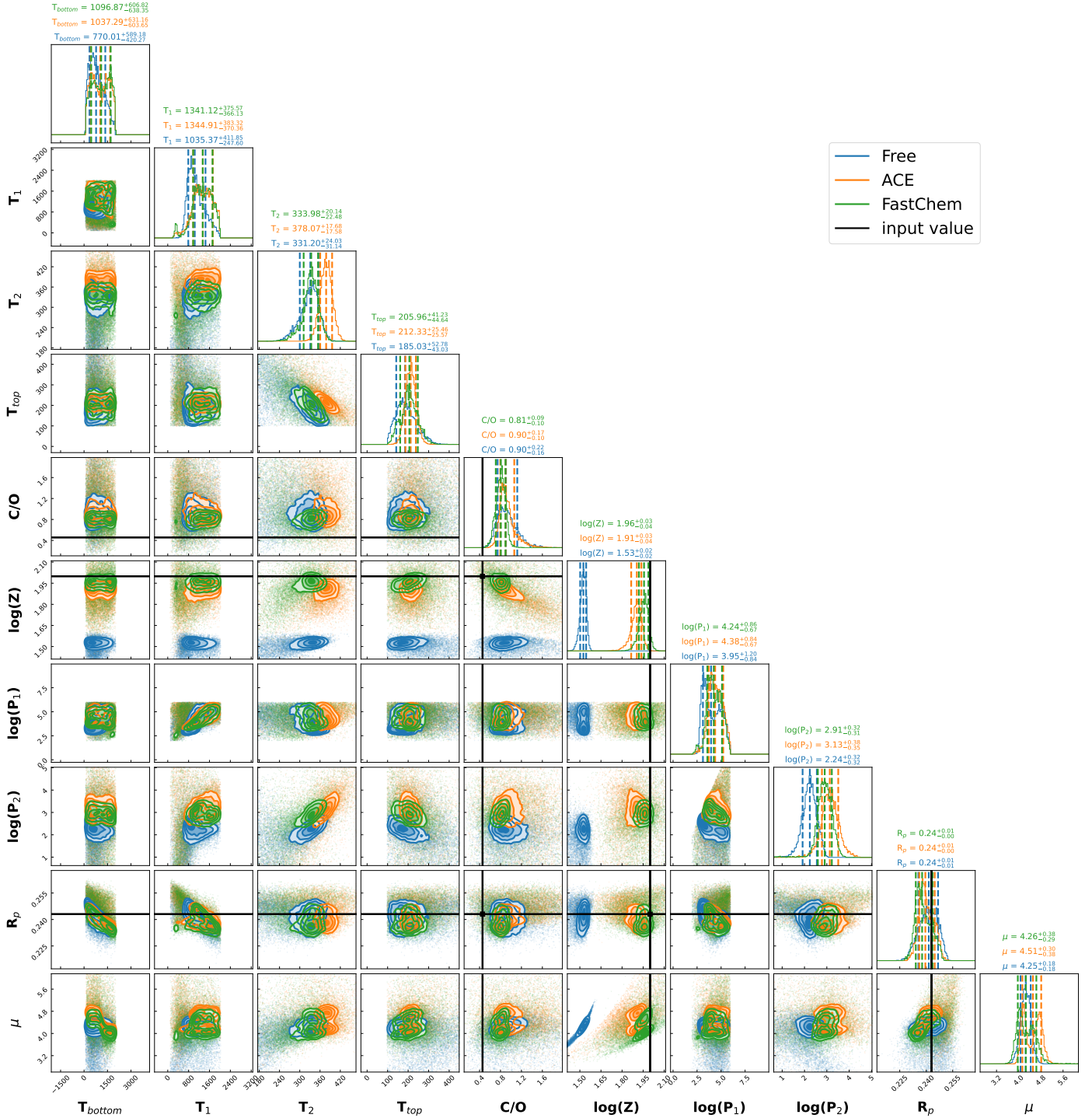
Fig. E.1. JWST simulation with equilibrium chemistry on GJ 1214 b assuming a 1D atmosphere for the thermal structure. Overplotted are Free, ACE, FastChem, and/or GGchem retrievals (blue, orange, green, and red, respectively). The input value is indicated by the black line for the input C/O ratio, metallicity (Z), and radius of the planet (R_p) values.
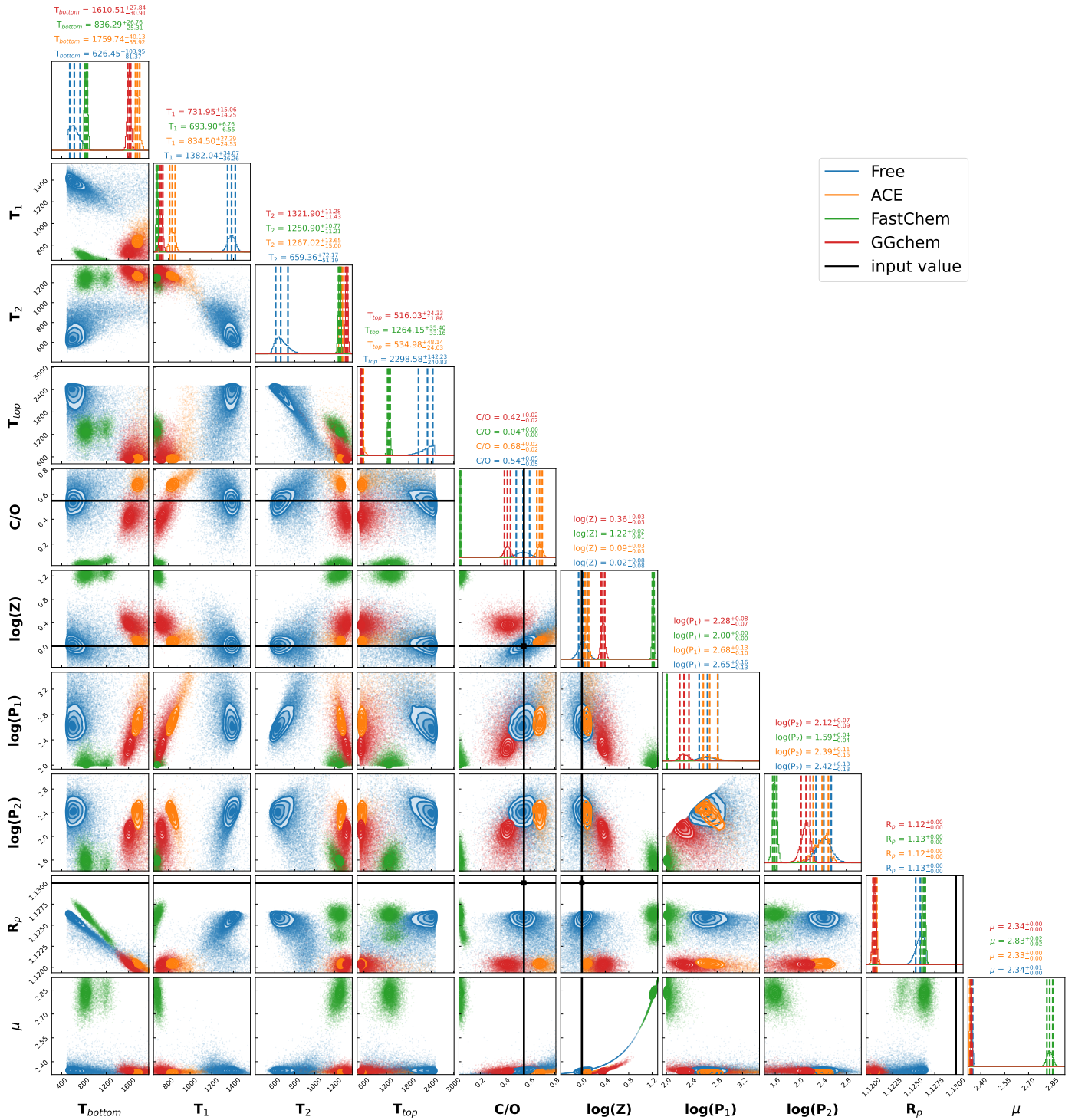
Fig. E.2. Same as Figure E.1, but for HD 189733 b.
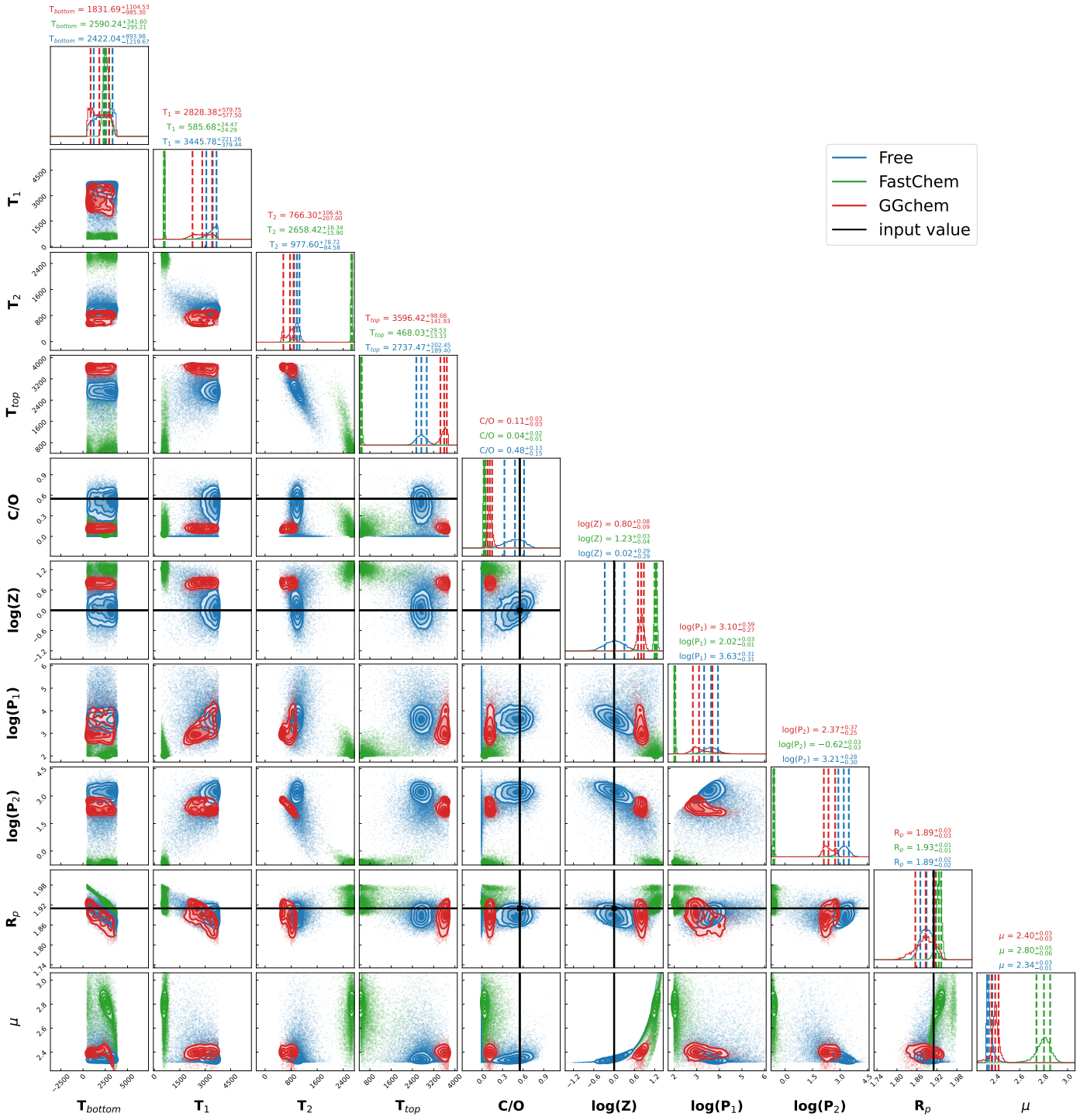
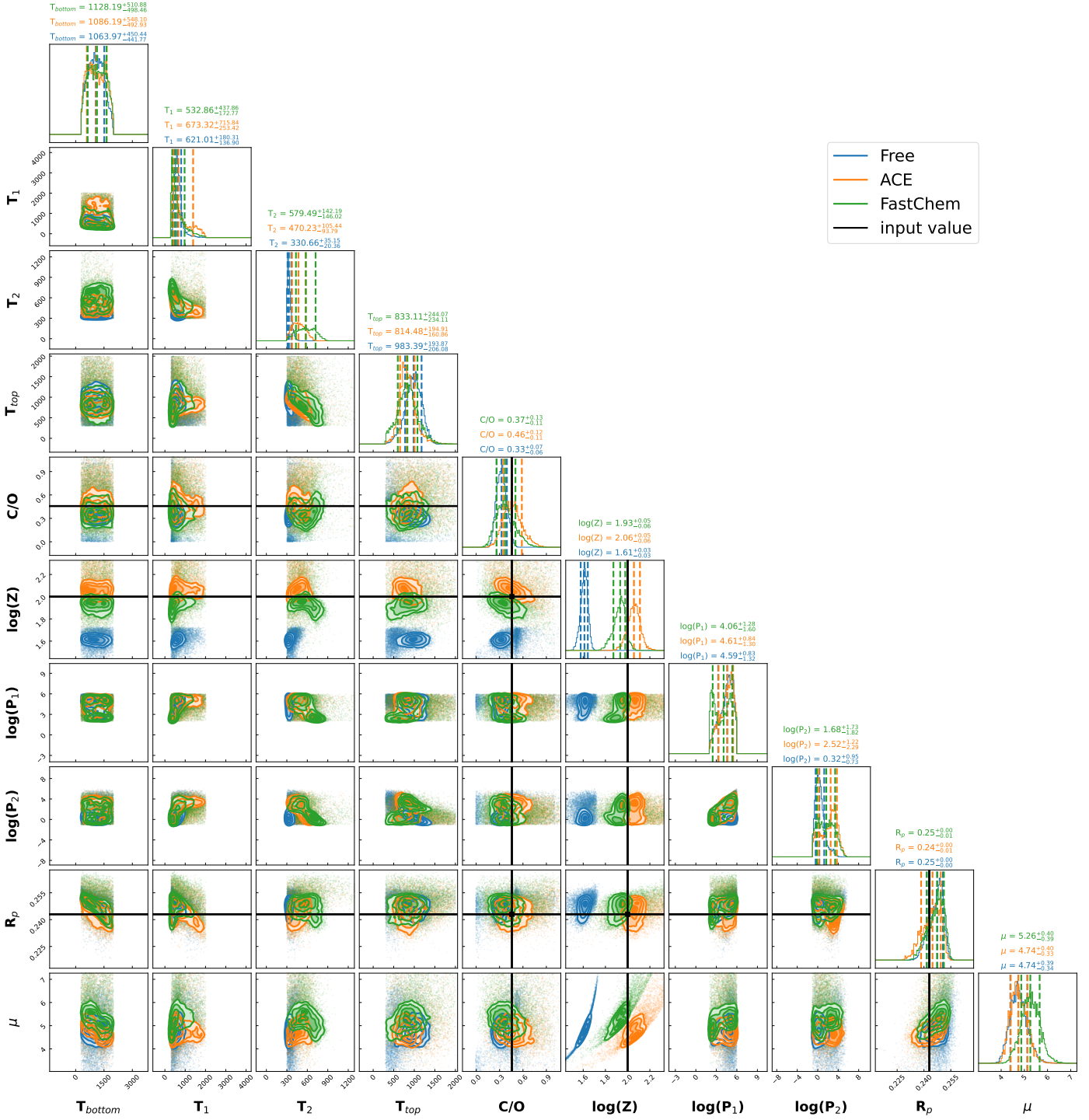**Fig. E.3.** Same as Figure E.1, but for WASP121 b.

Fig. E.4. Same as Figure E.1, but for GJ 1214 b with Ariel simulation assuming a 3D atmosphere for the thermal structure.
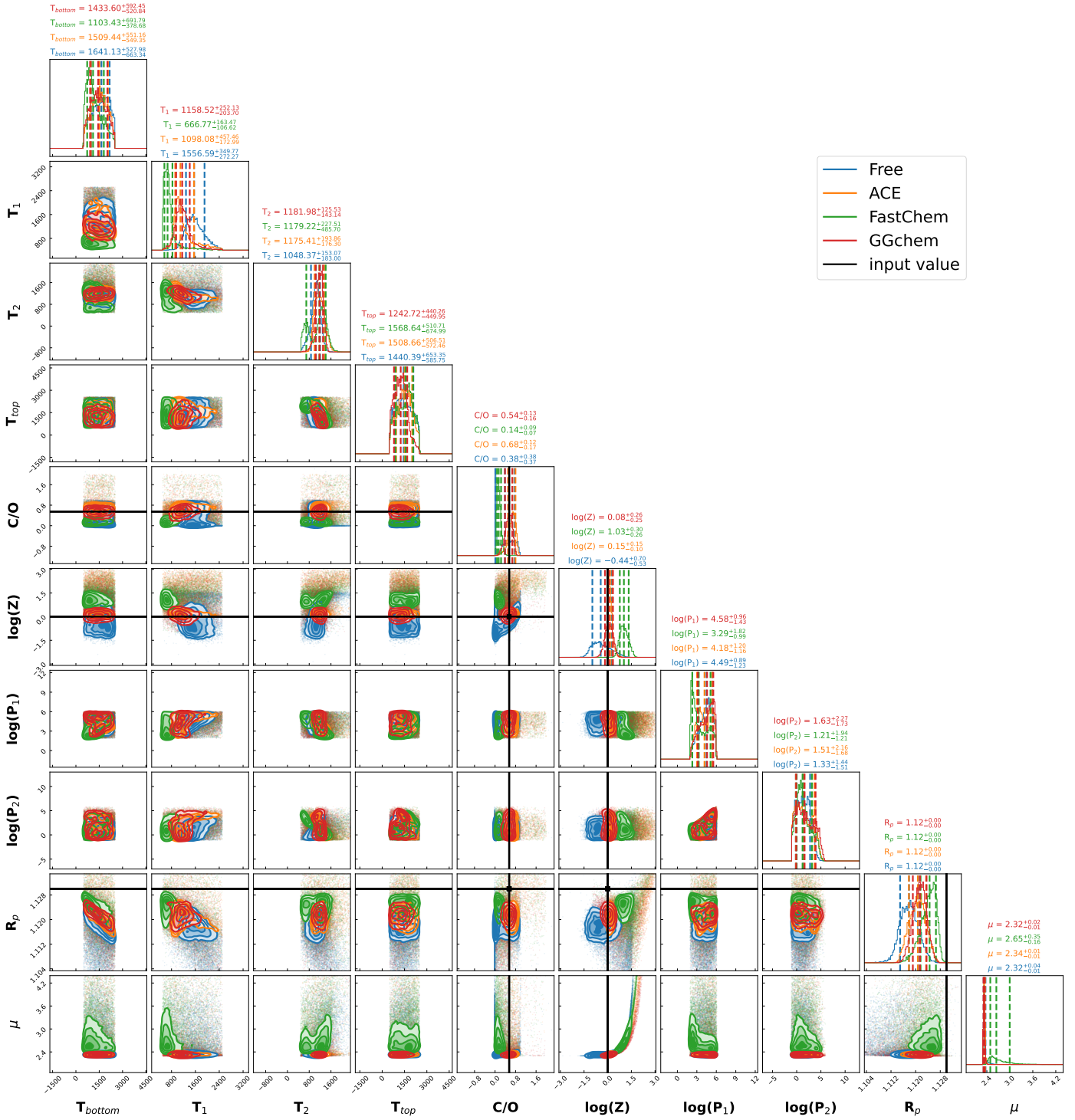
Fig. E.5. Same as Figure E.1, but for HD 189733 b with Ariel simulation assuming a 3D atmosphere for the thermal structure.

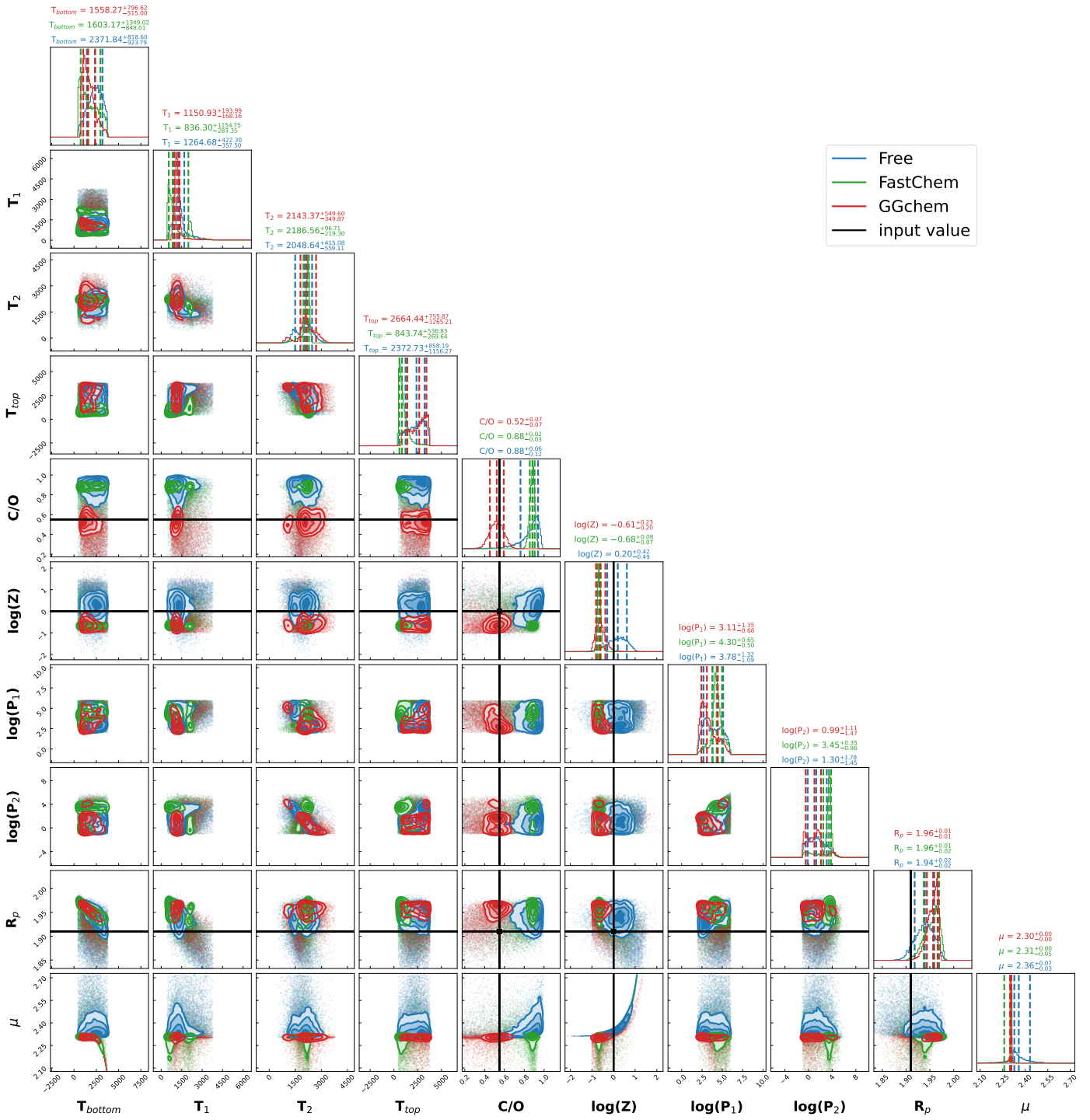**Fig. E.6.** Same as Figure E.1, but for WASP121 b with Ariel simulation assuming a 3D atmosphere for the thermal structure.