

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 21 Novembre 2022 par

Lorena ACUÑA AGUIRRE

Modélisation des populations des planètes de faible masse:
transition entre super-terres et mini-Neptunes

Discipline

Physique et Sciences de la Matière

Spécialité

Astrophysique et Cosmologie

École doctorale

ED 352 - Physique et Sciences de la Matière

Laboratoire/Partenaires de recherche

Laboratoire d'Astrophysique de Marseille

Composition du jury

•	Giovanna TINETTI	Rapporteure
•	University College London	
•	Pierre DROSSART	Rapporteur
•	IAP/PSL/LESIA	
•	Emmanuel MARCQ	Examineur
•	UVSQ/LATMOS	
•	Heike RAUER	Examinatrice
•	Freie Universität Berlin	
•	Jérémy LECONTE	Examineur
•	Laboratoire d'Astrophysique de Bordeaux	
•	Jean-Luc BEUZIT	Président du jury
•	LAM	
•	Magali DELEUIL	Directrice de thèse
•	LAM	
•	Olivier MOUSIS	co-Directeur de thèse
•	LAM	


Invité

Laura KREIDBERG
Max Planck Institute for Astronomy

I, undersigned, Lorena Acuña Aguirre, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific direction of Magali Deleuil and Olivier Mouis, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the french national charter for Research Integrity and the Aix-Marseille University charter on the fight against plagiarism.

This work has not been submitted previously either in this country or in another country in the same or in a similar version to any other examination body.

Marseille, 30 August 2022

A handwritten signature in black ink that reads "Lorena Acuña". The signature is written in a cursive style with a long horizontal line above the name and a shorter one below it.

Cette œuvre est mise à disposition selon les termes de la [Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Résumé

Les exoplanètes de faible masse montrent une diversité dans leurs densités et leurs conditions d'irradiation, allant de planètes fortement irradiées à des mondes plus tempérés. Pour estimer leur composition, nous avons développé un modèle de structure interne qui inclut de manière autoconsistante une atmosphère en équilibre radiatif-convectif. Le modèle peut reproduire des enveloppes d'eau et de CO₂ dans les conditions extrêmes des exoplanètes hautement irradiées. Il peut être utilisé comme modèle direct pour générer des relations masse-rayon, ainsi que dans le cadre de notre propre modèle bayésien adaptatif de type Markov Chain Monte Carlo, ce qui permet d'estimer les incertitudes des paramètres de composition de la planète compte tenu des barres d'erreur de la masse et du rayon observés. Nous présentons les différents développements que nous avons réalisés pour aboutir à l'implémentation cohérente du modèle couplé d'intérieur et d'atmosphère et à sa validation.

Le modèle permet d'estimer des paramètres de composition avec leurs incertitudes respectives à partir de la mesure de la masse et du rayon d'une planète donnée. Une fois la composition estimée, il permet aussi d'évaluer l'observabilité des spectres d'émission des exoplanètes rocheuses avec le télescope spatial James Webb. Dans le cas de systèmes multiplanétaires, l'analyse des compositions des différentes planètes fournit des contraintes quant à leurs mécanismes de formation. Nous discutons des propriétés que l'on peut déduire des compositions du large échantillon de 46 planètes que nous avons analysé tout au cours de cette thèse, ainsi que des limitations dues aux incertitudes de mesure auxquelles le modèle est confronté. Notre modèle peut servir de précédent à un cadre de modélisation construit pour interpréter les données fournies par les missions à venir telles que PLATO et Ariel. Notre modèle pourra également servir à interpréter les données fournies par les missions à venir telles que PLATO et Ariel et préparer des analyses atmosphériques plus fouillées.

Mots clés: Exoplanètes. Intérieurs et atmosphères planétaires. Modélisation.

Abstract

Low-mass exoplanets are showing a diversity in their densities and irradiation conditions, ranging from highly irradiated planets to temperate worlds. To estimate their composition, we have developed an interior structure model that includes self-consistently an atmosphere in radiative-convective equilibrium. The model can reproduce water and CO₂ envelopes under the extreme conditions of highly-irradiated exoplanets, which are the most numerous population. It can be used as a forward model to generate mass-radius relationships, as well as within our own Bayesian adaptive Markov Chain Monte Carlo (MCMC), which estimates the uncertainties of the compositional parameters, given the error bars of the observed mass and radius. We present the various developments we carried out to achieve the coherent implementation of the coupled interior and atmosphere model and its validation.

The applications of the model include the estimation of the compositional parameters with their respective uncertainties based on observations for a single planet, the analysis of planetary compositions within multiplanetary systems to constrain their formation mechanisms, and the assessment of the observability of emission spectra for rocky exoplanets with the James Webb Space Telescope (JWST). We discuss the properties that can be deduced from the compositions of the large sample of 46 planets that we have analysed throughout this thesis, as well as the limitations due to the measurement uncertainties and degeneracies. Our model may serve as a precedent to a modelling framework for the interpretation of mass, radius and atmospheric characterization data provided by upcoming missions such as PLATO and Ariel.

Keywords: Exoplanets. Planetary interiors and atmospheres. Modelling.

Acknowledgements

This PhD thesis has consisted of 3 years of my academic life. Many people have contributed and helped to the writing of it, so I will try my best to thank all of you as you deserve.

I would like to start thanking my PhD supervisors, Magali Deleuil and Olivier Mousis. Thanks for considering me a potential PhD candidate from the earliest of our meetings, while I was a master student in Sweden. Your outstanding vision and perspective of the field is clearly demonstrated in the initial conceptualization of this PhD project, I am glad I was capable of seeing that at such an early stage of my career, and decided to join you in this journey. The first 9 months of this PhD were very tough, mostly due to the lockdowns and the COVID-19 pandemic, when any in-person meeting was not possible. Despite these initial rocky times, I have greatly enjoyed working with you on this project for the two following years. I want to thank you for your guidance, leadership, creativity, and *savoir-faire* in research.

During this thesis, I have had the pleasure and honor to collaborate with many researchers, both outside and inside LAM. I would like to start thanking Emmanuel Marcq, for the enlightening discussions about atmospheric physics during the three years of the PhD, and as an examiner. To Paul Molliere and Eleonora Alei, for their help on using their atmospheric model. I would also like to thank my observer collaborators, whose work has provided me with a significant amount of data points in my analysis: Sergio Hoyer, Theo Lopez and Alexandre Santerne, Neda Heidari and Isabelle Boisse, Jorge Lillo-Box, Carina Persson, Iskra Georgieva, Romy Rodriguez and Thierry Morel. I could not finish this paragraph without thanking the jury of my defence for their questions and discussions, including Giovanna Tinetti, Heike Rauer, and the president Jean-Luc Beuzit. Special thanks to Jérémy Leconte for the discussion on atmospheric modelling, Pierre Drossart for encouraging me to discuss on the Solar System-exoplanet connection, and my future postdoc supervisor, Laura Kreidberg, for the discussion on observations with JWST. I'm looking forward to unveiling the secrets of exoplanets with JWST together with you and the people in MPIA in my postdoctoral phase.

To the other employees in LAM, for creating an environment where we can thrive while doing science. My huge thanks goes to the cluster and IT service, especially to Jean-Charles Lambert, for your patience and willingness to educate me on optimization and use of the LAM cluster.

To all the students in LAM, for making the breaks and lunches a moment for social life and support. It seems unfair to me that I do not have enough room to name and thank you one by one, as you deserve. Thanks to my friend and collaborator, Artem Aguichine. I doubt I will ever find an office mate that plays guitar in the breaks and brightens up the office with unicorn-themed decorations. Talking about office mates, the honorable mention also goes to the most easy-going, Marin Ferrais. I must thank Jorge Villa-Velez and Steven Rendon-Restrepo too, for being like my older brothers, and always showing me how to figure things out in France. Finally, big thanks to my gym buddies, Athanasia and Alexis, for accompanying me at the time I discovered that mind and body are more connected than what doctors (the medical ones) let us know.

Finally, to my family, especially to my parents. I wouldn't be here today with this thesis and career in STEM without the education you gave me from a very early age, and without your support.

Contents

Résumé	3
Abstract	4
Acknowledgements	5
Contents	6
1 Introduction	9
1.1 History of exoplanet characterization	9
1.2 Interior and atmosphere models	14
1.3 Outline	16
2 Interior model	19
2.1 Existing model in 2019	20
2.1.1 Ice implementation	23
2.1.2 Supercritical implementation	24
2.2 Supercritical water	25
2.2.1 Equation of State	25
2.2.2 Grüneisen parameter	28
2.3 Mass-radius relations	33
3 Interior-atmosphere coupling	36
3.1 Coupling algorithm	36
3.2 Low surface pressures	41
4 Atmospheric model	44
4.1 Numerical scheme	44
4.2 Pressure-temperature profile	46
4.3 Collision-induced absorption	49
4.4 K-correlated method	52
4.4.1 Formalism	52
4.4.2 K-table processing	54
4.4.3 Gas mixtures	59
4.5 Cloud and Rayleigh scattering	61
4.5.1 Opacities	61
4.5.2 Single scattering albedo	62
4.5.3 Phase function	62
4.6 Atmospheric model validation	65
4.6.1 Emission spectra	65
4.6.2 Reflection spectra and Bond albedo	69

4.6.3	Mass-radius relations and water phase diagram	72
5	Bayesian statistics	77
5.1	Algorithm	77
5.2	Adaptive MCMC	78
5.3	MCMC Validation	80
5.4	Optimization and Python interface	83
6	Analysis of planetary systems	85
6.1	TRAPPIST-1	85
6.1.1	Results	86
6.1.2	TRAPPIST-1 formation scenarios	90
6.2	Multiplanetary systems	93
6.2.1	K2-138	96
6.2.2	TOI-178	97
6.2.3	Kepler-11	97
6.2.4	Kepler-102	97
6.2.5	Kepler-80	98
6.2.6	Compositional trends	98
6.3	Low multiplicity systems	100
6.3.1	K2-32 and K2-233	102
6.3.2	LHS1140	102
6.3.3	TOI-220	103
6.3.4	HD207897	103
6.3.5	TOI-1233	103
6.3.6	TOI-969	105
6.3.7	TOI-2196	105
6.3.8	K2-106	105
6.4	JWST rocky planets	106
6.4.1	TRAPPIST-1 c	107
6.4.2	55 Cancri e	109
6.5	Planet composition properties	111
7	Discussion	118
8	Conclusion	121
A	First-author contributions	128
A.1	Characterisation of the hydrospheres of TRAPPIST-1 planets (Acuña et al. 2021)	128
A.2	Water content trends in K2-138 and other low-mass multi-planetary systems (Acuña et al. 2022)	140
A.3	Interior-atmosphere modelling to assess the observability of rocky planets with JWST (Acuña et al. Submitted)	156
B	Other publications and contributions	170
C	Résumé en français	174
C.1	Introduction	174
C.2	Modèle	175

C.2.1	Interieur	175
C.2.2	Atmosphère	176
C.2.3	Monte-Carlo par chaînes de Markov	178
C.3	Analyse de systèmes exoplanétaires	179
C.3.1	TRAPPIST-1	179
C.3.2	Systèmes multiplanétaires	180
C.3.3	Cibles d'intérêt du JWST	182
C.3.4	Propriétés déduites de la composition des planètes	182
C.4	Perspectives	183

1. Introduction

1.1. History of exoplanet characterization

The first detected exoplanet, 51 Peg b, was discovered by [Mayor and Queloz \(1995\)](#), expanding the field of planetary science beyond the Solar System. Since then, more than 5000 exoplanets¹ have been detected, showing the wide variety in masses, radii and stellar irradiation of the exoplanet population. There are several exoplanet detection methods, including microlensing, astrometry and direct imaging, but the most prolific ones are transit photometry and the radial velocity method. In transit photometry, we monitor the flux intensity of the light received from a star with time (see [Fig. 1.1](#), left panel). If a planet passes in front of the star while it is orbiting, it will produce a drop in the star's brightness. The depth of this drop is connected to the radius of the planet. Thus, the analysis of the transit light curves allows us to estimate the radius of the planet with respect to the radius of the star. On the other hand, the radial velocity method relies on the gravitational effects caused by a planet on the star it orbits. The gravitational influence of the planet makes the star orbit around the center of mass of the two bodies (see [Fig. 1.1](#), right panel). This motion affects the spectrum of the star seen from Earth's position by the Doppler effect: the spectrum is blueshifted when the star moves towards us ([Fig. 1.1](#), bottom right), while it is redshifted when it moves away ([Fig. 1.1](#), top right). By analysing the variation of the spectral shift in time, we can estimate the distance of the star from the center of mass of the two-body system, which is related to the mass of the planet and the star. Additionally, the transit timing variation method (TTV) is also extremely useful to discover and characterise the masses of exoplanets. In planetary systems with more than one planet, gravitational perturbations are introduced in the orbit of each planet by the other planets. These perturbations can slow down or accelerate the orbits, resulting in a varying periodicity of the orbits (or a variation in the expected timing of the transit). The time the transit deviates from a transit with constant period is related to the mass and number of the other planets in the system. Therefore, if we measure the timing difference of enough transits, we can estimate the masses of the planets in the system ([Agol et al. 2005](#)). These methods are biased towards close-in and large planets, making hot Jupiters like 51 Peg b the first class of exoplanets to be detected and characterized. However, 10 years after the discovery of the first exoplanet, the first low-mass planet ($M < 20 M_{\oplus}$) was detected. This was the super-Earth GJ 876 d ([Rivera et al. 2005](#)), and although the radial velocity method with which it was detected only provided its mass, it was the precedent to the detection and characterization of a smaller class of exoplanets ($R < 4 R_{\oplus}$). This led to the design of space missions and ground-based telescopes with the precision necessary to detect these planets, as well as the development of interior models to interpret the data that they would generate.

NASA's Kepler ([Borucki et al. 2006](#)) and its extended mission, K2, probed a large sample of exoplanets by the transit photometry method, obtaining their radius. It showed that low-mass planets are very common and diverse, constituting $\simeq 85\%$ of detected exoplanets, and presenting both rocky planets and planets with volatile envelopes ([Weiss and Marcy 2014](#)). The sample of

¹NASA Exoplanet Archive, updated on August 1, 2022 (<https://exoplanetarchive.ipac.caltech.edu>).

planets that Kepler provided was statistically analysed by [Fulton et al. \(2017\)](#); [Fulton and Petigura \(2018\)](#), who obtained its radius distribution. This distribution presents two peaks, one at $R = 1.3 R_{\oplus}$, and another one at $R = 2.4 R_{\oplus}$. The size of the sample has been increased and the precision of the radii improved by other transit photometry missions following Kepler, confirming this characteristic distribution ([Petigura et al. 2022](#)). These missions are CHEOPS ([Benz et al. 2021](#)), and TESS ([Ricker et al. 2015](#)), who have continued the legacy of Kepler by discovering more super-Earths and sub-Neptunes, while refining their radii with a precision down to 1-2% in some cases. This has enabled a comparison of the samples with planet interior and evolution models, confirming that the first peak corresponds to planets mainly composed of Fe and Si-bearing rocks, known as super-Earths; whereas the planets found in the second maximum of the distribution have a significant volatile (H/He, water) content. In addition, it was been suggested that super-Earths could either form with very little or no volatile content from the very beginning, or could have been born as sub-Neptunes with H/He atmospheres that lost these due to atmospheric escape processes, transitioning to the super-Earth regime. The most likely atmospheric escape mechanisms that produce the Fulton gap are XUV photoevaporation and/or core-powered mass loss ([Rogers et al. 2021](#)). In the former, hydrogen in the upper layers of the atmosphere escape due to the erosion of X-UV light from the stellar host, which heats the upper atmosphere and induces an hydrodynamic flow ([Lammer et al. 2003](#); [Owen and Jackson 2012](#)). Meanwhile, in core-powered mass loss, the core of the planet emits heat that still remains from the accretion stage. If the primordial hydrogen envelope cools down slower than the core, the thermal energy emitted by the core drives further the loss of atmospheric mass from beneath the envelope. The amount of primordial atmosphere that is lost depends on the planetary mass and the equilibrium temperature of the planet ([Ginzburg et al. 2016](#)).

To constrain the composition of exoplanets, not only the radius is required, but the planetary mass is needed to obtain the total density. This means that transit photometry missions need a radial velocity follow-up with a ground-based spectrograph, such as SOPHIE ([Perruchot et al. 2008](#)), HARPS ([Mayor et al. 2003](#)), HIRES ([Vogt et al. 1994](#)), and ESPRESSO ([Pepe et al. 2021](#)). The first low-mass planet with a complete characterization of its density was super-Earth CoRoT-7 b ([Léger et al. 2009](#); [Queloz et al. 2009](#)), which is compatible with the presence of a water layer or a secondary atmosphere, ruling out a dry rocky planet or a H/He atmosphere ([Valencia et al. 2010](#)). The availability of more instruments at the end of the 2010s allowed us to obtain both the masses and radii of more low-mass planets, including Earth-sized planets around M dwarfs, which are the coolest and smallest stars in the main sequence. The discovery of the seven Earth-sized planets around the M-dwarf TRAPPIST-1 was a break-through in exoplanetary science. The innermost two planets (b and c) plus a third planetary signal were detected via the transit method by [Gillon et al. \(2016\)](#), but a later analysis with more transit photometry data found that the third planet was the combined signal of five planets (e to h), yielding a total of seven planets ([Gillon et al. 2017](#)). Three of the planets in this system (e to g) are within the habitable zone (HZ), which is defined as the range of distances from the host star at which surface water could be in liquid phase. Even though the radii of the TRAPPIST-1 planets were characterised with a precision of 5%, their masses, which were obtained via the radial velocity method, had uncertainties that were too large to provide tight constraints on their densities. This motivated a follow-up with the TTV method, to refine their masses uncertainties ([Grimm et al. 2018](#); [Agol et al. 2021](#)). In the case of TRAPPIST-1, the masses obtained by [Grimm et al. \(2018\)](#) were precise enough to provide individual estimates on the core, water and gas mass fractions with uncertainties, hinting at a dichotomy between the inner water-poor planets and the outer, icy planets. However, the scatter in water content within the system was difficult to explain from the point of view of planet formation ([Dorn et al. 2018](#)). A refinement of the mass estimates by [Agol et al. \(2021\)](#) cleared this scatter, unveiling an increasing water content

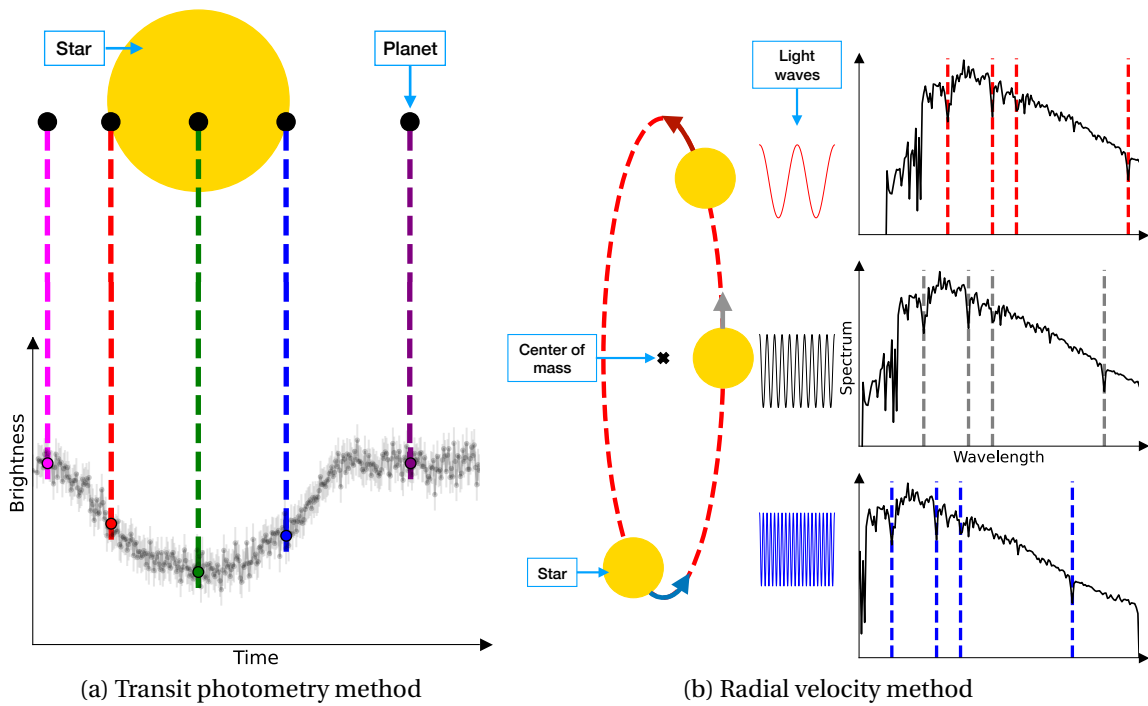


Figure 1.1.: Illustration of the exoplanet detection methods that provide the planetary radius (transit), and mass (radial velocity). The left panel shows the light curve of the hot Jupiter WASP-74 b during one transit (*Courtesy: N. Crouzet*). The right-hand plots in the right panel display the spectrum of a F9-type star when redshifted (top), in the rest frame (middle), and when blueshifted (bottom). The same set of three spectral lines is indicated as dashed lines.

with semi-major axis (Agol et al. 2021; Acuña et al. 2021). The discovery of the TRAPPIST-1 planets were soon followed by other planets in the habitable zone of M dwarfs, such as rocky LHS 1140 b (Dittmann et al. 2017), and the sub-Neptune K2-18 b (Montet et al. 2015; Benneke et al. 2019), as well as multiplanetary systems, such as K2-138 (Christiansen et al. 2018) and TOI-178 (Leleu et al. 2019).

The next step to characterise a planet, after detecting it and obtaining its density, is atmospheric characterization. Atmospheric data bring important information about its composition, formation and evolution. There are several methods to characterize the atmosphere of a planet, which include the following:

- **Transmission spectroscopy.** Similarly to the transit photometry detection method, transmission spectroscopy monitors the brightness of a star during a transit. In this case, the light gathered by the telescope is passed through a diffracting element, such as a diffraction grid, to decompose the white light into different wavelengths. During the transit, the light of the star passes through the atmosphere of the planet, interacting with its gaseous species, which imprint the photons at different wavelengths with spectral absorption lines. This causes the radius of the planet to vary with wavelength, which is the final transmission spectrum. Seager and Sasselov (2000) laid out and proposed this method initially, while the first detection of an atmospheric species was done by Charbonneau et al. (2002), who found sodium lines in the transit spectrum of a hot Jupiter. Since then, several molecular species have been found in Jupiter-sized planets, including water, as well as Rayleigh scattering in the optical caused by

clouds and hazes (Sing et al. 2016).

- **Emission spectroscopy.** As the planet orbits its host star, there will be a point in time when it will pass behind the star as seen from Earth, which is the secondary eclipse (see Fig. 1.2). Just shortly after and before the secondary eclipse, spectroscopic measurements of the flux probe the dayside emission of the exoplanet's atmosphere. Simultaneously, part of the stellar light will be scattered by the atmosphere, generating the reflection spectrum. This corresponds to the variation of the albedo with wavelength. The albedo is defined as the fraction of the stellar light that is reflected by the planet. This means that an albedo close to 1 shows a fully reflecting atmosphere, which is characteristic of clouds (MacDonald et al. 2018). To obtain the dayside emission, the flux of the star plus the planet is compared with the stellar flux alone during the secondary eclipse. In the near infrared (NIR), molecular spectral lines can be detected, while in the optical, the emission spectrum can show reflected light from clouds and hazes. Absorption lines point out to a decreasing temperature gradient with height, whereas emission lines suggest a thermal inversion, which means that temperature increases with height (Perryman 2018).
- **Phase curves.** If the orientation of the planetary system allows to cover a complete orbit with photometric observations, the variation of the star-plus-planet brightness can be monitored, obtaining the phase curve. It will show a slow increase in flux as the dayside of the planet faces Earth (see 'Phase Curve' arrow in Fig. 1.2). Then the phase curve will have a dip down to the single star's flux during the secondary eclipse, followed by a second dip that corresponds to the transit, whose flux decreases down to the star-minus-planet level. The amplitude of the variation between the dayside emission and the nightside, as well as the shape of the phase curve, provide information on the efficiency of the distribution of energy with the longitude of the planet. This is particularly interesting for tidally locked planets: if there is no efficiency in heat re-distribution and the dayside and nightside have very different temperatures, there might be a very thin atmosphere or no atmosphere at all. In the case of Jupiter-size planets, the efficiency in heat re-distribution is very dependent on winds. Moreover, the phase curve can be diffracted in different wavelengths, obtaining the phase-dependent vertical temperatures, which provides the thermal structure and winds speed in hot Jupiters (Stevenson et al. 2014).

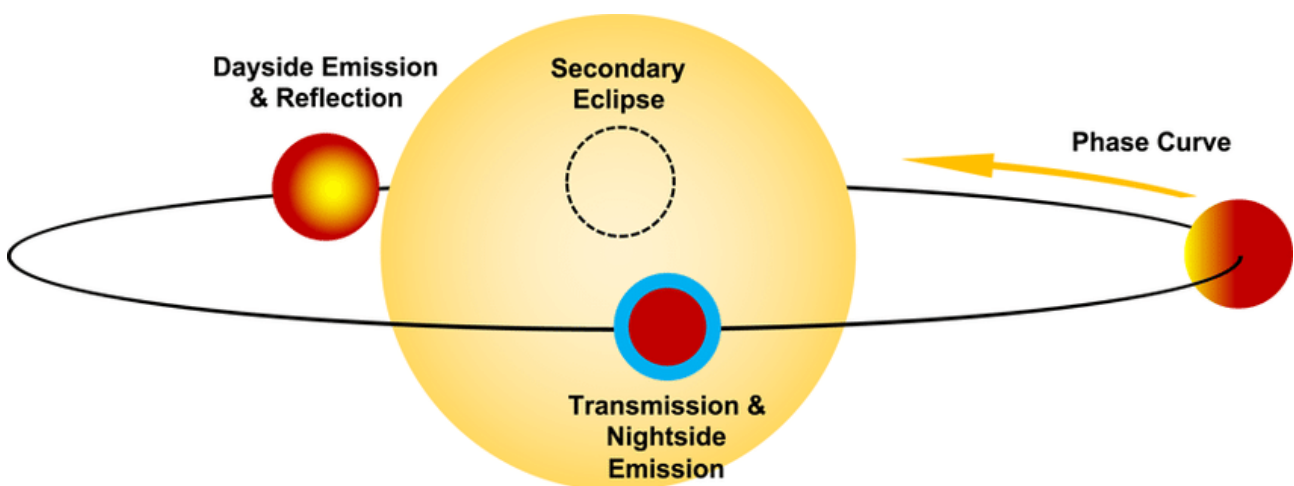


Figure 1.2.: Geometry of a transiting exoplanet for different atmospheric characterization techniques. After Gao et al. (2021).

The first atmospheric characterization data were the transmission spectra of hot Jupiters (Coustenis et al. 1997; Charbonneau et al. 2002), obtained with space-based telescopes. Later on, the two observatories that have played an important role in probing the atmospheres of low-mass planets are Spitzer (Werner et al. 2004), and the Hubble Space Telescope (HST). The first super-Earth to show a spectrum with spectral features and a non-flat phase curve was 55 Cancri e. Its HST transmission spectrum revealed a light atmosphere that could contain H/He, with hints of HCN (Tsiaras et al. 2016). Its Spitzer phase curve showed a large contrast between the dayside and the nightside, suggesting that 55 Cancri e has a thick atmosphere with low efficiency heat re-distribution (Demory et al. 2016). Later on, more exotic compositions, such as a silicate atmospheres, have been looked for with ground-based transmission spectroscopy. However, the non-detection of mineral species has enabled to set absorption limits on these species in 55 Cancri e (Keles et al. 2022). Subsequently, the transmission spectrum of another super-Earth was obtained, GJ1132 b, being featureless and flat, which suggests that there is no atmosphere, or that the atmosphere has a high mean molecular weight or clouds/hazes (Diamond-Lowe et al. 2018; Mugnai et al. 2021). In contrast, a spectral feature around $1.4 \mu\text{m}$ was detected in K2-18 b (Tsiaras et al. 2019; Benneke et al. 2019). It is under debate whether this absorption line in the transmission spectrum is due to water, which could indicate the presence of a condensed water layer underneath the atmosphere (Madhusudhan et al. 2020), or methane (Bézard et al. 2020). The case of K2-18 b shows that atmospheric retrieval is a very degenerate problem, where not only the gaseous species and their abundances have a primary influence in the final spectrum, but also the thermal structure, non-equilibrium chemistry and aerosol parametrization. Additionally, stellar contamination can also hamper atmospheric characterization observations, especially of exoplanets around M-dwarfs. Ducrot et al. (2018) obtained the combined transmission spectrum of TRAPPIST-1 planets by gathering data with different instruments, including Spitzer and HST. This spectrum shows a strong inverted water feature, which is due to the formation of water vapour in the cold dark spots on the surface of the host star (Zhang et al. 2018). Other stellar contamination models with higher spot temperatures and a lower spot surface coverage could not account for the complete set of data, which could lead to the conclusion that the spectral water signals come from the planetary atmosphere, and not the stellar one. This shows the importance of stellar contamination modelling in atmospheric characterization. Very recently, the transmission spectrum of LHS1140 b was acquired with HST, showing a tentative water feature. This water feature can be fully explained by stellar contamination models, or could be a combination of contributions from the planet and the star (Edwards et al. 2021). Simultaneous works also include the atmospheric characterization of sub-Neptune GJ1214 b, which has a tentative detection of He (Orell-Miquel et al. 2022); hot super-Earth π Mensae c, whose detection of C II ions indicates atmospheric escape of a high-molecular atmosphere (García Muñoz et al. 2021); and K2-141 b, which might present a tenuous silicate atmosphere (Zieba et al. 2022). Finally, the atmospheric characterization techniques can also be used to confirm the absence of an atmosphere, which is the case of the transmission spectrum of LHS 3844 b (Diamond-Lowe et al. 2020). Moreover, Kreidberg et al. (2019) were able to constrain the material the surface is made of, based on the emission flux, which favours a basaltic rock. These examples show the diversity in bulk and atmospheric compositions in low-mass planets, which require detailed interior and atmosphere models to interpret the upcoming data with on-going and future observatories.

1.2. Interior and atmosphere models

Given the mass and radius of a planet, M_P and R_P respectively, its density can be calculated assuming a perfect sphere shape:

$$\rho_P = \frac{M_P}{\frac{4}{3}\pi R_P^3} \quad (1.1)$$

The density can be used as a proxy of the composition in first approximation. Planets rich in rocks and iron have densities similar to that of the terrestrial planets in the Solar System ($\rho_P \approx 3.9 - 5.5 \text{ g/cm}^3$), while planets with a volatile composition have low densities, $\rho_P < 1.65 \text{ g/cm}^3$. Nonetheless, the density does not only depend on the composition, but also on the mass and the dependence of the temperature with radius that is assumed, which is the thermal structure of the interior. This is why detailed interior models are necessary to obtain an accurate estimate of the composition of a planet. Interior models compute the radius of the planet given their mass and their composition. This is done by solving a set of equations that describe the density, the transport of heat, and the behaviour under the effects of gravity and pressure of the different materials (Fe, rock, water, hydrogen). For a first approximation of the planetary composition, the mass and radius can be compared to mass-radius relationships. These are computed by interior models, and express the radius as a function of mass for a constant interior configuration or composition (see Fig. 1.3 for an example of mass-radius diagram and relationships). These relations were derived for the gaseous giants of the Solar System as far back as the 1970s (Hubbard 1973), followed years later by the mass-radius relations for the first super-Earth, GJ 876 d (Valencia et al. 2006; Seager et al. 2007).

A variety of interior models has been developed to describe the diversity found in exoplanets. For terrestrial planets, interior models take as reference the Earth, Venus and Mars (Valencia et al. 2006), which present an inner Fe core, surrounded by a silicate mantle. Exoplanets with significant water/ice layers were proposed by Léger et al. (2004), named ocean planets, containing a Fe core and rock layer with a global ocean layer of 100 km. Further works propose the formation of a high pressure, high temperature layer in these ocean planets, especially for highly-irradiated planets, where water can reach supercritical and plasma phases (French et al. 2016; Mazevet et al. 2019; Mousis et al. 2020), which have very different densities compared to liquid and ice water. Seager et al. (2007), Fortney et al. (2007), and Lopez and Fortney (2014) considered H/He-dominated interiors instead of water-rich compositions by using single-material layers of Fe for the core, perovskite (MgSiO_3) for the mantle, and a H and He mixture for the envelope. It becomes clear that there are intrinsic degeneracies in the composition of exoplanets. We say a problem is degenerate when many possible solutions yield the same result. In this case, this means that several composition configurations produce the same planetary density. Fig. 1.3 shows an example of this. The mass and radius of 55 Cancri e could be explained by a pure silicate rock composition (100% mantle), or by an interior that contains Fe, with a water atmosphere on top. This is because Fe increases the planetary density, while a water atmosphere decreases it, compensating each other's density changes with respect to the 100% mantle case. Another degeneracy would be between the fraction of mass that constitutes the atmosphere, and the atmospheric composition. An example of this is HD 207897 b (Heidari et al. 2022), whose density could be accounted for with a water atmosphere that constitutes 25% of its mass, or a H_2 envelope that entails only 0.2% of its mass, assuming the same rocky core underneath (see Fig. 1.3). To reduce the former degeneracy, it has been proposed to consider the Fe and Si abundances of the host star (Valencia et al. 2007; Dorn et al. 2015; Brugger et al. 2016; Unterborn et al. 2018; Wang et al. 2018). The abundances of the refractory elements

in the protoplanetary disk and the stellar host are similar. Formation models indicate that the planets formed in the planetary disk retain this Fe-to-Si abundance ratio (or Fe/Si mole ratio) by accreting material from the protoplanetary disk (Thiabaud et al. 2015). Therefore, it is widely extended in interior modelling to use the chemical stellar abundances of Fe and Si to constrain the core-to-mantle ratio, as most of the Fe is found in the core, and the mantle contains Si. While the core mass fraction (CMF) is constrained by the Fe/Si mole ratio, the remaining compositional parameter, which is the volatile mass fraction (or water mass fraction, for models that only consider a water layer), is controlled by the density. Nonetheless, it is still under discussion if the stellar Fe/Si ratio represents accurately the real Fe/Si of the planet interior (Dorn et al. 2015; Adibekyan et al. 2021). Therefore, the most complete interior structure analyses should include scenarios with different Fe/Si ratios.

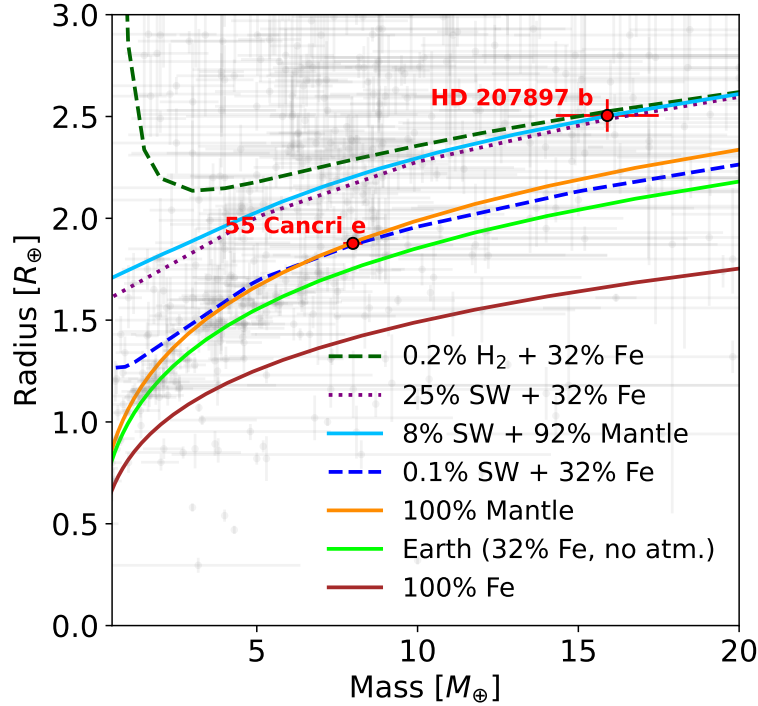


Figure 1.3.: Mass-radius diagram for super-Earth 55 Cancri e, and sub-Neptune HD 207897 b. Mass-radius relationships for supercritical water (SW) (Acuña et al. 2021), rocky dry planets (Brugger et al. 2017), and H/He atmosphere planets (Zeng et al. 2019) are shown to illustrate the degeneracy in planetary interiors. Grey points and error bars show the discovered exoplanets with mass and radius measurements (NASA Exoplanet Archive).

Back in 2019, at the start of this PhD thesis, the interior models developed to explore different Fe/Si ratios for low-mass planets had limitations: they were only applicable to planets with liquid water surface conditions (Brugger et al. 2016; Unterborn et al. 2018), or they considered a H/He atmosphere on top of an ice layer (Dorn et al. 2017). However, many of the exoplanets we characterize are highly irradiated, where water would be in the form of gas instead of liquid or ice. Applying an interior model with liquid/ice water to these planets would overestimate the real water content, since ice and liquid water are more dense than steam. Therefore, there was a need to develop an interior model that could calculate consistently the effects of an atmosphere on the surface conditions and the radius of super-Earths and sub-Neptunes, while it could simultaneously explore different core-to-mantle fractions. In addition, super-Earths cannot present H/He envelopes given

their small radius (Lopez and Fortney 2014), making secondary atmospheres with water and CO₂ more appropriate to assume in the modelling of their interior. It was still to be seen if the densities of sub-Neptunes could be explained only with H/He envelopes (Zeng et al. 2019), or if they were also consistent with pure water envelopes.

An envelope needs to be implemented to account for the effects of a highly irradiated atmosphere. Until 2019, the atmosphere prescriptions within interior models were typically analytical radiative models. For example, Valencia et al. (2013), Jin et al. (2014) and Dorn et al. (2017) used an analytical model that is benchmarked with a numerical full-physics atmosphere model for a fixed atmospheric solar composition. Their advantage is that they are computationally fast compared to numerical atmosphere models. Analytical models are derived by assuming an approximation to the radiative transfer equation and prescribe the pressure, temperature and opacity (P-T- τ) profile depending on different free parameters (Guillot 2010; Heng et al. 2014). The opacity is a proxy of how opaque is an atmosphere layer to radiation, being close to 0 if the gas layer is transparent and lets all radiation pass through. They require the calculation of the mean opacity, which is often assumed to be independent of wavelength. This is known as a grey opacity, and it is only dependent on the pressure and the temperature. Tabulated values of grey opacities were publicly available (Freedman et al. 2014), but only for H/He-dominated atmospheres. Consequently, these tabulated opacities do not reflect the absorption and emission behaviour of a secondary atmosphere that is mainly composed of H₂O, and/or CO₂. There are atmosphere analytical models that allow for non-grey opacities, such as the picket fence model (Parmentier and Guillot 2014); as well as analytical models adapted to the secondary atmospheres of Solar System bodies, such as Venus and Titan (Robinson and Catling 2012; Tolento and Robinson 2019). However, they would still require a numerical non-grey atmospheric model to calibrate their free parameters. Moreover, they would not be able to generate spectra to use for predictions of atmospheric characterization observations. Therefore, for our purposes, numerical atmosphere models are more appropriate to couple with an interior model. One-dimensional atmospheric models can be divided in two types of models depending on how they compute their opacities: line-by-line models, and k-correlated models. Line-by-line models are able to compute the shape and intensity of each individual opacity line to produce high-resolution spectra; whereas k-correlated models perform an approximation in the integration of the opacity over wavelength to reduce the data and computation time.

Finally, the interior models mentioned so far, such as the one developed by Brugger et al. (2017), are forward models. This means that their input parameters are not necessarily observable, such as the water or Fe content of the planet. To have as input the observable parameters (mass and radius), and as output the non-observables, it is necessary to invert the forward problem. Bayesian inversion methods are used to do this, such as Markov Chain Monte Carlo (Mosegaard and Tarantola 1995) and nested sampling (Buchner 2021). These are also necessary to take into account the error bars of the observables to estimate the uncertainties of the non-observable parameters. While these methods were widely used in atmospheric composition retrieval, very few studies had implemented these to retrieve the interior composition (Dorn et al. 2015).

1.3. Outline

At the beginning of this PhD thesis, the interior model developed by Brugger et al. (2017) had a detailed implementation of a Fe core, a Si-rich mantle and a condensed water layer with liquid and ice VII phases. I started my PhD by continuing the development of this interior model so it would be applicable to highly-irradiated planets. This involved changing the data that express

the relations between the pressure, temperature, and the density in the water layer of the interior model, which are the Equation of State (EOS) and the Grüneisen parameter. In Chapter 2, I detail the implementation of the water layer in the interior model for highly-irradiated conditions, namely the supercritical water layer. Moreover, I also carry out a comparison of different EOS for supercritical water, their validity and effects on the computation of mass-radius relationships of water-rich planets.

In addition, I coupled the interior model with a one-dimensional k-correlated model for water atmospheres to account for the effects the atmosphere produces on the surface conditions and on the total planetary radius. In 2019, only planet formation models had been consistently coupled with atmospheric models (Linder et al. 2019; Marleau et al. 2019). Consequently, I had to develop an algorithm specifically designed to couple our interior model with the atmospheric model developed by Marcq et al. (2017) and Pluriel et al. (2019). In Chapter 3, I detail the algorithm to couple self-consistently the interior model exposed in the previous chapter with this atmospheric model. The initial coupling was done by using pre-generated files of the atmospheric parameters, instead of running the interior and the atmosphere at the same time. I updated the EOS and opacity data, and propose a k-uncorrelated approximation for the atmospheric model (see Chapter 4) to produce low-resolution emission spectra and radius estimates simultaneously. In the time I completed this PhD thesis (2019-2022), several interior models have been coupled to numerical atmospheric models. These include line-by-line models coupled to generate high-resolution spectra to predict observations (Katyal et al. 2020; Madhusudhan et al. 2020), as well as k-correlated models to compute the interior conditions of habitable ocean planets (Marounina and Rogers 2020), and the radius of highly-irradiated water-rich planets (Turbet et al. 2020). This shows that the development of interior-atmosphere models for super-Earth and sub-Neptune exoplanets is an active research field, and the work presented in this thesis contributes with a different approach compared to independent and simultaneous studies. Finally, in Chapter 5 I explain the MCMC Bayesian algorithm I implemented to obtain the posterior distribution functions of the compositional parameters and basic atmospheric parameters. In this Chapter, I also include a brief explanation of the user-friendly interface I built to use the interior and atmosphere models in Python 3, as well as optimization of the interior model to facilitate its use with MCMC methods and packages.

The aim of this work is to provide a modelling framework to obtain the compositional parameters with their respective uncertainties based on the error bars of the observed mass and radius, under the assumption of a water-dominated volatile layer. The applications of our model are shown in Chapter 6. This includes TRAPPIST-1, the planetary system that is well-known for hosting seven terrestrial planets, three of which are in the habitable zone. Agol et al. (2021) obtain estimates with an interior model on the core and water contents with the most precise mass data on the system. However, our self-consistent interior-atmosphere presents a few improvements in comparison to their modelling work, including the inclusion of a CO₂ atmosphere composition for the inner planets (see Sect. 6.1).

Multiplanetary systems act as laboratories to explore the diverse formation and evolution pathways of low-mass planets, as they highlight differences in composition caused by different formation mechanisms within the same environment. A self-consistent interior structure model that is applicable to a wide range of irradiations such as ours allows us to carry out a homogeneous analysis of a sample of multiplanetary systems (see Sect. 6.2). By doing this, we overcome the differences in volatile mass fraction estimates of multi-planetary systems due to the different assumptions taken by each interior structure model in the literature.

In addition, I also perform an interior analysis of several low-multiplicity systems, or multiplanetary systems with only a few planets whose masses and radii are available. These systems include

LHS 1140, which hosts one Earth-sized planet in the habitable zone, LHS 1140 b; and TOI-220 b, which is a highly-irradiated sub-Neptune located in the Fulton gap, which is interesting from the point of view of atmospheric escape. The final application of our model is the generation of emission spectra to assess the observability of two targets of interest for the James Webb Space Telescope (JWST), TRAPPIST-1 c and 55 Cancri e. I close Chapter 6 with a discussion on the complete sample I have analysed, with insights into the composition of the low-mass planet population and uncertainty trends.

In Chapter 8, I summarize the key findings of our model, while also commenting on the future work necessary to keep reducing the degeneracies found in interior and atmosphere modelling. I conclude this thesis by elaborating on the application of self-consistent interior-atmosphere models to the data that will be provided by on-going and future missions such as JWST, Ariel, and PLATO.

2. Interior model

This chapter introduces the new developments implemented in the interior model with respect to the versions that existed in 2019, when I started this PhD thesis. The first version of the model is for planets with liquid water on their surface (Brugger et al. 2016, 2017), which is presented at the beginning of Section 2.1. The second version of the model, which is applicable to planets with ice phases on their surface (see Fig. 2.1, blue), was initially implemented by M. Levesque (Levesque 2019). The improvements introduced in this version are detailed in Sect. 2.1.1. The third version has the aim of making the model applicable to volatile-rich planets whose surface conditions do not allow for the presence of liquid water by implementing a supercritical water layer. The existing supercritical version in 2019 is introduced in Sect. 2.1.2.

The updates I performed in the supercritical water model with respect to the existing version in 2019 are indicated in Sect. 2.2.

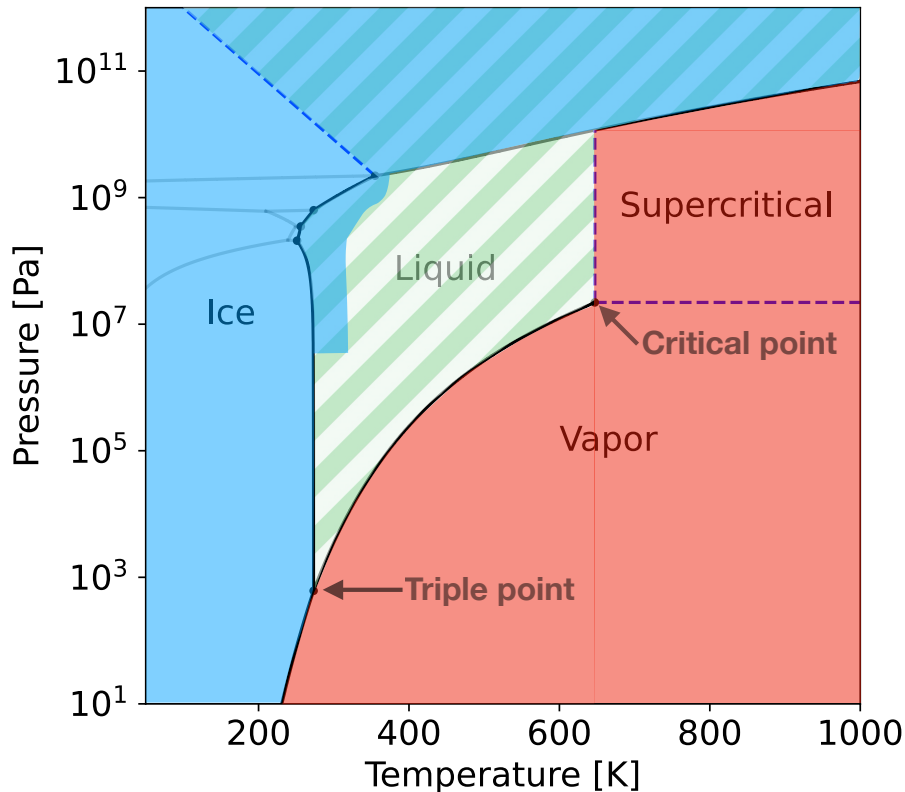


Figure 2.1.: Water phase diagram showing the four main phase of water: liquid, ice, vapor, and supercritical. The version of the model for planets with liquid water (Brugger et al. 2017) covers the area indicated in green stripes. The ice (see Sect. 2.1.1) and supercritical (see Sect. 2.1.2) versions of the model cover the regions marked in blue and red, respectively. Dashed blue lines mark the boundary between supercritical and liquid and vapor phases (right), as well as the lowest temperature in ice VII that the model of Brugger et al. (2017) would cover (left). Grey lines delimit the ice transitions in the ice version of the model.

2.1. Existing model in 2019

The interior structure model for low-mass planets was initially developed by [Brugger et al. \(2016, 2017\)](#), which is a 1D model that presents three separate layers: a Fe-rich core, a Si-rich mantle and a condensed water layer (see Fig. 2.2). The input of the model are the planetary mass and two compositional parameters. These parameters are the core mass fraction (CMF) and the water mass fraction (WMF), which are calculated as the mass of the corresponding layer (core or hydrosphere, respectively) divided by the total mass of the planet. The one-dimensional grid in the model represents the radius, ranging from zero at the center of the planet, up to the surface. Along the spatial grid, we calculate the pressure, $P(r)$; the temperature, $T(r)$; the gravity acceleration, $g(r)$; and the density, $\rho(r)$. These are computed by solving a set of four differential equations (see Eqs. 2.1 to 2.5), which correspond to the equation of hydrostatic equilibrium for the pressure, Gauss's theorem for the gravity acceleration, the adiabatic gradient for the temperature, and the equation of state (EOS) for the density. In this work, we adopt the Vinet EOS ([Vinet et al. 1989](#)) with a thermal correction for the core, the lower mantle, and the upper mantle. Reference parameters of the Vinet EOS for these three layers can be found in Tables 2.1 to 2.3 of [Brugger \(2018\)](#).

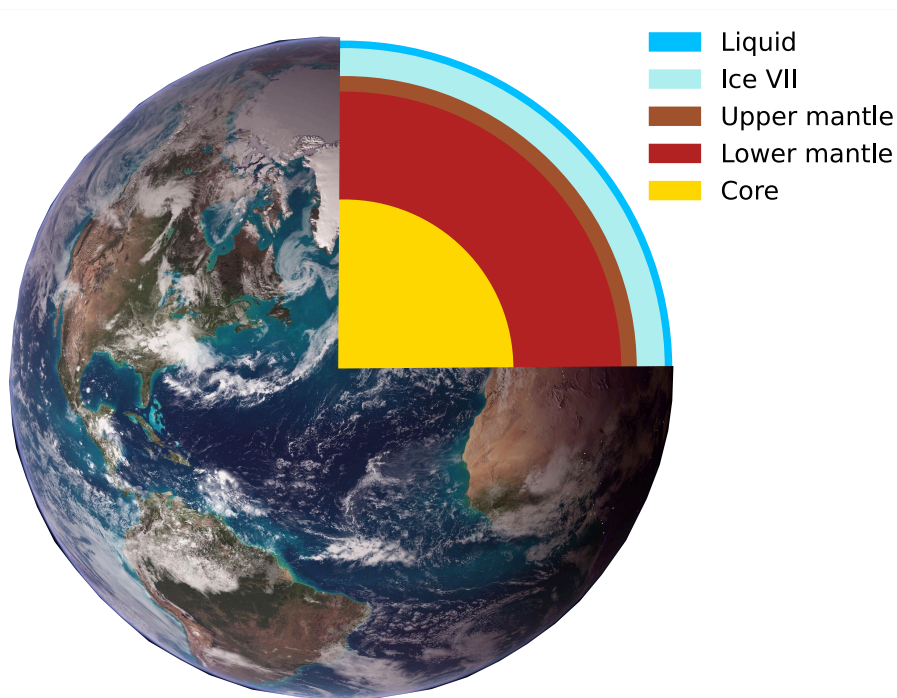


Figure 2.2.: Interior layers for a water-rich, Earth-sized planet according to our model for temperate planets. In the interior structure model for highly-irradiated planets, the ice VII and liquid water layers are substituted by a supercritical water layer. (*Image credits: National Geographic Society*).

In Eqs. 2.1 to 2.5, $G = 6.674 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$ is the gravitational constant. m represents the mass contained within the radius r , whereas γ and ϕ are the Grüneisen and the seismic parameters, respectively. The Grüneisen parameter represents the dependence of the vibrational properties of a crystal with the size of its lattice. It establishes the relation between the temperature in a solid to its total density. The seismic parameter characterises the propagation of seismic waves inside a solid. These two parameters are defined in Eq. 2.5, where E corresponds to the internal energy, and V is

volume.

The differential equations require boundary conditions to be solved. These conditions consist of the gravitational acceleration at the planetary center, $g(r = 0) = 0$, and the surface temperature and pressure, $P(r = R) = P_{surf}$ and $T(r = R) = T_{surf}$. Eqs. 2.1 to 2.5 are solved iteratively until the pressure, temperature, gravity and density profiles have converged (see Fig. 2.3). Once the density profile is known, the mass of each layer can be computed by integrating Eq. 2.6, which is the equation of mass conservation (Brugger et al. 2017; Sotin et al. 2007). The sum of the masses of all layers is then the total planetary mass, whereas the CMF and the WMF are calculated as the mass of the core and the hydrosphere layers divided by the total mass, respectively. The interior model reaches convergence when the total integrated mass is equal to the input planetary mass.

$$\frac{dP}{dr} = -\rho g \quad (2.1)$$

$$\frac{dg}{dr} = 4\pi G\rho - \frac{2Gm}{r^3} \quad (2.2)$$

$$\frac{dT}{dr} = -g \frac{\gamma T}{\phi} \quad (2.3)$$

$$\phi = \frac{dP}{d\rho} \quad (2.4)$$

$$\gamma = V \left(\frac{dP}{dE} \right)_V \quad (2.5)$$

$$\frac{dm}{dr} = 4\pi r^2 \rho \quad (2.6)$$

In addition to the CMF and the WMF, the interior structure model also calculates a third compositional parameter, which is the Fe/Si mole ratio. This is defined as:

$$\left(\frac{Fe}{Si} \right)_P = \frac{\sum_{i=1}^3 n_i (Fe)_i}{\sum_{i=1}^3 n_i (Si)_i} \quad (2.7)$$

where $n_i = M_i / M_{mol,i}$ is the ratio between the mass of layer i , and the mean molecular mass of the material that composes the layer. $(Fe)_i$ and $(Mg)_i$ are the mole fractions of Fe and Si in layer i , respectively. The detailed distribution of the materials and chemical species in the core and mantle can be found in Brugger (2018). The Fe/Si mole ratio can be derived analytically as well (Brugger et al. 2017):

$$\left(\frac{Fe}{Si} \right)_P = \frac{1}{Mg\#} \left(\frac{Mg}{Si} \right)_P \left[1 - Mg\# + \frac{M_{mol,2}}{M_{mol,1}} \frac{CMF}{1 - CMF - WMF} \right] \quad (2.8)$$

where Mg# is the Mg number, and $M_{mol,1}$ and $M_{mol,2}$ are the mean molecular weight of the core and the lower mantle, respectively. The Mg number is defined as $Mg\# = (Mg / (Mg + Fe))_{mantle}$. If the planet is fully differentiated, in other words, all Fe is found in the core and all Mg is found in the mantle, the Mg number is equal to 1. Therefore, Mg# indicates the level of differentiation of the planet. In Eq. 2.8, for constant values of Mg#, molecular weights of the core and lower mantle, and Mg/Si mole ratio, the Fe/Si is only dependent on the CMF and the WMF. Fig. 2.4 shows the varying WMFs and CMFs values along a line where the Fe/Si mole ratio is constant. We denote these Fe/Si

isolines. The Fe/Si mole ratio can help break the degeneracy of the WMF and CMF for a given pair of planetary mass and radius [Brugger et al. \(2017\)](#).

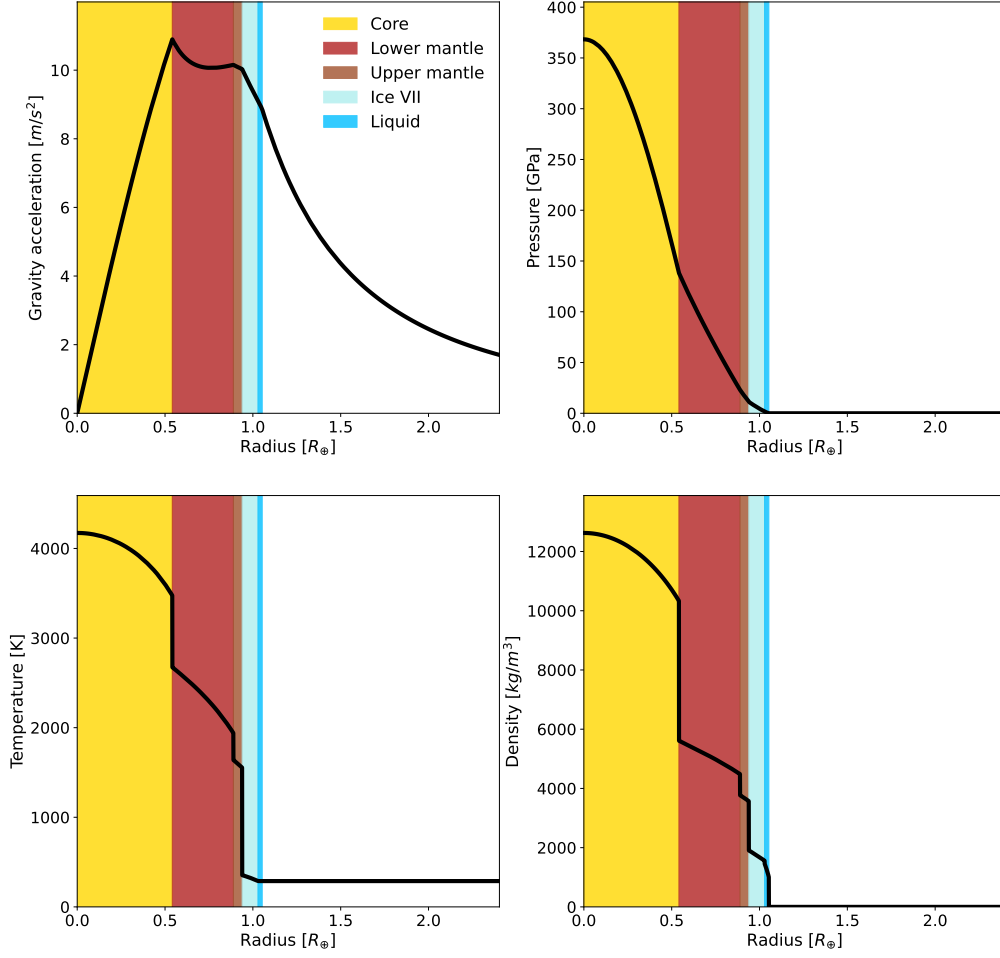


Figure 2.3.: Planetary profiles for a water-rich, Earth-sized planet calculated by our model for temperate planets. The assumed parameters are $M = 1 M_{\oplus}$, $CMF = 0.32$, $WMF = 0.10$, $P_{surf} = 1$ bar, and $T_{surf} = 288$ K.

Depending on the surface conditions, the water layer, which is the outermost layer, can present different phases. [Brugger et al. \(2016, 2017\)](#) implement liquid and ice VII phases, which are relevant for temperate planets whose irradiation conditions are similar to that of Earth. Nonetheless, many of the extrasolar planets that we detect and characterise currently are in very close-in orbits, showing irradiation temperatures far above the critical point of water. This makes necessary the implementation of high-temperature, high-pressure phases in the interior structure model, which we explain in Sect. 2.2.

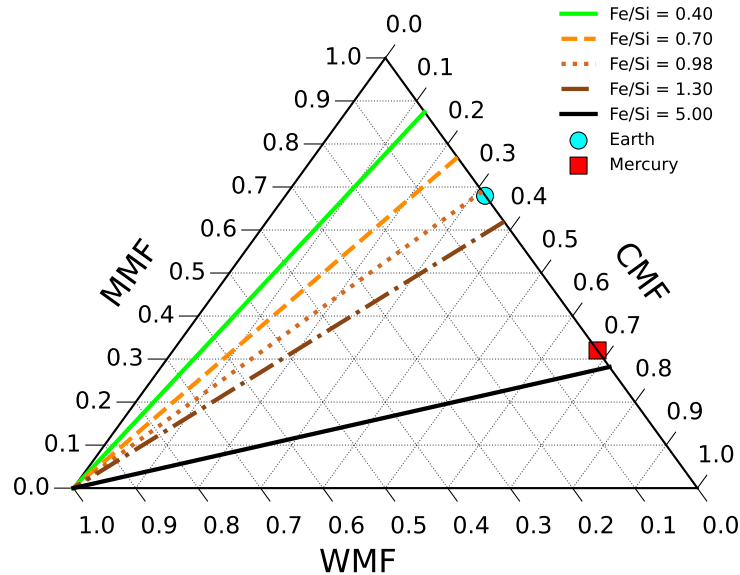


Figure 2.4.: Ternary diagram with isolines of constant Fe/Si mole ratio.

2.1.1. Ice implementation

The version of the model described above assumes that the outermost layer of the planet is liquid water, with a high-pressure ice VII layer underneath. In addition to this version, there existed a version of the model that had implemented more ice phases so the model could be applicable to icy planets. Icy planets have lower surface temperatures than planets that can sustain liquid water, which correspond to temperatures below 250 K. This version of the model was developed by Maëva Levesque (Levesque 2019). In the sample of planets I have analysed in this thesis, three of them required to use this version of the model due to their low temperatures: TRAPPIST-1 f, g and h (see Chapter 6).

I revisited the original code and modified the implementation of the interfaces between the layers of ice to represent more accurately their respective phase transitions. These phase changes are ice VII to VI, VI to II, III to V, I to II, and I to III (see Fig. 2.5). These are now fully based on the phase transition functions provided by Dunaeva et al. (2010).

Further details on this version of the model can be found in section 2.2 of Acuña et al. (2021) (see Sect. A.1 in the Appendix), including the calculation of the EOS of the ice phases and their corresponding references.

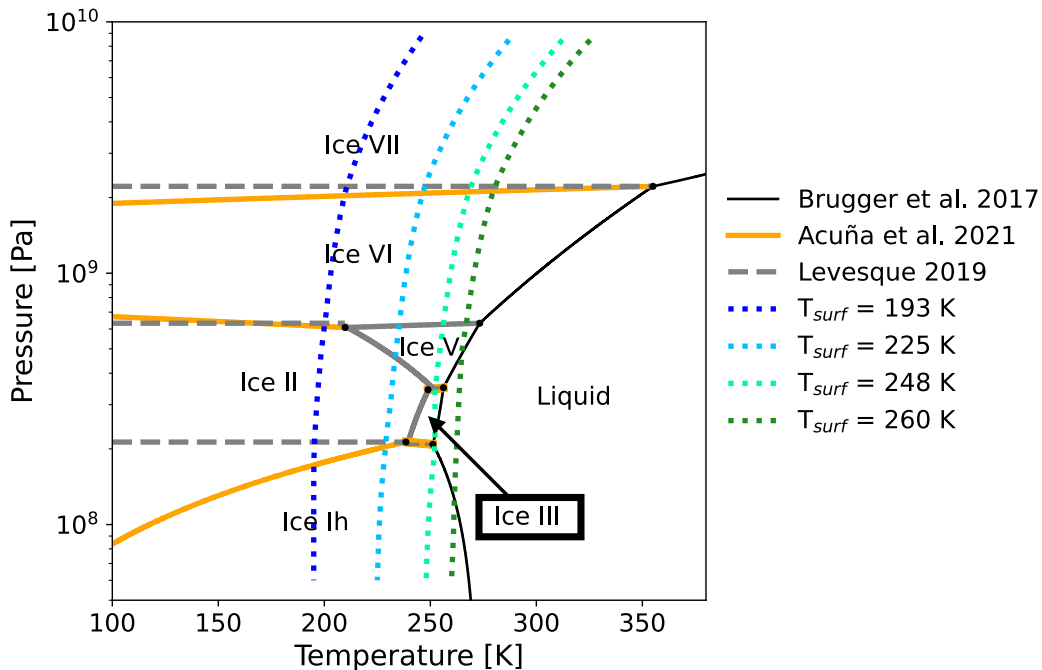


Figure 2.5.: Water phase diagram for ice phases. Grey dashed lines and solid orange ones highlight the differences in the implementation of the phase changes between the interior model version in 2019 and the model used in (Acuña et al. 2021) (see Appendix A.1). Dotted lines show the pressure-temperature profiles for a planet with 41% Fe, 5% water, $M = 1.36 M_{\oplus}$, and a surface pressure of 600 bar.

2.1.2. Supercritical implementation

The high-pressure and high-temperature phases of water correspond to the supercritical and superionic phases in ocean planets (Zeng and Sasselov 2014; Thomas and Madhusudhan 2016), whose density is significantly lower than that of liquid and ice water. The Equation of State (EOS) is the mathematical expression that yields the pressure of a material given a certain density and temperature, $p = p(\rho, T)$. Consequently, the first step in the inclusion of the supercritical phase in the interior model is to substitute the EOSs of the condensed water layer by an EOS that is valid in the supercritical pressure and temperature ranges.

J. Naar implemented a supercritical water layer in 2019 (Naar 2019). I revised and corrected this code in collaboration with A. Aguichine. We corrected the computation of the complete temperature profile and the density at the top interface of the supercritical layer, in addition to solving convergence issues. This corrected version was used to calculate pressure-temperature profiles and mass-radius diagrams in Mousis et al. (2020). Nonetheless, the EOS implemented for supercritical water in this version by J. Naar corresponds to that of Duan and Zhang (2006) (see Sect. 2.2.1.2), which is based on quantum molecular dynamics simulations, rather than experimental data as the standard EOS for steam (Wagner and Pruß 2002). I carried out a bibliographic revision, where I found an EOS that combines both experimental data and theoretical simulations (Mazevet et al. 2019). I present a comparison between these three formulations in the next section.

2.2. Supercritical water

In Section 2.2.1, I explain the formalism of the three EOSs developed for the supercritical regime of water, and I compare them. These EOSs are the International Association for the Properties of Water and Steam (IAPWS-95) EOS (Wagner and Pruß 2002), Duan and Zhang (2006) (DZ06) and Mazevet et al. (2019) (M19) EOSs. Similarly, of these three formulations, two of them also present a formalism for the Grüneisen parameter. In Sect. 2.2.2, I describe their implementation for the high-pressure, high-temperature supercritical regime. Finally, in Section 2.3, I compare the resulting mass-radius relationships with varying water contents and formulations.

2.2.1. Equation of State

2.2.1.1. IAPWS-95 formulation

The IAPWS-95 EOS is based on a formulation that fits experimental data within its validity region, and shows reasonable behaviour within its extrapolation region in the water phase diagram. The validity region is 251 to 1273 K in temperature and up to 1 GPa in pressure; whereas the extrapolation area ranges from 1273 K to 5000 K in temperature, and 1 GPa to 100 GPa in pressure.

The IAPWS-95 EOS is formulated as a function of the Helmholtz free energy. In its dimensionless form, $f(\rho, T)$, it can be separated in two terms: an ideal-gas term, ϕ^0 , and a residual one, ϕ^r :

$$\frac{f(\rho, T)}{RT} = \phi(\delta, \tau) = \phi^0(\delta, \tau) + \phi^r(\delta, \tau) \quad (2.9)$$

where $\delta = \rho/\rho_c$ and $\tau = T_c/T$. $\rho_c = 322 \text{ kg/m}^3$ and $T_c = 647.096 \text{ K}$ are the density and temperature at the critical point of water, respectively. The pressure (in kPa) is defined as:

$$p(\delta, \tau) = (1 + \delta\phi_\delta^r) \rho RT \quad (2.10)$$

where R is the specific gas constant, $R = 0.461 \text{ kJ kg}^{-1} \text{ K}^{-1}$. We define ϕ_δ^r as the first partial derivative of the residual term of the Helmholtz free energy with respect to δ . This is expressed as the sum of four terms:

$$\phi_\delta^r = \phi_{\delta,1}^r + \phi_{\delta,2}^r + \phi_{\delta,3}^r + \phi_{\delta,4}^r \quad (2.11)$$

Each of these parts are defined as:

$$\phi_{\delta,1}^r = \sum_{i=1}^7 n_i d_i \delta^{d_i-1} \tau^{t_i} \quad (2.12)$$

$$\phi_{\delta,2}^r = \sum_{i=8}^{51} n_i e^{-\delta^{c_i}} \left[\delta^{d_i-1} \tau^{t_i} (d_i - c_i \delta^{c_i}) \right] \quad (2.13)$$

$$\phi_{\delta,3}^r = \sum_{i=52}^{54} n_i \delta^{d_i} \tau^{t_i} e^{-\alpha_i(\delta-\varepsilon_i)^2 - \beta_i(\tau-\gamma_i)^2} \left[\frac{d_i}{\delta} - 2\alpha_i(\delta-\varepsilon_i) \right] \quad (2.14)$$

$$\phi_{\delta,4}^r = \sum_{i=55}^{56} n_i \left[\Delta^{b_i} \left(\Psi + \delta \frac{\partial \Psi}{\partial \delta} \right) + \frac{\partial \Delta^{b_i}}{\partial \delta} \delta \Psi \right] \quad (2.15)$$

where n_i , d_i , t_i , c_i , α_i , β_i , γ_i , ε_i , b_i , A_i , B_i , C_i and D_i are tabulated coefficients that can be found in Table 2 of Wagner and Pruß (2002). For the fourth term, the variables Ψ and Δ are expressed as:

$$\Psi = e^{-C_i(\delta-1)^2 - D_i(\tau-1)^2} \quad (2.16)$$

$$\Delta = \theta^2 + B_i [(\delta - 1)^2]^{a_i} \quad (2.17)$$

where $\theta = (1 - \tau) + A_i [(\delta - 1)^2]^{1/2\beta_i}$. Finally, their partial derivatives with respect to δ are:

$$\frac{\partial \Psi}{\partial \delta} = -2C_i(\delta - 1)\Psi \quad (2.18)$$

$$\frac{\partial \Delta^{b_i}}{\partial \delta} = b_i \Delta^{b_i-1} \frac{\partial \Delta}{\partial \delta} \quad (2.19)$$

where:

$$\frac{\partial \Delta}{\partial \delta} = (\delta - 1) \left\{ A_i \theta \frac{2}{\beta_i} [(\delta - 1)^2]^{1/2\beta_i-1} - 2B_i a_i [(\delta - 1)^2]^{a_i-1} \right\} \quad (2.20)$$

2.2.1.2. Duan & Zhang formulation

The EOS developed by [Duan and Zhang \(2006\)](#) consists on a fit to molecular dynamics simulations that assume an ab initio molecular potential. Its validity region extends up to 2573 K in temperature, and up to 10 GPa in pressure. In comparison to the experimental data from the IAPWS-95, DZ06 presents a deviation of 0.1%, while in its remaining validity region, the deviation from the IAPWS-95 EOS is 1.3%. The extended region spans up to $T = 2800$ K and $p = 35$ GPa, with deviations from other benchmark EOS of up to 5% ([Duan et al. 1996](#)).

The pressure is formulated as:

$$p(V, T) = \frac{ZRT}{V} \quad (2.21)$$

where R is the universal gas constant, $R = 83.14467 \text{ cm}^3 \text{ bar K}^{-1} \text{ mol}^{-1}$, and V is the mole density in mol/cm^3 . The adimensional compressibility factor, Z , is expressed as a sum of several end-members:

$$Z = 1 + \frac{BV_c}{V} + \frac{CV_c^2}{V^2} + \frac{DV_c^4}{V^4} + \frac{EV_c^5}{V^5} + \frac{FV_c^2}{V^2} \times \left(\beta + \frac{\gamma V_c^2}{V^2} \right) e^{-\gamma V_c^2/V^2} \quad (2.22)$$

The parameters of these end-members are defined as:

$$B = a_1 + \frac{a_2}{T_r^2} + \frac{a_3}{T_r^3} \quad (2.23)$$

$$C = a_4 + \frac{a_5}{T_r^2} + \frac{a_6}{T_r^3} \quad (2.24)$$

$$D = a_7 + \frac{a_8}{T_r^2} + \frac{a_9}{T_r^3} \quad (2.25)$$

$$E = a_{10} + \frac{a_{11}}{T_r^2} + \frac{a_{12}}{T_r^3} \quad (2.26)$$

$$F = \frac{\alpha}{T_r^3} \quad (2.27)$$

$$T_r = \frac{T}{T_c} \quad (2.28)$$

$$V_c = \frac{RT_c}{P_c} \quad (2.29)$$

where the a_i coefficients are tabulated values shown in Table 4 of [Duan and Zhang \(2006\)](#); T_c and P_c are the temperature and pressure at the critical point of water, which are tabulated as $T_c = 647.25$ K and $P_c = 221.19$ bar.

2.2.1.3. Mazevet formulation

The EOS presented in [Mazevet et al. \(2019\)](#) consists on a fit to the experimental data from [Wagner and Pruß \(2002\)](#), and molecular quantum simulation data for pressure and temperature conditions outside the experimental range. For the high-pressure region, the simulations consider first-principles quantum molecular dynamics and the Thomas-Fermi extension limit. This region extends from $T = 1000$ to 10^5 K and $\rho = 1$ to 100 g/cm³, which corresponds to pressures of 100-150 GPa approximately (see Fig. 4 in [Mazevet et al. 2019](#)). These ranges comprise the plasma and superionic phases. The Fortran implementation of this EOS is publicly available in the Strasbourg astronomical Data Center (CDS)¹.

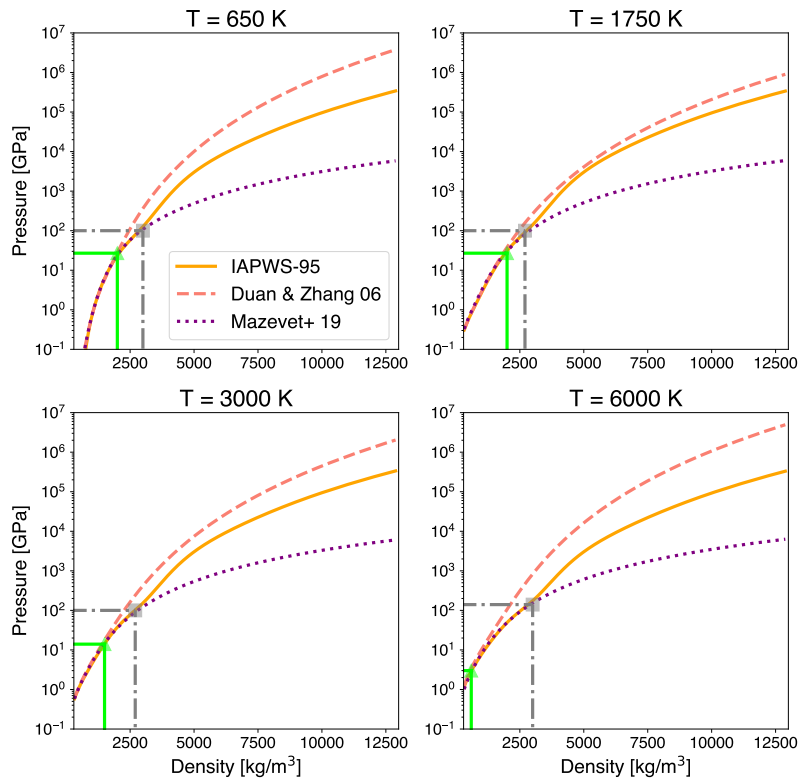


Figure 2.6.: Supercritical EOS comparison for four different constant temperatures: 650, 1750, 3000 and 6000 K. Green triangles indicate when DZ06 starts to depart noticeably from the other two EOS, whereas grey squares mark the pressure and density at which IAPWS-95 and M19 begin to differ significantly.

¹<http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/621/A128>

Figure 2.6 shows a comparison of the three EOSs for four different constant temperatures. At $T = 650$ K (upper left panel), where all EOSs are within their temperature validity range, and experimental data is available up to 1 GPa, we observe that the DZ06 EOS starts to deviate noticeably from the other two EOS at a pressure of ~ 27 GPa. This is close to the maximum validity pressure established by Duan et al. (1996), which is 35 GPa. In contrast, the IAPWS-95 and M19 are in good agreement up to 100 GPa throughout the whole temperature range. We do not have experimental data available at $T = 1750$ K, although all three EOSs are within their validity or extrapolation range, maintaining the maximum pressure at which they agree well. Nonetheless, at $T = 3000$ K the DZ06 EOS is not valid, and starts to differ from the other two EOS at a pressure of 14 GPa. Finally, at $T = 6000$ K only the M19 EOS is still applicable, which agrees well with the IAPWS-95 EOS up to ~ 140 GPa, while DZ06 starts showing a significant disagreement at pressures above 1 to 3 GPa.

2.2.2. Grüneisen parameter

As seen in Eq. 2.5, the Grüneisen parameter depends on the derivative of the pressure as a function of energy. For each EOS formulation, the dependence of the pressure on the internal energy is different. Therefore, in the following subsections I describe the calculation of the Grüneisen parameter for the IAPWS-95 and M19 formulations. Duan and Zhang (2006) do not provide the relation between pressure and internal energy, only its relation to the density and temperature.

2.2.2.1. IAPWS-95 formulation

The internal energy per unit volume is $u = E/V$, which is related to the Grüneisen parameter as:

$$\gamma = V \left(\frac{\partial p}{\partial u} \right)_V \left(\frac{\partial u}{\partial E} \right)_V = V \left(\frac{\partial p}{\partial u} \frac{1}{V} \right)_V = \left(\frac{\partial p}{\partial u} \right)_V \quad (2.30)$$

We can decompose the partial derivative of the pressure with respect to u in two terms, knowing that $p = p(\rho, T)$:

$$\frac{\partial p}{\partial u} = \frac{\partial p}{\partial T} \frac{\partial T}{\partial u} + \frac{\partial p}{\partial \rho} \frac{\partial \rho}{\partial u} \quad (2.31)$$

As indicated in Eqs. 2.5 and 2.30, the partial derivative that defines the Grüneisen parameter is performed at constant volume. Mass is also constant due to the mass conservation principle, which together with constant volume yields a constant density. Therefore, the derivative of the density is equal to zero, resulting in:

$$\gamma = \left(\frac{\partial p}{\partial u} \right)_V = \left(\frac{\partial p}{\partial T} \right)_V \left(\frac{\partial T}{\partial u} \right)_V \quad (2.32)$$

In the IAPWS-95 formulation, the pressure and the internal energy are defined as functions of the Helmholtz free energy f (see Eq. 2.9):

$$p = \rho^2 \left(\frac{\partial f}{\partial \rho} \right)_T \quad (2.33)$$

$$u = f - T \left(\frac{\partial f}{\partial T} \right)_\rho \quad (2.34)$$

The partial derivatives in Eq. 2.32 can then be calculated as functions of the Helmholtz free energy by deriving Eqs. 2.33 and 2.34:

$$\frac{\partial p}{\partial T} = \rho^2 \frac{\partial}{\partial T} \left[\left(\frac{\partial f}{\partial \rho} \right)_T \right] \quad (2.35)$$

$$\frac{\partial u}{\partial T} = \frac{\partial f}{\partial T} - \left[\left(\frac{\partial f}{\partial T} \right)_\rho + T \frac{\partial}{\partial T} \left(\frac{\partial f}{\partial T} \right)_\rho \right] = -T \left(\frac{\partial^2 f}{\partial T^2} \right)_\rho \quad (2.36)$$

We therefore need the derivative of the Helmholtz free energy with respect to the density. For this, we derive Eq. 2.9:

$$\frac{\partial f}{\partial \rho} = RT \left(\frac{\partial \phi^0}{\partial \rho} + \frac{\partial \phi^r}{\partial \rho} \right) = \frac{RT}{\rho_c} (\phi_\delta^0 + \phi_\delta^r) \quad (2.37)$$

$$\text{since } \frac{\partial \phi^0}{\partial \rho} = \frac{\partial \phi^0}{\partial \delta} \frac{\partial \delta}{\partial \rho} = \frac{\phi_\delta^0}{\rho_c} \text{ and } \frac{\partial \phi^r}{\partial \rho} = \frac{\partial \phi^r}{\partial \delta} \frac{\partial \delta}{\partial \rho} = \frac{\phi_\delta^r}{\rho_c}.$$

See immediately after Eq. 2.9 for the definition of δ and the density at the critical point of water. ϕ_δ^r was previously defined as well in Eqs. 2.11 to 2.20.

We proceed to calculate the derivative of the Helmholtz free energy with respect to the temperature by differentiating Eq. 2.9:

$$\frac{\partial f}{\partial T} = R(\phi^0 + \phi^r) + RT \left(\frac{\partial \phi^0}{\partial T} + \frac{\partial \phi^r}{\partial T} \right) \quad (2.38)$$

We recall that the ideal-gas and residual terms of f depend on T via τ , so $\frac{\partial \phi^0}{\partial T} = \frac{\partial \phi^0}{\partial \tau} \frac{\partial \tau}{\partial T} = \frac{-T_c}{T^2} \phi_\tau^0$ and $\frac{\partial \phi^r}{\partial T} = \frac{\partial \phi^r}{\partial \tau} \frac{\partial \tau}{\partial T} = \frac{-T_c}{T^2} \phi_\tau^r$.

We substitute these expressions in Eq. 2.38:

$$\frac{\partial f}{\partial T} = R(\phi^0 + \phi^r) - R\tau (\phi_\tau^0 + \phi_\tau^r) \quad (2.39)$$

We differentiate Eq. 2.37 to obtain the second partial derivative of the Helmholtz free energy with respect to the temperature and the density, which is necessary for Eq. 2.35:

$$\frac{\partial}{\partial T} \left[\frac{\partial f}{\partial \rho} \right] = \frac{\partial}{\partial T} \left[\frac{RT}{\rho_c} (\phi_\delta^0 + \phi_\delta^r) \right] = \frac{R}{\rho_c} (\phi_\delta^0 + \phi_\delta^r) + \frac{RT}{\rho_c} \left(\frac{-T_c}{T^2} \right) (\phi_{\delta\tau}^0 + \phi_{\delta\tau}^r) \quad (2.40)$$

By definition of ϕ^0 in Wagner and Pruß (2002), the second partial derivatives of the ideal-gas term are zero, $\phi_{\tau\delta}^0 = \phi_{\delta\tau}^0 = 0$. Consequently, Eq. 2.40 is simplified as:

$$\frac{\partial}{\partial T} \left(\frac{\partial f}{\partial \rho} \right)_T = \frac{R}{\rho_c} (\phi_\delta^0 + \phi_\delta^r - \tau \phi_{\delta\tau}^r) \quad (2.41)$$

The second derivative of the Helmholtz free energy with respect to the temperature is calculated by deriving Eq. 2.39:

$$\frac{\partial^2 f}{\partial T^2} = R(\phi_\tau^0 + \phi_\tau^r) \left(\frac{-T_c}{T^2} \right) - \left[R \left(\frac{-T_c}{T^2} \right) (\phi_\tau^0 + \phi_\tau^r) + R\tau \left(\frac{-T_c}{T^2} \right) (\phi_{\tau\tau}^0 + \phi_{\tau\tau}^r) \right] =$$

$$= \frac{R\tau^2}{T} (\phi_{\tau\tau}^0 + \phi_{\tau\tau}^r) \quad (2.42)$$

The isochoric heat capacity, c_v , can be defined as a function of the partial derivatives of ϕ (see Table 3 in [Wagner and Pruß 2002](#)):

$$\frac{c_v(\delta, \tau)}{R} = -\tau^2 (\phi_{\tau\tau}^0 + \phi_{\tau\tau}^r) \quad (2.43)$$

If we substitute Eq. 2.43 in Eq. 2.42, we obtain:

$$\frac{\partial^2 f}{\partial T^2} = -\frac{c_v}{T} \quad (2.44)$$

We finally have the partial derivatives of the Helmholtz free energy necessary to calculate the partial derivatives of the pressure and the internal energy. We substitute Eq. 2.40 in Eq. 2.35:

$$\frac{\partial p}{\partial T} = \rho^2 \frac{R}{\rho_c} [\phi_{\delta}^0 + \phi_{\delta}^r - \tau \phi_{\delta\tau}^r] \quad (2.45)$$

Similarly, with Eqs. 2.44 and 2.36:

$$\frac{\partial u}{\partial T} = -T \left(\frac{c_v}{T} \right) = c_v \quad (2.46)$$

We substitute Eqs. 2.45 and 2.46 in Eq. 2.32 to express the Grüneisen parameter as a function of the Helmholtz free energy:

$$\gamma = \rho^2 \frac{R}{\rho_c} [\phi_{\delta}^0 + \phi_{\delta}^r - \tau \phi_{\delta\tau}^r] \frac{1}{c_v} \quad (2.47)$$

This expression can be further simplified by using the definitions of ϕ_{δ}^0 and ϕ_{δ}^r . The derivative of the ideal-gas term with respect to δ is computed as $\phi_{\delta}^0 = 1/\delta$ ([Wagner and Pruß 2002](#)). Thus, the sum of the two derivatives is (see Eq. 2.10):

$$\phi_{\delta}^0 + \phi_{\delta}^r = \frac{1}{\delta} + \phi_{\delta}^r = \frac{1}{\delta} (1 + \delta \phi_{\delta}^r) = \frac{1}{\delta} \frac{p}{\rho RT} \quad (2.48)$$

Eventually, substituting Eq. 2.48 in Eq. 2.47 yields:

$$\gamma = \frac{R}{c_v} \left(\frac{p(\rho, T)}{\rho RT} - \tau \delta \phi_{\delta\tau}^r \right) \quad (2.49)$$

The pressure, $p(\rho, T)$, is obtained with the EOS (see Section 2.2.1). The other three partial derivatives necessary for the calculation of the Grüneisen parameter according to Eq. 2.49 are $\phi_{\tau\tau}^0$, $\phi_{\tau\tau}^r$ and $\phi_{\delta\tau}^r$. These are calculated as a function of the coefficients tabulated in [Wagner and Pruß \(2002\)](#):

$$\phi_{\tau\tau}^0 = -n_3^0/\tau^2 - \sum_{i=4}^8 n_i^0 (\gamma_i^0)^2 e^{-\gamma_i^0 \tau} (1 - e^{-\gamma_i^0 \tau})^{-2} \quad (2.50)$$

Similarly to ϕ_{δ}^r (Eq. 2.11), $\phi_{\tau\tau}^r$ and $\phi_{\delta\tau}^r$ are expressed as the sum of four terms each:

$$\phi_{\tau\tau}^r = \phi_{\tau\tau,1}^r + \phi_{\tau\tau,2}^r + \phi_{\tau\tau,3}^r + \phi_{\tau\tau,4}^r \quad (2.51)$$

$$\phi_{\delta\tau}^r = \phi_{\delta\tau,1}^r + \phi_{\delta\tau,2}^r + \phi_{\delta\tau,3}^r + \phi_{\delta\tau,4}^r \quad (2.52)$$

These terms are computed as:

$$\phi_{\tau\tau,1}^r = \sum_{i=1}^7 n_i t_i (t_i - 1) \delta^{d_i} \tau^{t_i-2} \quad (2.53)$$

$$\phi_{\tau\tau,2}^r = \sum_{i=8}^{51} n_i t_i (t_i - 1) \delta^{d_i} \tau^{t_i-2} e^{-\delta^{c_i}} \quad (2.54)$$

$$\phi_{\tau\tau,3}^r = \sum_{i=52}^{54} n_i \delta^{d_i} \tau^{t_i} e^{-\alpha_i(\delta-\varepsilon_i)^2 - \beta_i(\tau-\gamma_i)^2} \left[\left(\frac{t_i}{\tau} - 2\beta_i(\tau-\gamma_i) \right)^2 - \frac{t_i}{\tau^2} - 2\beta_i \right] \quad (2.55)$$

$$\phi_{\tau\tau,4}^r = \sum_{i=55}^{56} n_i \delta \left[\frac{\partial^2 \Delta^{b_i}}{\partial \tau^2} \Psi + 2 \frac{\partial \Delta^{b_i}}{\partial \tau} \frac{\partial \Psi}{\partial \tau} + \Delta^{b_i} \frac{\partial^2 \Psi}{\partial \tau^2} \right] \quad (2.56)$$

$$\phi_{\delta\tau,1}^r = \sum_{i=1}^7 n_i d_i t_i \delta^{d_i-1} \tau^{t_i-1} \quad (2.57)$$

$$\phi_{\delta\tau,2}^r = \sum_{i=8}^{51} n_i t_i \delta^{d_i-1} \tau^{t_i-1} (d_i - c_i \delta^{c_i}) e^{-\delta^{c_i}} \quad (2.58)$$

$$\phi_{\delta\tau,3}^r = \sum_{i=52}^{54} n_i \delta^{d_i} \tau^{t_i} e^{-\alpha_i(\delta-\varepsilon_i)^2 - \beta_i(\tau-\gamma_i)^2} \left[\frac{d_i}{\delta} - \alpha_i(\delta-\varepsilon_i) \right] \left[\frac{t_i}{\tau} - 2\beta_i(\tau-\gamma_i) \right] \quad (2.59)$$

$$\phi_{\delta\tau,4}^r = \sum_{i=55}^{56} n_i \left[\Delta^{b_i} \left(\frac{\partial \Psi}{\partial \tau} + \delta \frac{\partial^2 \Psi}{\partial \delta \partial \tau} \right) + \delta \frac{\partial \Delta^{b_i}}{\partial \delta} \frac{\partial \Psi}{\partial \tau} + \frac{\partial \Delta^{b_i}}{\partial \tau} \left(\Psi + \delta \frac{\partial \Psi}{\partial \delta} \right) + \frac{\partial^2 \Delta^{b_i}}{\partial \delta \partial \tau} \delta \Psi \right] \quad (2.60)$$

The functions Δ and Ψ and their first partial derivatives with respect to δ were presented in Eqs. 2.17 to 2.19. The other first and second partial derivatives necessary for the calculation of the Grüneisen parameter are:

$$\frac{\partial \Delta^{b_i}}{\partial \tau} = -2\theta b_i \Delta^{b_i-1} \quad (2.61)$$

$$\frac{\partial^2 \Delta^{b_i}}{\partial \tau^2} = 2b_i \Delta^{b_i-1} + 4\theta^2 b_i (b_i - 1) \Delta^{b_i-2} \quad (2.62)$$

$$\frac{\partial^2 \Delta^{b_i}}{\partial \delta \partial \tau} = -A_i b_i \frac{2}{\beta_i} \Delta^{b_i-1} [(\delta-1)^2] \frac{1}{2\beta_i}{}^{-1} - 2\theta b_i (b_i - 1) \Delta^{b_i-2} \frac{\partial \Delta}{\partial \delta} \quad (2.63)$$

$$\frac{\partial \Psi}{\partial \tau} = -2D_i(\tau-1)\Psi \quad (2.64)$$

$$\frac{\partial^2 \Psi}{\partial \tau^2} = \{2D_i(\tau-1)^2 - 1\} 2D_i \Psi \quad (2.65)$$

$$\frac{\partial^2 \Psi}{\partial \delta \partial \tau} = 4C_i D_i (\delta-1)(\tau-1)\Psi \quad (2.66)$$

2.2.2.2. Mazevet formulation

The Fortran subroutine made available by [Mazevet et al. \(2019\)](#) provides both the pressure and the internal energy per unit volume as functions of the temperature and the density. To calculate the Grüneisen parameter at a given temperature T and density ρ , we define two values $T_1 = T - \Delta/2$ and $T_2 = T + \Delta/2$, where Δ is a constant value small enough to allow us to calculate an infinitesimal increase or decrease of the temperature. The Grüneisen parameter is then computed as:

$$\gamma(T, \rho) = \left(\frac{\partial p}{\partial u} \right)_V = \frac{p(T_2, \rho) - p(T_1, \rho)}{u(T_2, \rho) - u(T_1, \rho)} \quad (2.67)$$

Fig. 2.7 shows a comparison of the Grüneisen parameter calculated with the IAPWS-95 and M19 formulations for four temperatures. At $T = 650$ K, which is in the proximity of the critical point of water, both formulations present non-monotonic or asymptotic behaviours. In the case of the M19 formulation, the range of densities at which the non-monotonic behaviour occurs is close to the ice VII and ice X phases, where this formulation cannot be applied ([Mazevet et al. 2019](#)). For the IAPWS-95, the asymptotic behaviour appears outside of its extrapolation range, at $\rho > 3000$ kg/m³. At higher temperatures, the Grüneisen parameter in the M19 formulation increases monotonically with increasing density.

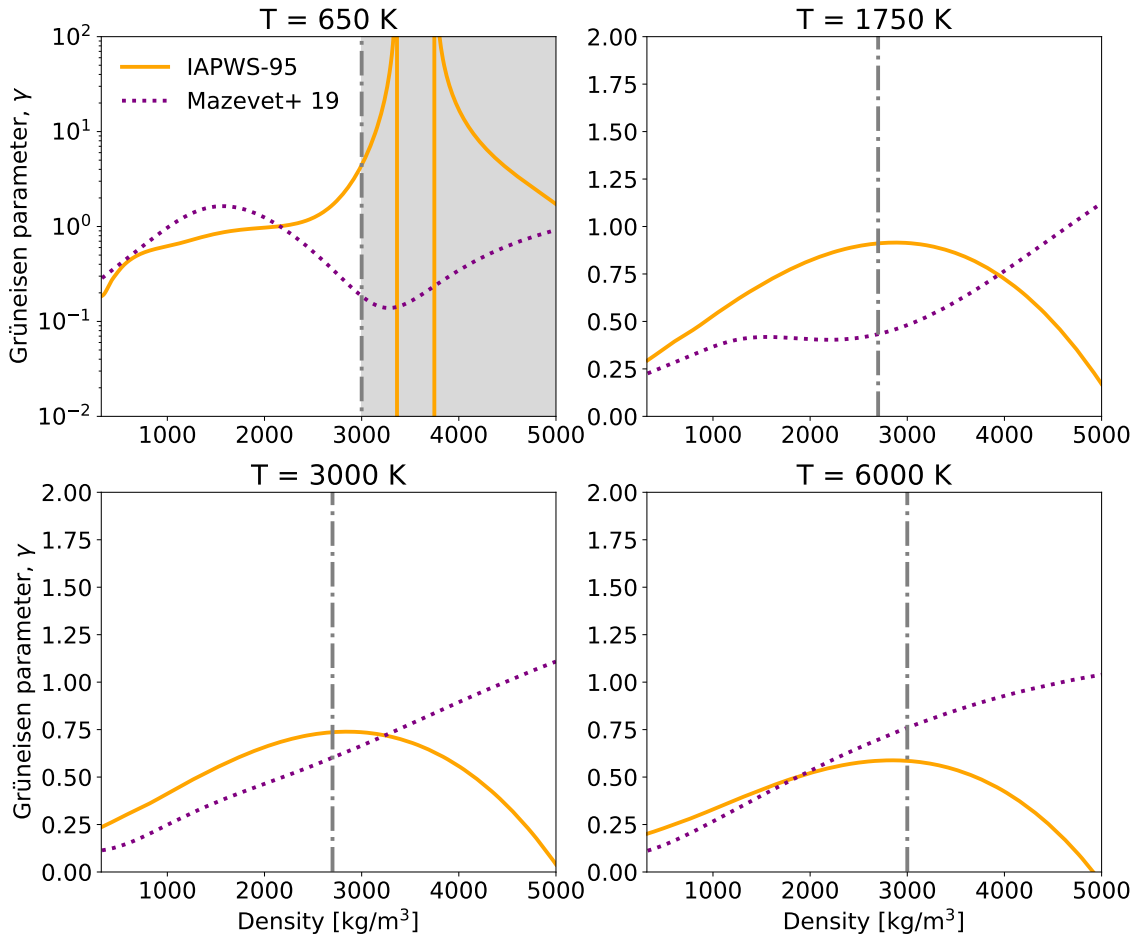


Figure 2.7.: Grüneisen parameter comparison for four different constant temperatures: 650, 1750, 3000 and 6000 K. Grey dashed-dotted lines and shaded areas mark the densities at which the IAPWS-95 formulation is outside of its extrapolation validity range.

Fig. 2.8 shows the temperature profile up to the highest pressures reached in the planetary mass regime we consider (up to $20 M_{\oplus}$). We assume three different equilibrium temperatures for both Grüneisen parameter formulations. We can see that the IAPWS-95 profile is more steep than that of the M19 formulation. This is expected from Fig. 2.7, where γ in the IAPWS-95 formulation is higher than in M19 at high temperatures ($T \geq 1750$ K) within its extrapolation range. Therefore, the IAPWS-95 formulation of the Grüneisen parameter will produce warmer interiors at a similar pressure compared to the M19 formulation.

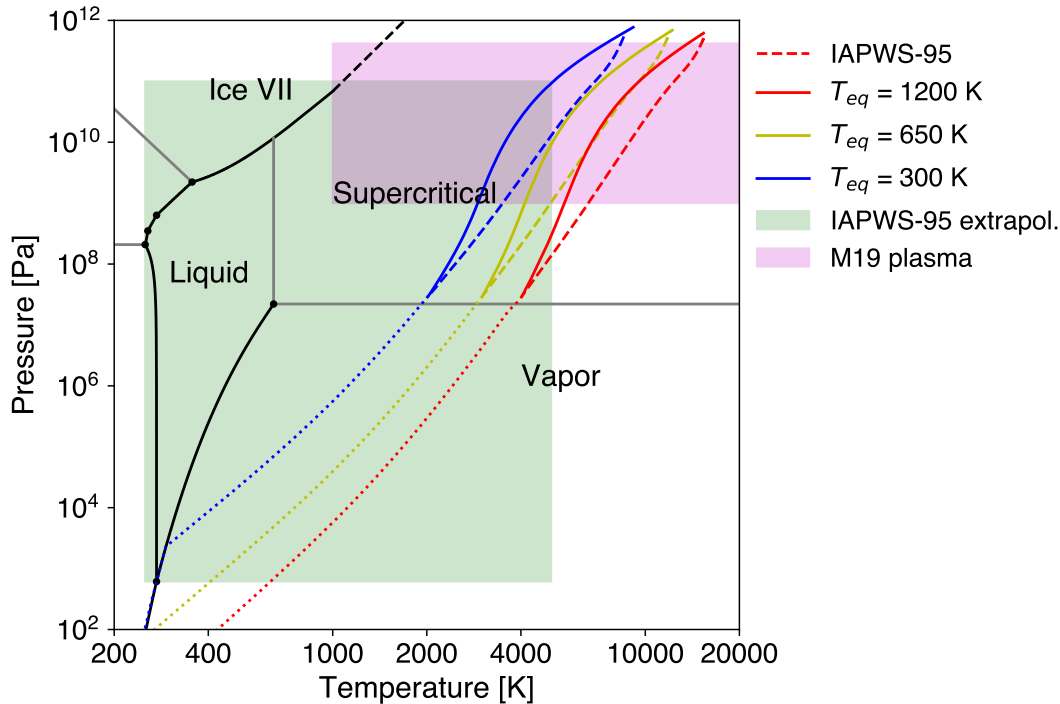


Figure 2.8.: Water phase diagram with adiabatic profiles of IAPWS-95 (solid) and M19 (dashed) formulations for three different equilibrium temperatures: $T_{eq} = 1200$ K (red), 650 K (yellow) and 300 K (blue). We assumed a theoretical 100% water composition, for a planet of $M = 15 M_{\oplus}$. The temperature-pressure profiles of the atmosphere calculated by the model in [Marcq et al. \(2017\)](#) are also shown (dotted). The green shaded area indicates the extrapolation range of the IAPWS-95 formulation, whereas the magenta shaded area marks the plasma validity region of the M19 formulation.

2.3. Mass-radius relations

Fig. 2.9 (top panel) shows a comparison of the mass-radius relations of the three formulations for four different compositions. In the case of the DZ06 EOS, we consider the IAPWS-95 prescription of the Grüneisen parameter described in Sect. 2.2.2.1, since [Duan and Zhang \(2006\)](#) do not provide the internal energy as a function of the density and temperature. We can see that in comparison with M19, who presents the broadest validity region of the three, the IAPWS-95 and DZ06 obtain higher total radii for a similar mass and composition. This overestimation of the radius occurs because for a given pressure and constant temperature, the IAPWS-95 and DZ06 EOSs underestimate the density (see Fig. 2.6), which yields more extended hydrospheres.

We calculate the absolute difference between these two formulations and M19:

$$\Delta R_{IAPWS-95/DZ06} = R_{IAPWS-95/DZ06} - R_{M19} \quad (2.68)$$

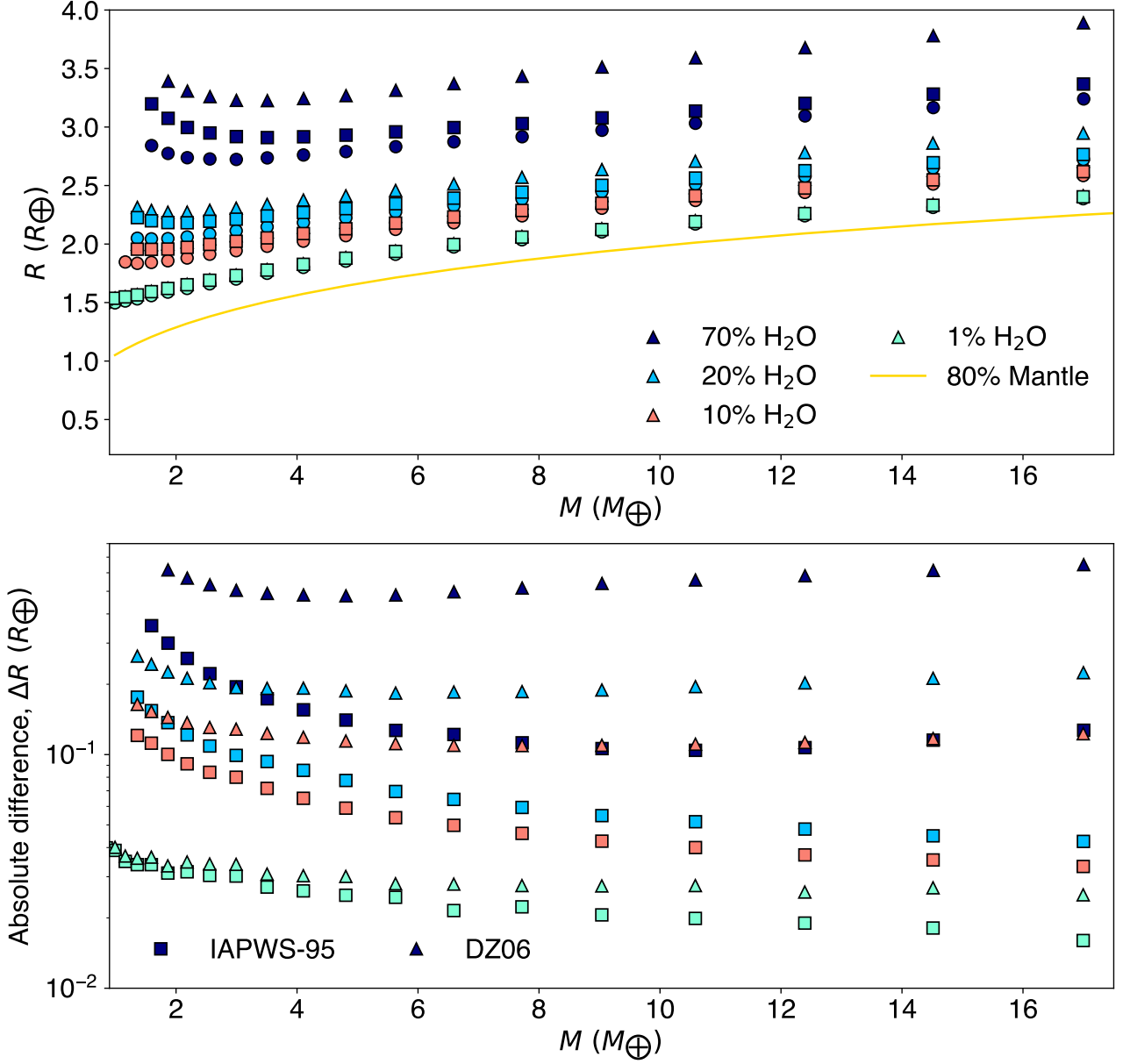


Figure 2.9.: Top panel: Mass-radius relationships for supercritical water planets with three different formulations: IAPWS-95 (squares), [Duan and Zhang \(2006\)](#) (DZ06, triangles) and [Mazevet et al. \(2019\)](#) (M19, circles). We assume a core composition of 100% mantle, and an equilibrium temperature of 1200 K. The atmospheric contribution to the total radius has been calculated with the model from [Marcq et al. \(2017\)](#), as described in [Acuña et al. \(2021\)](#) and Chapter 3. Bottom panel: Absolute difference (see text) between the IAPWS-95 (squares) and DZ06 (triangles) formulation and M19 as a function of the planetary mass.

ΔR_{DZ06} is always greater than $\Delta R_{IAPWS-95}$ (see Fig. 2.9, bottom panel), which is due to the DZ06 EOS having a greater difference with the M19 EOS for the same density and temperature than the IAPWS-95 EOS. The ΔR of both formulations increases with increasing WMF. Assuming that the

typical uncertainty in radius for a low-mass planet is $\Delta R \simeq 0.1 R_{\oplus}$, the IAPWS-95 exceeds this limit in absolute difference for planets whose WMF $\geq 20\%$, whereas DZ06 presents a difference greater than this value at WMF = 10%. The relative differences to the total planetary radius are between 5% and 20%. This is in agreement with the results of [Aguichine et al. \(2021\)](#), who also found that the IAPWS-95 and DZ06 formulations overestimate the total planetary radius, being this overestimation greater for higher water mass fractions. We can conclude that the total radius of water-dominated planets is sensitive to the choice of the EOS and adiabatic profile formulations. For this reason, it is necessary to take into account a formulation whose validity range includes the interior conditions of low-mass planets, which in our case is [Mazevet et al. \(2019\)](#).

3. Interior-atmosphere coupling

Highly-irradiated planets have a significant part of their radius constituted by their envelope. A complete water layer is comprised of an upper low-pressure atmosphere, and a high-pressure water layer, where the supercritical phase is reached. The thickness of the upper atmosphere depends on the temperature profile, since steam will tend to expand more at higher temperatures. The supercritical water layer underneath will also expand if the temperature at the bottom of the upper atmosphere is high too. Some interior structure models of volatile-rich planets have assumed isothermal upper atmospheric profiles with convection but with a simple implementation of radiative transfer (Dorn et al. 2017, 2018; Zeng et al. 2019; Thomas and Madhusudhan 2016). Nonetheless, radiative transfer is necessary in interior models to compute self-consistently the planetary energy balance and its effect on the transit radius (Nettelmann et al. 2011). Turbet et al. (2019) found that when considering radiative transfer, the atmosphere of highly-irradiated planets is expanded compared to their less irradiated counterparts. This effect, named runaway greenhouse radius inflation effect, is due to an increase in the atmospheric temperature as a function of altitude as well as an increase in atmospheric thickness.

To include the contribution of an atmosphere to the total radius, I couple our interior model for supercritical planets (Chapter 2) to a 1D atmosphere model. We first couple the interior model with the atmospheric model initially presented in Marcq (2012) and Marcq et al. (2017), and further developed by Pluriel et al. (2019), which considers radiative transfer and adiabatic profiles for water-dominated atmospheres. Later on, I substitute it by the atmosphere model detailed in Chapter 4. In the present chapter, I describe the algorithm used to couple self-consistently the interior and atmosphere models for supercritical planets in Sect. 3.1. In Sect. 3.2, I explain the trends of the atmospheric parameters (surface temperature, thickness, mass and Bond albedo) with surface pressure.

3.1. Coupling algorithm

Within our coupled interior-atmosphere model, if the surface pressure is below the 300 bar level, the atmosphere and the interior are coupled at the atmosphere-mantle interface and water is always in gaseous phase. If the surface pressure is above 300 bar, the atmosphere and the interior models are coupled at this pressure value. Underneath the upper atmosphere, a convective supercritical water layer extends down to the mantle. We choose the 300 bar pressure level to couple the interior and the atmosphere to be sufficiently close to the critical point of water ($P_{crit} = 220$ bar). This prevents the atmospheric model from taking over pressures where convection dominates over radiative transfer. The pressure at the top of the atmosphere corresponds to the observable transiting radius, which is $P = 20$ mbar (Grimm et al. 2018; Mousis et al. 2020).

The output parameters of the atmospheric model are the outgoing longwave radiation (OLR), Bond albedo, and atmospheric thickness and mass. The OLR represents the emitted energy in infrared (IR), while the Bond albedo corresponds to the percentage of light reflected by the planet coming from its host star. The input variables are the bulk mass and radius, and the temperature

at the bottom of the atmosphere. We define the bulk mass and radius, R_{bulk} and M_{bulk} , as the mass and radius comprised from the center of the planet, up to the bottom of the atmosphere. In our interior-atmosphere coupling, this level corresponds to the atmosphere-mantle interface for planets whose surface pressure is less than 300 bar, whereas for planets with more massive envelopes, this level corresponds to the top of the interior supercritical layer. Both the OLR and the Bond albedo are necessary to determine if an atmosphere is in radiative-convective equilibrium. If this is the case, the OLR, which represents the emitted energy in the infrared (IR), is equal to the radiation absorbed by the planet from its host star, F_{abs} . This is calculated as:

$$F_{abs} = \sigma T_{eq}^4 \quad (3.1)$$

where $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$ is the Stefan-Boltzmann constant, and T_{eq} is the equilibrium temperature.

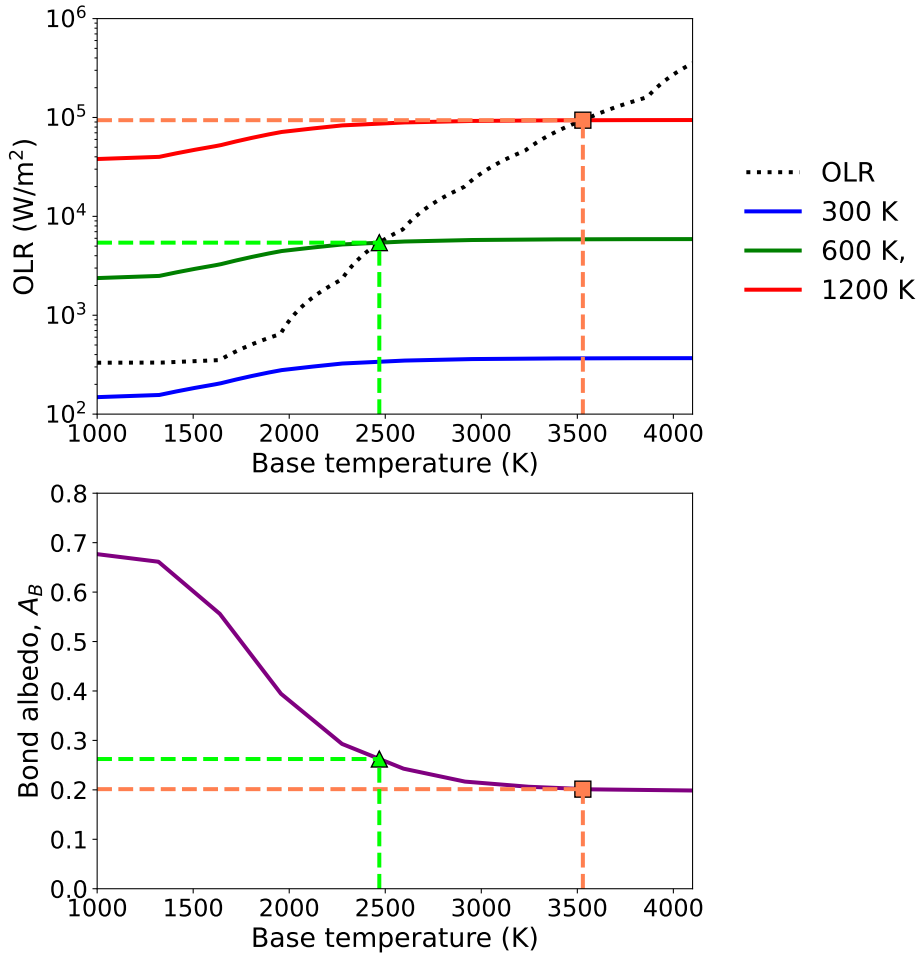


Figure 3.1.: Top panel: OLR (dotted) and absorbed flux (solid) for a planet of $M_{bulk} = 5 M_{\oplus}$ and $R_{bulk} = 1 R_{\oplus}$ at different equilibrium temperatures: 300 K, 600 K and 1200 K. The square and triangle indicate the base temperature and OLR at which the atmosphere is in radiative-convective equilibrium for $T_{eq} = 600$ and 1200 K, respectively. We consider a water-dominated atmospheric composition and a base pressure of 300 bar. Bottom panel: Bond albedo for a planet with $M_{bulk} = 5 M_{\oplus}$ and $R_{bulk} = 1 R_{\oplus}$. We assumed a Sun-like star with $T_{\star} = 5777$ K. The square and triangle mark the base temperature and Bond albedo at which radiative-convective equilibrium is reached.

The planet's equilibrium temperature, T_{eq} , is defined as:

$$T_{eq} = (1 - A_B)^{0.25} \left(\frac{R_\star}{2a_d} \right)^{0.5} T_\star \quad (3.2)$$

where A_B is the Bond albedo, a_d is the planetary semi-major axis, and R_\star and T_\star are the host stellar radius and effective temperature, respectively.

Given two certain values of the bulk mass and radius, the atmospheric model obtains the OLR and Bond albedo as functions of the temperature at the bottom of the atmosphere, T_{base} . We solve for radiative-convective equilibrium by finding the temperature at which the following function, f , is zero with a root-finding method (in our case, the bisection method):

$$f(T_{base}) = OLR(T_{base}) - F_{abs}(A_B(T_{base})) = 0 \quad (3.3)$$

where the absorbed flux depends on T_{base} via the Bond albedo as seen in Eqs. 3.1 and 3.2.

Figure 3.1 shows the OLR and absorbed flux of three planets with similar bulk mass and radius, but three different equilibrium temperatures. We can see that the OLR is independent of the equilibrium temperature, since it depends on the surface conditions, the bulk mass and radius, and the composition of the atmosphere; whereas the absorbed flux also depends on the equilibrium temperature. The base temperatures at which the two warmest planets are in radiative-convective equilibrium are computed by finding the root of Eq. 3.3 with the bisection method. These temperatures correspond to $T_{base} = 2469$ K and 3529 K for the planets with equilibrium temperatures of 600 K and 1200 K, respectively. For the planet with $T_{eq} = 300$ K, its absorbed radiation never exceeds its OLR, which means that this planet would be cooling down. In order to reach radiative-convective equilibrium, it would be necessary to supply an internal heat flux of $F_{int} = OLR - F_{abs} \simeq 180$ W/m². Once the base temperature is obtained for a given set of bulk mass and radius and equilibrium temperature, the atmospheric thickness is calculated as $z_{atm} = z_{atm}(T_{base})$.

For the same irradiation conditions, we obtain the base temperature, Bond albedo and atmospheric thickness as functions of the bulk mass and radius. This is done with three 2-dimensional grids (one grid per output parameter of the atmospheric model) where the two dimensions are the bulk mass and radius. The three parameters are interpolated by bilinear interpolation:

$$f(M, R) = \frac{1}{(M_2 - M_1)(R_2 - R_1)} [f_{11}(M_2 - M)(R_2 - R) + f_{21}(M - M_1)(R_2 - R) + f_{12}(M_2 - M)(R - R_1) + f_{22}(M - M_1)(R - R_1)] \quad (3.4)$$

where M and R are the bulk mass and radius at which we want to interpolate. The tabulated values of the mass and radius are M_1 , M_2 , R_1 and R_2 , where $M_1 < M < M_2$, and $R_1 < R < R_2$. The constant values f_{ij} are the function we wish to interpolate evaluated at the tabulated masses and radii: $f_{11} = f(M_1, R_1)$, $f_{12} = f(M_1, R_2)$, $f_{21} = f(M_2, R_1)$ and $f_{22} = f(M_2, R_2)$. In our case, the function $f(M, R)$ are T_{surf} , A_B or z_{atm} .

Figure 3.2 shows the corresponding grids for the surface temperature and Bond albedo of a planet with $T_{eq} = 1200$ K. The dashed line indicates the limit of surface gravity $g_{surf} = 2$ m/s², which is the lower limit established by the atmospheric model of [Pluriel et al. \(2019\)](#) for a planet to retain its atmosphere via hydrostatic equilibrium. We observe that the surface temperature and Bond albedo are higher for planets whose bulk is less dense (high radius and low mass), and lower for more dense bulks (low radius and high mass). For low bulk densities, the atmosphere is more extended, which becomes more optically thick, having a higher reflectance and being less efficient at irradiating the heat from the warmer inner layers.

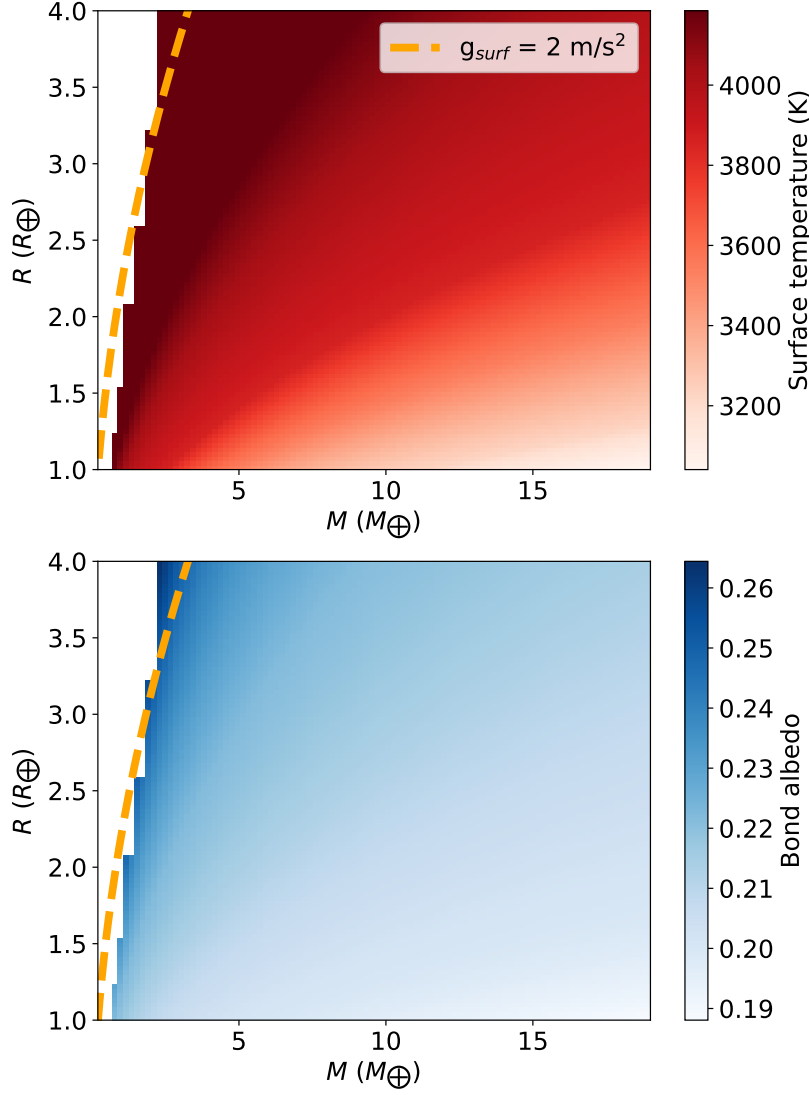


Figure 3.2.: Top panel: Surface temperature as a function of bulk mass and radius for a planet with a water-dominated atmosphere and $T_{eq} = 1200 \text{ K}$ in radiative-convective equilibrium. We assume a base pressure of 300 bar. Bottom panel: Bond albedo as a function of bulk mass and radius for $T_{eq} = 1200 \text{ K}$. We assumed a Sun-like star with $T_{\star} = 5777 \text{ K}$.

Since the bulk radius, which is an input to the atmospheric model grids, is an output of the interior model, we first need to compute the base temperature for a set of mass, WMF and CMF values with an initial guess bulk radius. The recently calculated base temperature is the input surface temperature for the interior model, which obtains a new bulk radius that can be the input for the atmospheric model. We iterate this scheme until the bulk radius converges to a constant value. This is checked by calculating the difference between the new bulk radius, $R_{interior}$, and the guess bulk radius, R_{bulk} . If this difference is below a given tolerance value, we consider convergence has been reached. We set this tolerance to a maximum value of $0.02 R_{\oplus}$, which is usually reached in 1-3 iterations. This iterative algorithm is shown in Fig. 3.3. Once convergence is reached, the total radius of the planet is calculated as the sum of the atmospheric thickness and the latest output of the interior model, $R_{interior}$. Similarly, the total mass of the planet is updated as the sum of the bulk mass M_{bulk} , and the atmospheric mass M_{atm} .

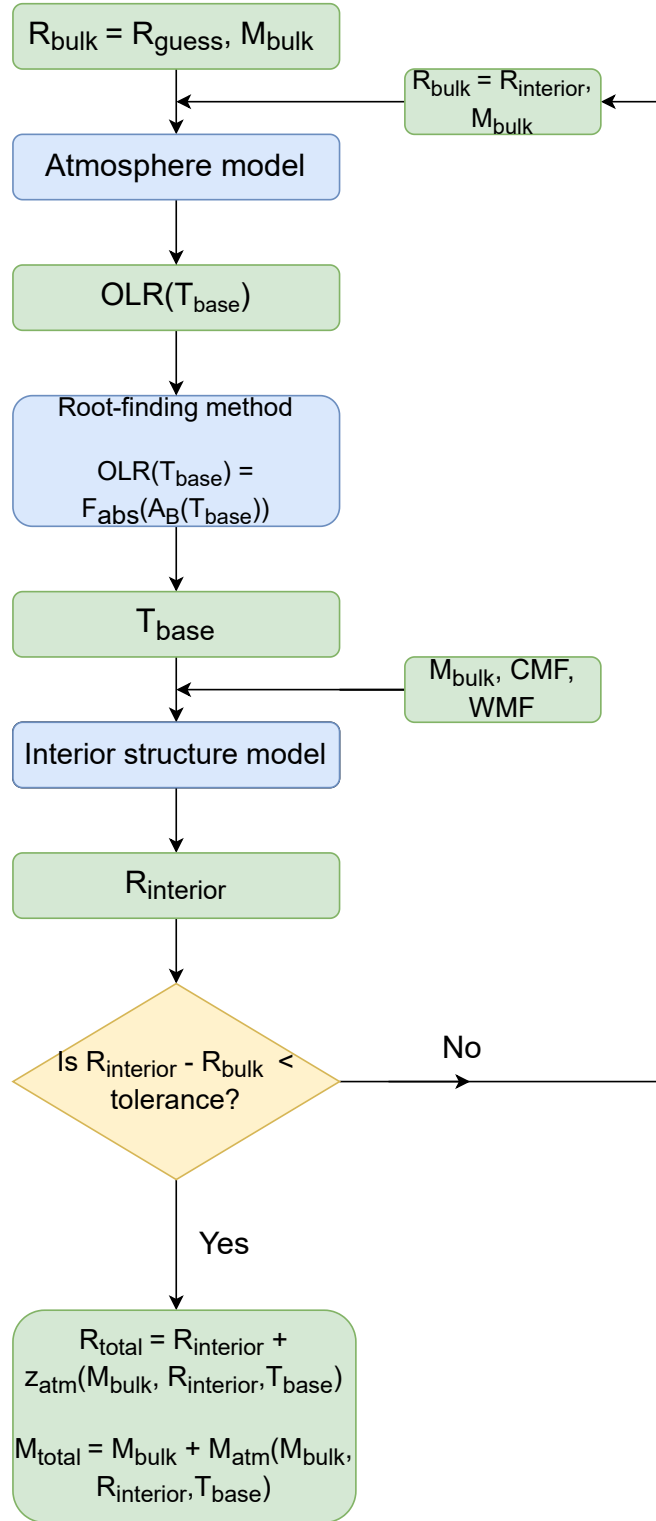


Figure 3.3.: Diagram of the interior-atmosphere coupling algorithm. T_{base} denotes the base temperature at the bottom of the atmosphere, while z_{atm} and M_{atm} correspond to the atmospheric thickness and mass, respectively. R_{bulk} and M_{bulk} are the bulk radius and mass, while R_{guess} is the initial of the bulk radius. $R_{interior}$ refers to the output bulk radius of the interior model in each iteration. Green boxes indicate the input of the previous step and output of the following one, while blue boxes corresponds to calculations and models. The yellow box is a conditional operation that represents a decision (yes/no).

3.2. Low surface pressures

In Sect. 3.1, I have always considered that the surface pressure at the interface between the upper atmosphere and supercritical layer is constant to 300 bar. This makes the coupling relatively simple, since I only need one grid per atmospheric parameter, such as the ones shown in Fig. 3.2. However, rocky planets ($R < 1.5 R_{\oplus}$) have densities that are not compatible with the presence of a massive atmosphere whose surface pressure is 300 bar. In these cases, no supercritical water layer exists. The pressure at the bottom of the atmosphere is less than 300 bar, and the interior and the atmosphere are coupled at the interface between the upper atmosphere and the mantle, which represents the surface of the planet. The atmospheric parameters, such as the emitted flux and the atmospheric thickness, are dependent on the atmospheric surface pressure, which is not constant. I describe how the dependence of the atmospheric parameters on the atmospheric surface pressure is implemented in the context of the interior-atmospheric coupling hereunder.

In the case of a constant surface pressure, I only needed one grid per atmospheric parameter, adding up to four grids in total: T_{surf} , A_B , z_{atm} and M_{atm} . Now these exact same parameters need more than one grid, each for a different surface pressure. I choose to compute the grids for surface pressures at 1 bar, 10 bar, 100 bar, in addition to the 300 bar grid that I had initially. This means that instead of four grids, I handle $4 \times 4 = 16$ grids when the surface pressure is less than 300 bar.

In Fig. 3.4 we show the OLR and absorbed flux of the same atmosphere for different surface pressures, where we can see that the surface temperatures at which the atmospheres are in equilibrium (dotted lines) are significantly different. The surface temperature increases as the surface pressure is higher because less massive atmospheres emit more IR radiation than heavier ones, which is due to decrease of the the opacity as the pressure is lower.

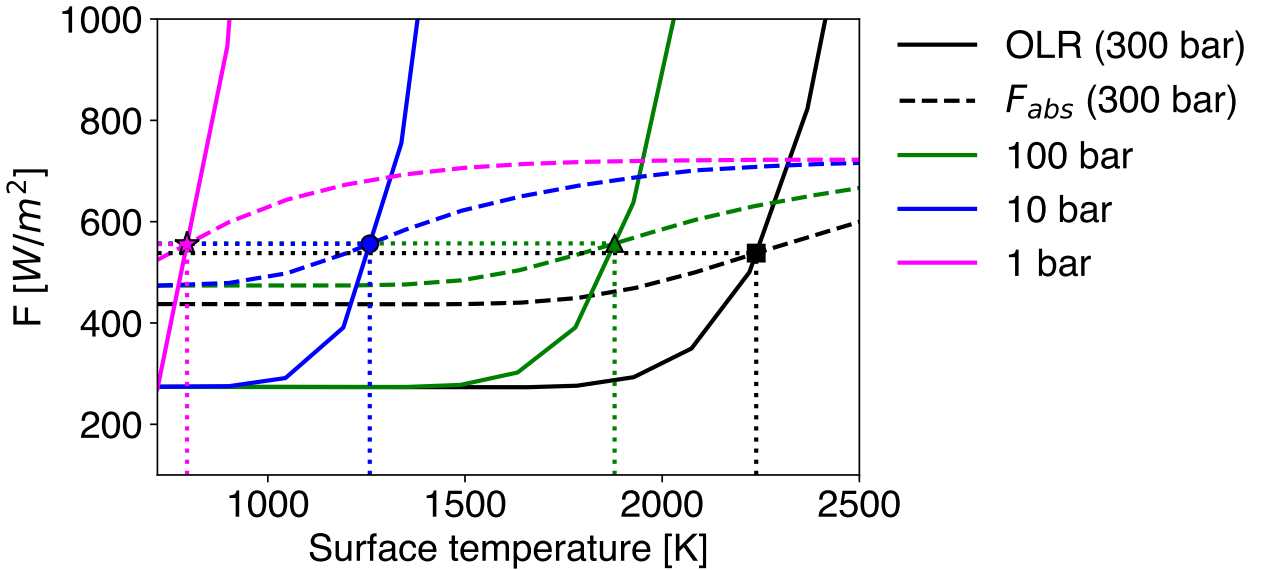


Figure 3.4.: OLR and absorbed flux as functions of the surface temperature for atmospheres with different surface pressures (300 bar, 100 bar, 10 bar and 1 bar). We assume a water-dominated composition, together with the irradiation conditions, bulk mass, and radius of TRAPPIST-1 c ($a_d = 0.0158$ AU, $M_{bulk} = 1.308 M_{\oplus}$ and $R_{bulk} = 1.308 R_{\oplus}$), as provided by Agol et al. (2021).

I parameterize the base temperature and the other atmospheric properties (Bond albedo, thickness and mass) as functions of the base pressure for different pairs of the bulk mass and radius. I

show the behaviour of these parameters with surface pressure in Fig. 3.5. The equilibrium surface temperature and the atmospheric thickness are fitted well by quadratic polynomials of the logarithm of the surface pressure, whereas the atmospheric mass is linear with respect to the pressure at the bottom of the atmosphere. This linear behaviour is expected, because the atmospheric mass can be estimated as:

$$M_{atm} = \frac{P_{base} 4\pi R_{bulk}^2}{g_{surf}} \quad (3.5)$$

where g_{surf} [m/s^2] = $9.8 \times (M_{bulk}/R_{bulk}^2)$ is the acceleration of gravity on the surface. The bulk mass and radius are expressed in Earth mass and radius units, respectively. For a given pair of the bulk mass and radius, Eq. 3.5 becomes a constant value times P_{base} , which makes the atmospheric mass linear with the pressure at the base of the atmosphere.

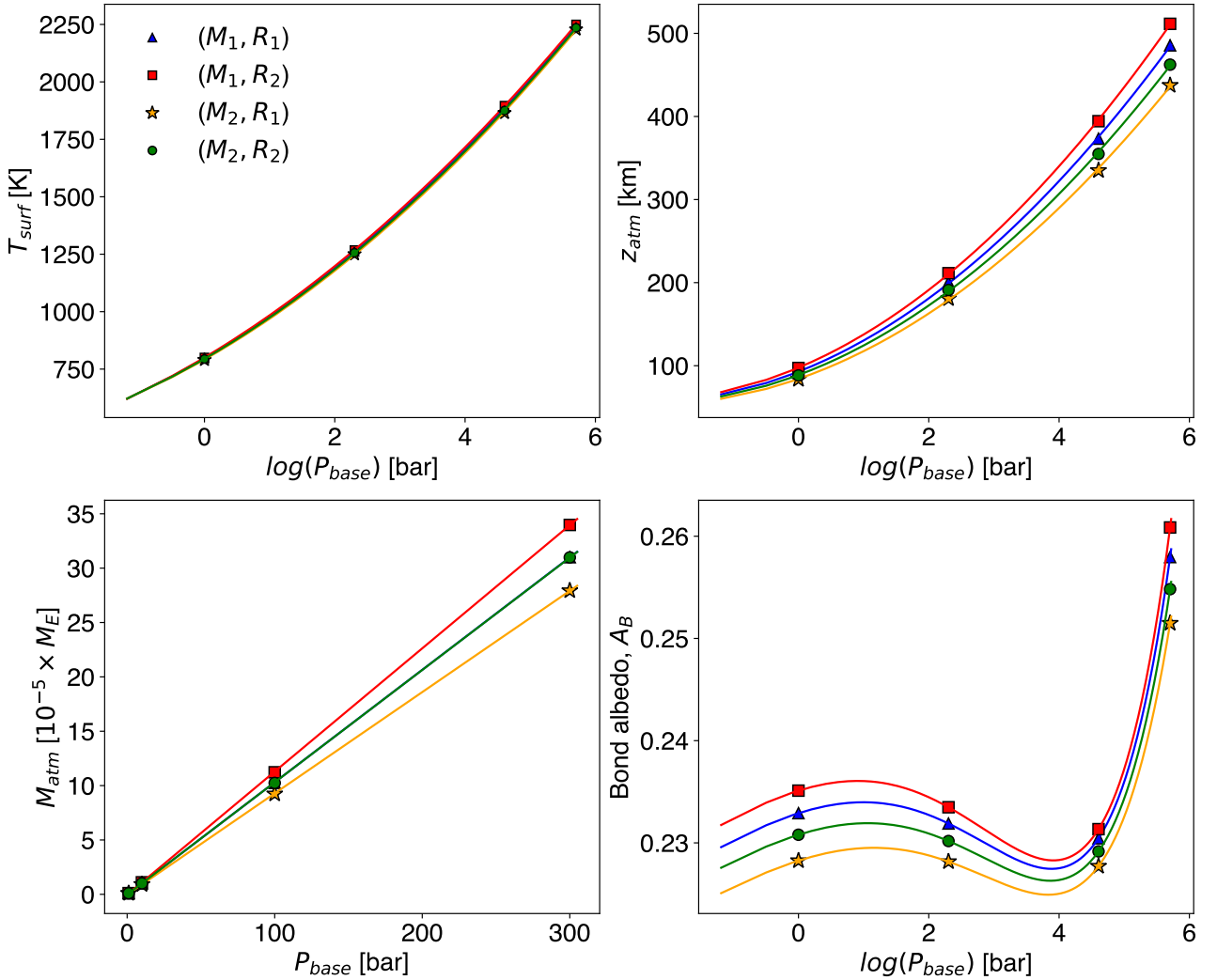


Figure 3.5.: Equilibrium surface temperature, atmospheric thickness, mass and Bond albedo as functions of surface pressure for 4 different pairs of mass and radius. The bulk mass and radius values are $M_1 = 1.252 M_\oplus$, $M_2 = 1.364 M_\oplus$, $R_1 = 1.085 R_\oplus$ and $R_2 = 1.109 R_\oplus$, which are the limits of the 1σ confidence intervals of the mass and radius of TRAPPIST-1 *c* (Agol et al. 2021).

The Bond albedo has a more complex relationship to the logarithm of P_{base} (see Fig. 3.5, bottom

right) than the other three atmospheric parameters, which can be fitted with a polynomial of 6th order. Except for the case of the atmospheric mass, the form of the function used for the fitting of the other three variables is difficult to determine from first principles, since the calculations of these parameters are based on a non-ideal water EOS and hydrostatic equilibrium for the thickness; and radiative transfer for the surface temperature and the Bond albedo. I store the coefficients of the fitting polynomials for the four variables as functions of different (M,R) pairs in our low-pressure atmosphere grids.

4. Atmospheric model

In Chapter 3, I explained how the interior model is coupled with an atmospheric model to calculate the total planetary radius and boundary conditions using the data grids generated by the atmospheric model. The interpolation of the atmospheric parameters from these grids requires loading a significant number of tables, and making assumptions on the behaviour of the parameters. In addition, this approach cannot generate the emission spectrum simultaneously, which is useful to assess the observability in emission spectroscopy. These disadvantages motivate me to implement our own atmospheric model (MSEIRADTRAN) to couple with the interior model, which I describe in this chapter. In Sect. 4.1, I lay out the numerical scheme of the atmospheric model. In Sect. 4.2, I describe how the pressure-temperature profiles are calculated by MSEIRADTRAN. In Sect. 4.3 to 4.5, I explain how the four contributions to the opacity are taken into account. These are collision-induced absorption (CIA), the line opacities, which are obtained with the k-correlated method, and scattering by clouds and gases (Rayleigh scattering). Finally, in Sect. 4.6 I compare the OLR, Bond albedo and total planetary radius computed with MSEIRADTRAN to reference atmospheric models.

4.1. Numerical scheme

Our atmospheric model presents a similar numerical structure to that of [Marcq et al. \(2017\)](#), which is publicly available¹. We first propose a pressure-temperature (PT) profile, that consists of a near-surface, dry convective layer; a wet convective region where condensation takes place, and an isothermal mesosphere on top. The calculations of the emission spectrum and the Bond albedo are performed by bands. [Marcq et al. \(2017\)](#) divide the spectrum from 0 to 10100 cm⁻¹ (equivalent to $\geq 1 \mu\text{m}$ in wavelength) in 36 bands to obtain the planetary emission. For each band:

1. Calculate the optical depth. The optical depth in each atmospheric layer is calculated as:

$$\tau = \kappa \rho \Delta z \quad (4.1)$$

where κ is the opacity in m²/kg, ρ is the mass density in kg/m³, and Δz is the thickness of the atmospheric layer in meters ([Liou 1980](#)). The total optical depth per layer has four contributions from different sources: atomic lines, Rayleigh scattering, collision-induced absorption (CIA) and scattering by clouds:

$$\tau_{total} = \tau_{lines} + \tau_{Rayleigh} + \tau_{CIA} + \tau_{clouds} \quad (4.2)$$

The contributions due to Rayleigh scattering and clouds are parameterized, while the optical depth caused by CIA is computed with tabulated opacities. Line opacities are obtained with the k-correlated method, which reproduces accurately line-by-line calculations. The line

¹<http://marcq.page.latmos.ipsl.fr/radconv1d.html>

contribution can also be modelled as a grey opacity. This will be computationally faster than the k-correlated method, but it yields a less accurate OLR. Once all the contributions to the optical depth are added, we obtain the total optical depth as a function of pressure or altitude. This is the input for the radiative transfer (RT) solver DISORT (Stamnes et al. 2017), which calculates the flux as a function of optical depth.

2. Obtain the upward flux at the top of the atmosphere (TOA). TOA corresponds to the topmost level in the optical depth ($\tau \ll 1$), which means the OLR corresponds to the flux obtained by DISORT at that level.
3. Store the upward flux at TOA for each band, $F_{TOA, i}^\uparrow$, where i is the band number. By doing this, we obtain the OLR as a function of wavelength, which is the emission spectrum.
4. Finally, the bolometric OLR, $F_{TOA, bol}^\uparrow$, is computed by integrating the emission spectrum over wavelength (Pluriel et al. 2019), which in our case is the total sum of the OLR:

$$OLR = F_{TOA, bol}^\uparrow = \int_0^\infty F_{TOA}^\uparrow(\nu) d\nu = \sum_{i=1}^{N_{bands}} F_{TOA, i}^\uparrow \quad (4.3)$$

Once the bolometric OLR is obtained, we initiate the calculation of the reflectivity in 30 bands, from 5 to 0.29 μm , to obtain the Bond albedo. This is done by calculating the optical depth from all four contributions similar to step 1, and then inputting the optical depth to DISORT to calculate the reflectivity. The bands for which we calculate both the flux at TOA and the reflectivity (from 1 to 5 μm) cannot have the two quantities calculated simultaneously since DISORT requires different input settings to calculate them. For the emission, we assume zero illumination from the top of the atmosphere, as well as an upward flux that forms 90 degrees with the surface of the planet, which corresponds to a polar angle equal to zero. To obtain the reflectivity, we assume an isotropically-incident source of radiation at the top of the atmosphere, while turning off all thermal emission sources. DISORT calculates the reflectivity of the atmosphere as a function of incident beam angle, which corresponds to the solar zenith angle (SZA). The SZA is the angle that the incident light forms with the normal of the incident surface. Once we obtain the dependence of the reflectivity with SZA, we can average it as (Simonelli and Veverka 1988):

$$A_B(\nu) = 2 \int_0^{\pi/2} A_B(\nu, i) \cos(i) \sin(i) di \quad (4.4)$$

where i is the SZA, and ν is wavenumber.

After averaging the reflectivity over SZA, we obtain the reflection spectrum, which is the dependence of the albedo as a function of wavenumber. To obtain the bolometric Bond albedo, we integrate Eq. 4.5:

$$A_{Bond, bol} = \frac{\pi \int_0^\infty A_B(\nu) B_\nu(T_\star) d\nu}{\sigma T_\star^4} \quad (4.5)$$

where $A_B(\nu)$ is the reflectivity as a function of wavenumber; $B_\nu(T_\star)$ is Planck's function for a temperature equal to the effective temperature of the host star T_\star , and σ is the Stefan-Boltzmann constant.

4.2. Pressure-temperature profile

The thermal structure of the atmosphere is divided in two layers. The near-surface layer is adiabatic, which means convection takes place, while the layer on top is a mesosphere with constant temperature. This is set to 200 K, which is representative of the cool temperatures that hot low-mass planets present in their mesospheres (Lupu et al. 2014; Leconte et al. 2013). In the adiabatic layer, condensation may take place depending on the value of the pressure compared to the water saturation pressure. Therefore, we assume an atmosphere with one condensable species, H₂O, and one non-condensable gas, CO₂. The adiabatic temperature profile is calculated as:

$$\left(\frac{\partial T}{\partial P}\right)_S \simeq \frac{\Delta T}{\Delta P} = \frac{T_{i-1} - T_i}{P_{i-1} - P_i} \quad (4.6)$$

where the index i specifies the location of each PT point in the atmosphere. The differential element of the pressure can be expressed as $\Delta P = P_i \Delta \ln(P)$ because $\frac{d \ln(P)}{dP} \sim \frac{\Delta \ln(P)}{\Delta P} = \frac{1}{P_i}$.

If we isolate T_i from Eq. 4.6, we obtain:

$$T_i = T_{i-1} - \left(\frac{\partial T}{\partial P}\right)_S P_i \Delta \ln(P_i) \quad (4.7)$$

The derivative of the temperature with respect to the pressure at constant entropy, $(dT/dP)_S$, is related to the adiabatic coefficient κ_{ad} by:

$$\kappa_{ad}(P, T) = \left(\frac{\partial \ln(T)}{\partial \ln(P)}\right)_S = \frac{P}{T} \left(\frac{\partial T}{\partial P}\right)_S \quad (4.8)$$

If the pressure of water vapour is below the gas saturation pressure, $P_v < P_{sat}(T)$, or if its temperature is above the temperature of the critical point of water, $T > T_{crit}$, we are under dry convection (Marcq 2012; Marcq et al. 2017). We calculate the derivative $(dT/dP)_S$ in the dry case as (Marcq et al. 2017):

$$\left(\frac{\partial T}{\partial P}\right)_{S, dry} = \frac{\rho_v T (\partial V_v / \partial T)_P}{\rho_v C_{p,v} + \rho_c C_{p,c}} \quad (4.9)$$

where ρ_v and ρ_c are the densities of water vapour and CO₂, respectively; and $C_{p,v}$ and $C_{p,c}$, their heat capacities. $V_v = 1/\rho_v$ is the specific volume of water vapour.

The atmospheric model of Marcq et al. (2017) uses the steam tables provided by Haar et al. (1984) to calculate the density and heat capacity of water vapour. These tables treat water as a non-ideal gas, although they are not valid for $T > 2500$ K. Therefore, for higher temperatures, I implement the use of the tables from Haldemann et al. (2020) to calculate the thermodynamic properties of water. These tables are a compilation of equations of state (EOS), where each EOS is applied in its validity region of the water phase diagram. There are two EOSs that are used in the region relevant for the atmospheres of low-mass, highly-irradiated planets. The first EOS is the IAPWS95 (Wagner and Pruß 2002), whose validity for the high-pressure supercritical regime was explored in Section 2.2. For lower pressures, such as liquid and cold gas phase, it is a reference EOS since the experimental data on which this EOS is based extends up to 1273 K. Haldemann et al. (2020) transition to an EOS valid at low pressures and high temperatures at 1200 K. This second EOS is the CEA (Chemical Equilibrium with Applications) package (Gordon 1994; McBride 1996). This package incorporates the effects of single ionisation and thermal dissociation, which are processes that occur only at high temperatures in gas phase. Figure 4.1 shows the dry adiabatic coefficient

as a function of pressure and temperature in the region of the water phase diagram relevant for hot planetary atmospheres. The reduction of the dry adiabatic coefficient at $T = 1000$ to 2500 K is due to thermal dissociation, whereas the decrease at higher temperatures ($T \geq 6000$ K) is caused by thermal ionisation (Haldemann et al. 2020).

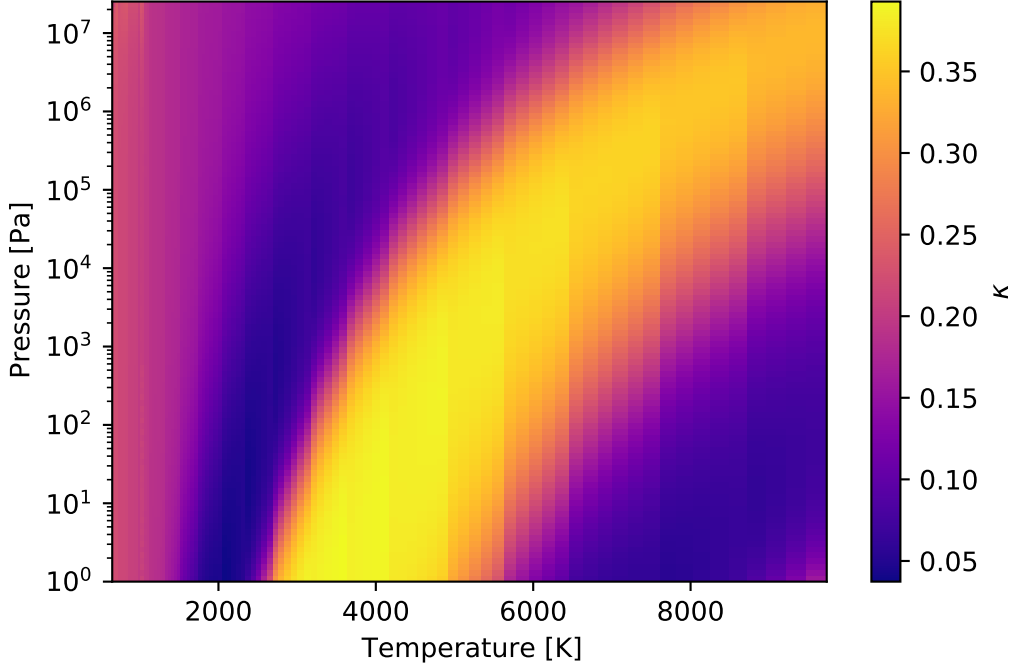


Figure 4.1.: Dry water adiabatic coefficient, κ , as a function of pressure and temperature. The region covers the cold and hot gas phase of water up to pressures close to the critical point.

When the atmospheric pressure reaches the water saturation curve, $P = P_{sat}(T)$, water vapour starts to condense out and clouds form. Since the phase change requires energy in the form of latent heat, the wet adiabatic coefficient is different from the dry adiabatic one (Eq. 4.9). The expression for the derivative $(dT/dP)_S$ in the wet case with CO_2 as the only non-condensable gas is (Marcq et al. 2017):

$$\left(\frac{\partial T}{\partial P}\right)_{S, wet} = \frac{1}{(dP_{sat}/dT) + \rho_c R/M_c (1 + \partial \ln(\rho_v)/\partial \ln(T) - \partial \ln(\alpha_v)/\partial \ln(T))} \quad (4.10)$$

where M_c is the molecular weight of carbon dioxide, and R is the ideal gas constant. α_v is the mixing ratio of water vapour density relative to CO_2 , $\alpha_v = \rho_v/\rho_c$. Its derivative $\partial \ln(\alpha_v)/\partial \ln(T)$ is defined as (Marcq et al. 2017; Marcq 2012; Kasting 1988):

$$\frac{\partial \ln(\alpha_v)}{\partial \ln(T)} = \frac{R/M_c (\partial \ln(\rho_v)/\partial \ln(T)) - C_{v,c} - \alpha_v (\partial s_v/\partial \ln(T))}{\alpha_v (s_v - s_c) + R/M_c} \quad (4.11)$$

where s_v and s_c are the specific entropies of water vapour and liquid water, respectively. $C_{v,c}$ is the specific heat capacity at constant volume of CO_2 . As we treat CO_2 as an ideal gas, we can calculate $C_{v,c} = C_{p,c} - R/M_c$. In this expression and in Eq. 4.9, the specific heat of CO_2 is obtained with Shomate's equation (Eq. 4.12), whose tabulated coefficients A to E are provided by Chase (1998). This is part of the NIST (National Institute of Standards and Technology) chemistry webbook

database², and it is valid for temperatures up to 6000 K.

$$C_{p,c} = A + B T + C T^2 + D T^3 + E/T^2 \quad (4.12)$$

The thermodynamic properties of water for the wet adiabat are computed with the tables provided by Haar et al. (1984), since in these conditions of low temperature they are valid.

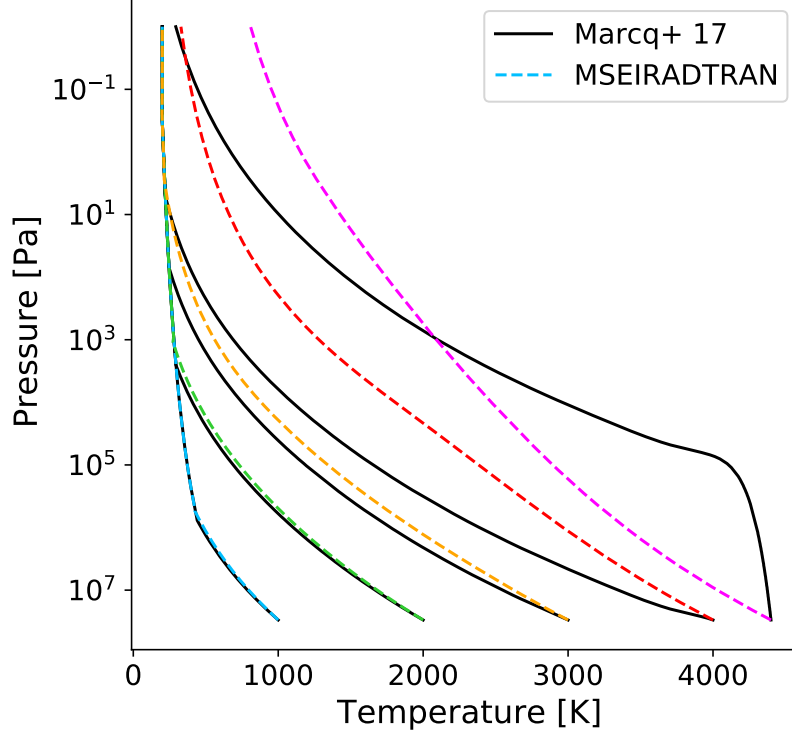


Figure 4.2.: PT profiles of a water-dominated atmosphere for surface temperatures $T_{surf} = 1000, 2000, 3000, 4000, 4400$ K. Black solid lines are profiles calculated with the EOS used in Marcq et al. (2017), while dashed lines indicates the PT profiles obtained with the MSEIRADTRAN.

Once the temperature profile is calculated as a function of pressure, the gravitational acceleration is computed with Gauss' law for gravity. The gravitational acceleration at the bottom of the atmosphere can be calculated as:

$$g_0 = g_{\oplus} \frac{M_T}{R_{bulk}^2} \quad (4.13)$$

where $g_{\oplus} = 9.8 \text{ m/s}^2$, M_T is the total mass of the planet in Earth mass units, and R_{bulk} is the bulk radius of the planet in Earth radius units, from the center of the planet to the planetary surface. The gravitational acceleration at an altitude z_i can then be expressed as:

$$g_i \simeq g_{\oplus} \frac{M_T}{(R_{bulk} + z_i)^2} = g_0 \frac{R_{bulk}^2}{(R_{bulk} + z_i)^2} \quad (4.14)$$

where the mass enclosed from the center of the planet to the radius $R_i = R_{bulk} + z_i$ is approximated as $M_i = M_T + M_{atm}(z_i) \simeq M_T$. This approximation is based on the typical value of the atmospheric

²<https://webbook.nist.gov/cgi/cbook.cgi?ID=C124389&Mask=1>

mass for surface pressures below 300 bar, $M_{atm} \leq 10^{-3} M_{\oplus}$, which is negligible compared to the total mass of the planet.

Finally, we can determine the altitude as a function of pressure under the assumption of hydrostatic equilibrium:

$$\frac{dP}{dz} \simeq \frac{\Delta P}{\Delta z} = -\rho g \quad (4.15)$$

If we isolate z_i from $\Delta z = z_{i-1} - z_i$ in Eq. 4.15, the altitude for a given point in the 1D grid is obtained as:

$$z_i = z_{i-1} + \frac{P_i \Delta \ln(P)}{g_i \rho_{total,i}} \quad (4.16)$$

where $\rho_{total,i}$ is the total mass density at altitude z_i , $\rho_{total,i} = \rho_{v,i} + \rho_{c,i}$.

Fig. 4.2 displays the PT profiles calculated with MSEIRADTRAN, and the model from Marcq et al. (2017), which uses the EOS and tables from Haar et al. (1984) for all temperatures, even those above 2500 K. It can be seen that the first two profiles, with surface temperatures of 1000 K and 2000 K, agree very well between the two models. Nonetheless, the profiles start to differ significantly as the surface temperatures increases, because the EOS used by Marcq et al. (2017) is out of its validity range.

4.3. Collision-induced absorption

In Eq. 4.2, one of the contributions to the total optical depth is collision-induced absorption (CIA). CIA, or continuum absorption features, are produced by inelastic collisions of the molecules in a gas. These collisions might induce quantum transitions that absorb and emit energy, which contributes to the line transition opacity. CIA absorption is particularly important in dense gases, such as steam and CO₂ at high pressures, specially if the line opacity is weak (Pluriel et al. 2019). CIA opacities are usually provided in tabulated data that can be interpolated in wavelength and temperature. I use CIA absorption data for H₂O-CO₂ and H₂O-H₂O provided by Ma and Tipping (1992) and Tran et al. (2018)³, respectively. We can see in Fig. 4.3 that the highest contribution to the opacity comes from H₂O-H₂O CIA in a water-dominated atmosphere, followed by H₂O-CO₂ and CO₂-CO₂, which is in agreement with Turbet et al. (2017) (see their Figure 1).

CO₂-CO₂ CIA opacities are read from a look-up table obtained by Bézard et al. (2011) and Marcq et al. (2008), which is also used in the atmospheric model by Marcq et al. (2017). Our H₂O-H₂O CIA table covers the complete spectral range where we calculate both our emission and reflection spectra, while the H₂O-CO₂ CIA table covers the bands with wavelength $\geq 1 \mu\text{m}$, which corresponds to the bands necessary for the emission spectrum only. For the bands whose wavelength is outside the spectral range of the CIA table, we assume a constant CIA opacity value equal to the opacity at the limit band of the table. This means that we assume the CIA opacity value at $1 \mu\text{m}$ for the visible bands when we calculate the reflection spectrum.

At each point of the 1D PT grid, I calculate the contribution to the optical depth by CIA, which is τ_{CIA} in Eq. 4.2. This optical depth is constituted of three contributions, which correspond to the three CIA pairs that are present in an atmosphere composed of H₂O and CO₂:

$$\tau_{CIA} = \tau_{WW} + \tau_{CC} + \tau_{CW} \quad (4.17)$$

³https://www.lmd.jussieu.fr/~lmdz/planets/LMDZ.GENERIC/datagcm/continuum_data/

where τ_{WW} is CIA absorption due to H₂O-H₂O collisions, τ_{CC} corresponds to CO₂-CO₂ interactions, and τ_{CW} to H₂O-CO₂.

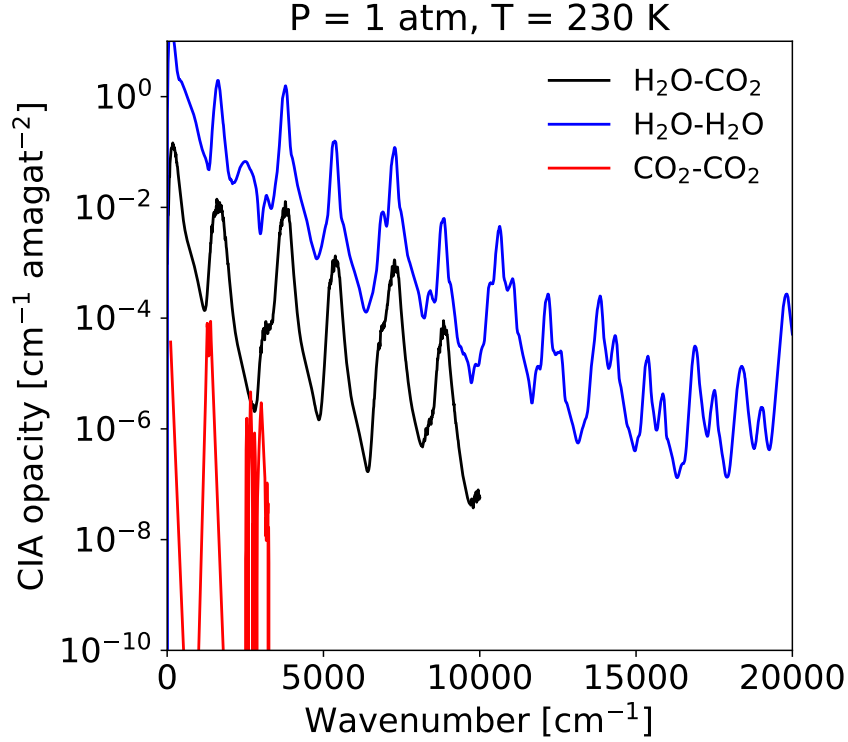


Figure 4.3.: CIA opacities for a water-dominated atmosphere with 1% CO₂, for T = 230 K and P = 1 atm.

The opacities due to each type of CIA are read from their respective tables. Consecutively, the optical depth of each contribution is calculated with Eq. 4.1, or its alternative form depending on the original units of the CIA data tables. For instance, to calculate τ_{CC} we use this alternative form of Eq. 4.1:

$$\tau_{CC} = n_c C_c \omega \quad (4.18)$$

where n_c is the number density of CO₂ in the middle of the atmospheric layer, ω is the column abundance of the gas, and C_c is the opacity provided by the opacity table. For CO₂, we convert the mass density in kg/m³, ρ_c , into number density in molecules/cm³, n_c :

$$n_c = \rho_c 10^{-3} \frac{1}{44 \text{ g/mol}} 6.022 \times 10^{23} = 1.3686 \times 10^{19} \rho_c \quad (4.19)$$

The column abundance of the gas, ω , represents the thickness of an equivalent atmospheric column of standard pressure and temperature. It is calculated as (Sanchez-Lavega 2011):

$$\omega = \frac{1}{L_0} n_c \Delta z \quad (4.20)$$

where $L_0 = 2.687 \times 10^{19} \text{ cm}^{-3}$ is Loschmidt's constant, n_c is the number density calculated in Eq. 4.19. The thickness of the atmospheric layer, Δz , must be in cm. The final units of ω are cm-amagat.

Finally, C_c is provided by the opacity table in $\text{cm}^{-1} \text{ amagat}^{-2}$ units⁴, but it needs to be expressed in $\text{cm}^2/\text{molecule}$ to be consistent with the units of ω and n in Eq. 4.18. For that, we convert C_c from $\text{cm}^{-1} \text{ amagat}^{-2}$ to $\text{cm}^5/\text{molecule}^2$ (Richard et al. 2012):

$$C_c [\text{cm}^5/\text{molecule}^2] = 1.385 \times 10^{-39} C_c [\text{cm}^{-1}/\text{amagat}^2] \quad (4.21)$$

Then we must convert the opacity from $\text{cm}^5/\text{molecule}^2$ to $\text{cm}^2/\text{molecule}$:

$$C_c [\text{cm}^2/\text{molecule}] = 2.687 \times 10^{19} C_c [\text{cm}^5/\text{molecule}^2] \quad (4.22)$$

where the constant factor comes from the definition of amagat in $\text{molecules}/\text{cm}^3$ units (Richard et al. 2012).

For τ_{WW} and τ_{CW} , the tables provide the opacity in $\text{cm}^{-2} \text{ molecule}^{-1} \text{ atm}^{-1}$. To obtain the product $\kappa \times \rho$ in Eq. 4.1, we convert the opacity, C_{ij} :

$$(\kappa \times \rho)_{ij} [\text{cm}^{-1}] = \frac{C_{ij}}{1.013 \times 10^6} \frac{P_i P_j}{T k_B} \quad (4.23)$$

where ij is the pair $\text{H}_2\text{O}-\text{H}_2\text{O}$ or $\text{H}_2\text{O}-\text{CO}_2$. The pressures, temperature and the Boltzmann constant, $k_B = 1.38 \times 10^{-16} \text{ cm}^2 \text{ g}/\text{s}^2 \text{ K}$, are in cgs units. The factor 1.013×10^6 is the equivalence of 1 atm in cgs units.

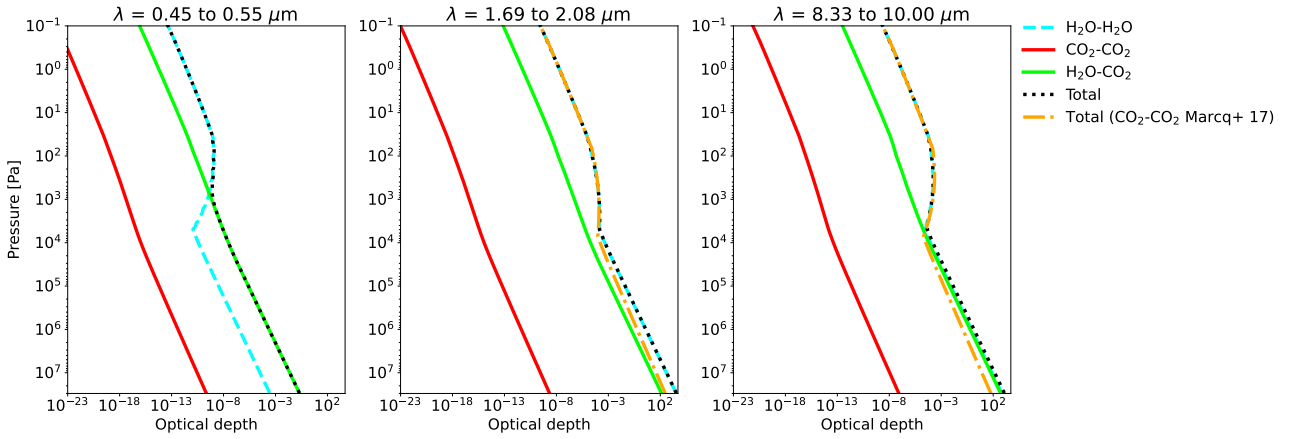


Figure 4.4.: Optical depth of the three collision-induced absorption (CIA) contributions for three spectral windows; one in the optical (left panel), one in the near IR (middle panel), and one in the far IR (right panel). The total optical depth for the same atmosphere obtained with the model by Marcq et al. (2017) is also shown for comparison in the IR windows.

The resulting optical depth for one spectral band in the optical and two in the IR are shown in Figure 4.4, where we can see that the total CIA optical depth agrees well between our model and Marcq et al. (2017) for the IR bins. We also obtain the expected result that for a given point in the PT grid, $\tau_{CC} < \tau_{CW} < \tau_{WW}$ (see Fig. 4.3), with the exception of the optical window at high pressures, where $\tau_{CW} > \tau_{WW}$. This is because the value of the $\text{H}_2\text{O}-\text{CO}_2$ opacity in the visible is set to the limit of the opacity table, which is at $\lambda = 1^4 \mu\text{m}$ ($\nu = 10^4 \text{ cm}^{-1}$ in Fig. 4.3). This has a negligible effect in our calculation of the Bond albedo.

⁴The amagat is a unit of number density, and it is defined as the number of molecules per unit volume of an ideal gas at a pressure of 1 atm, and a temperature of 0°C .

4.4. K-correlated method

4.4.1. Formalism

In this section I explain how the term τ_{lines} in Eq. 4.1 is obtained. To determine how opaque or transparent is the atmosphere within a spectral window, we need to integrate the line opacity over wavelength. In Figure 4.5 we can see that the opacity is a non-monotonic function that varies very rapidly with wavelength, making its integration computationally expensive. One way to reduce this computational effort is to integrate the opacity over its cumulative probability distribution, which is a monotonic and smooth function (Fig. 4.5, middle panel). The cumulative distribution function of the opacity, $G(X = \kappa)$, expresses the probability of finding an opacity value that is less or equal to κ . This means that it is zero if κ is the minimum, $G(X = \kappa_{min}) = 0$, and 1 if κ is the maximum, $G(X = \kappa_{max}) = 1$. The inverse of the cumulative distribution function of the opacity is $\kappa(G)$, which is known as the k-distribution function (Fig. 4.5, left panel). This function can be discretized into the k-coefficients, κ_i . These are defined as the values of the continuous k-distribution function evaluated at specific probabilities, known as G-points (G_i), so $\kappa_i = \kappa(G = G_i)$. In addition to G , the k-coefficients are also dependent on spectral band, pressure and temperature, which is explicitly expressed with the notation $\kappa_i^b(P, T)$.

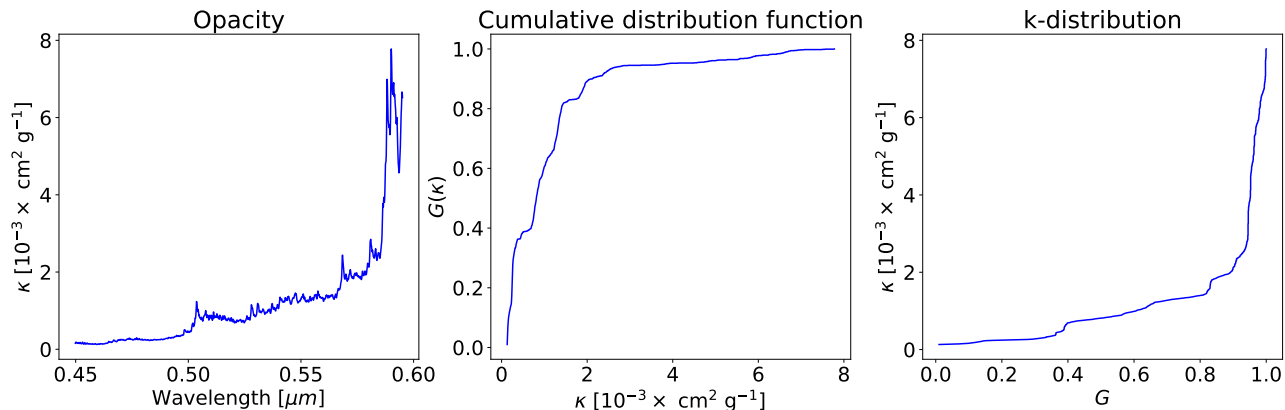


Figure 4.5.: Right panel: opacity as a function of wavelength for water vapor at 2550 K of temperature and 215 bar of pressure. The spectral band shown here ranges from 0.45 to 0.60 μm . Middle panel: Cumulative distribution function of the opacity, $G(\kappa)$, for the same spectral band, pressure and temperature in the right panel. Left panel: Inverse of the cumulative distribution function displayed in the middle panel.

If we consider that the optical depth - or spectral transmittance, related to the optical depth with $\tau = -\ln(T_\nu)$ - is independent of the order of the opacities, κ , within a spectral band, we can replace wavenumber integration by integration in κ -space. Moreover, given the properties of the k-distribution function of κ , we can derive an analytic expression for the cumulative probability function of the opacity, $G(\kappa)$ (Liou 1980). This function is monotonically increasing and smooth in κ -space, so we can calculate the integral necessary for the spectral transmittance with a Laplace-Gaussian quadrature. Hence, the integration in G -space substitutes the integration in wavenumber (ν -space), since we assume that the atmospheric layers are homogeneous. This is obtained as a finite number of exponential sums:

$$T^b(m) = \int_0^1 e^{-\kappa^v(G) m} dG \approx \sum_{i=1}^{N_G} e^{-\kappa_i^b m(P,T)} w_i \quad (4.24)$$

where κ_i are the values of the opacity at the G-points G_i , m is the column density of the atmospheric level, and w_i are the associated Legendre-Gaussian weights. The Legendre-Gaussian weights are defined as:

$$w_i = \frac{2}{(1 - y_i^2) P'_{N_G}(y_i)^2} \quad (4.25)$$

where y_i is the i th root of the Legendre polynomial of order N_G , and P'_{N_G} is the derivative of this polynomial (Malik et al. 2017). The G-points are fixed by the roots of the Legendre polynomial:

$$G_i = \frac{(1 + y_i)}{2} \quad (4.26)$$

The transmittance is different for each spectral band. We can calculate the line optical depth in each band as $\tau^b = -\ln(T^b)$ (Malik et al. 2017; Leconte 2021; Sanchez-Lavega 2011). Then, we can input this optical depth to the RT solver, and obtain the upward flux at TOA in each spectral bin, $F_{TOA, b}^\dagger$. I will refer to this as the uncorrelated-k approach or method henceforth.

In each atmospheric layer, the pressure and temperature are considered constant. Under this condition and within each spectral band, we can exchange wavenumber with G (Mollière 2017). Then, we can solve the RT equations, and obtain the flux at TOA. Consequently, we integrate the upward flux over G to obtain the OLR within each band:

$$F_{TOA, b}^\dagger = \int_0^1 F^b(G) dG = \sum_{i=1}^{N_G} F_{G_i}^b \Delta G_i \quad (4.27)$$

where $F_{G_i}^b$ is the flux at TOA calculated by the RT solver, which has as input the optical depth $\tau_{G_i}^b = \kappa_{G_i}^b m(P, T)$. ΔG_i are the widths of the bins in G -space. This is the k-correlated method that Marcq et al. (2017) and Malik et al. (2017) use.

In both methods, the column density is calculated in each atmospheric layer as:

$$m = (n_v + n_c) \Delta z \quad (4.28)$$

where n_v and n_c are the number densities of water and CO_2 , respectively. These are computed from the mass density as $n_{v,c} = \rho_{v,c} N_A / MW_{v,c}$, where N_A is Avogadro's number, and MW the molecular weight. Δz was defined in Sect. 4.2 as the increase in altitude within an atmospheric layer.

Consequently, in the atmospheric model we iterate over two variables:

- over G -points, to obtain the F_{G_i} values within each band, and then sum all the contributions to yield the TOA flux $F_{TOA, b}^\dagger$.
- over bands, to add the upward TOA fluxes of all bands to obtain the bolometric OLR.

The difference between the uncorrelated-k and the correlated-k methods is where the iteration over G -points is located in the algorithm. In the k-correlated method, the spectral lines are assumed to be correlated from an atmospheric layer to the next one, making it necessary to propagate the fluxes through the atmosphere for all g-points before performing the integral. Figure 4.6 shows a schematic of these sums over G -points and bands to obtain the total integrated OLR for ζ -correlated

method. As it can be seen in Fig. 4.6, the sum over G -points ends after the RT solver outputs $F_{G_i}^{bj}$. However, in the uncorrelated- k approach, the iteration over G -points ends before inputting the optical depth to the RT solver. The final line optical depth that is the input for the RT solver in the k -uncorrelated method assumes that all the contributions from the G -points are included before the sum in Eq. 4.24. I consider the uncorrelated- k method to compute the boundary conditions for the interior model since it is faster than the k -correlated one. This is because the RT solver is the most time-consuming subroutine in the atmospheric model. In the former case, it takes 0.22 seconds per spectral band to obtain the OLR, while in the latter, the time per band is 0.40 seconds, which is almost as twice as the k -uncorrelated approach. I therefore compare the k -uncorrelated method to several k -correlated models for its implementation in MSEIRADTRAN.

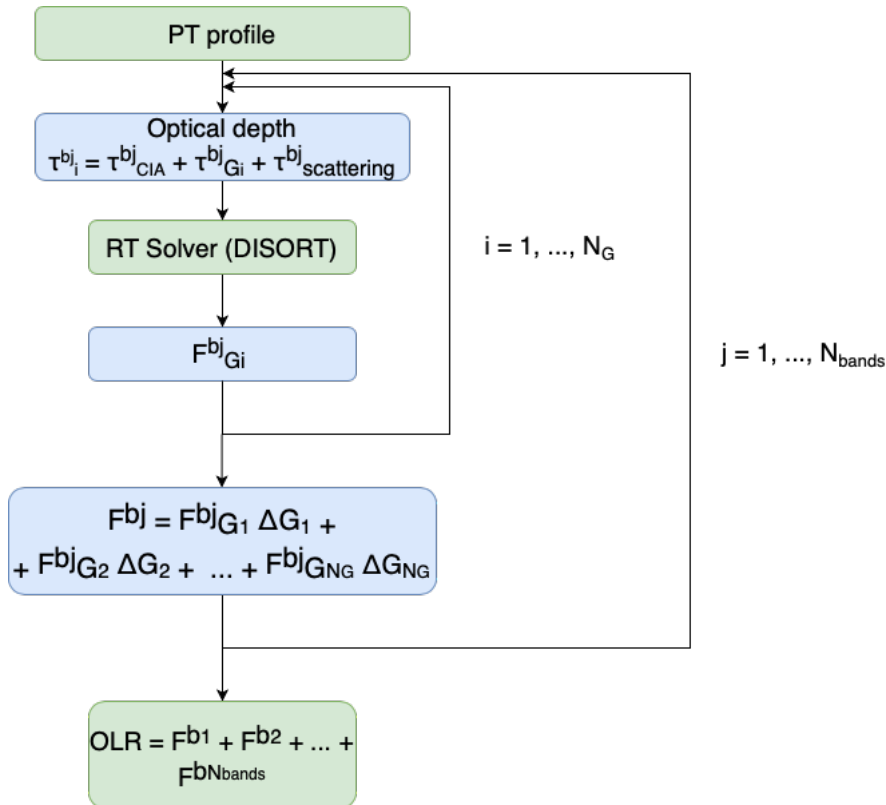


Figure 4.6.: Algorithm in the atmospheric model to calculate the OLR with the k -correlated method (see text). $\tau_{scattering}^{bj}$ includes the contributions from Rayleigh scattering and clouds.

4.4.2. K-table processing

The k -coefficients are stored in k -tables, which are calculated from an opacity table with the wavelength-dependent line opacities. Grimm and Heng (2015) provide a database⁵ of pre-calculated HELIOS-K output files for different species and line lists. For water and CO_2 -dominated atmospheres, I need opacities for these two absorber species. The line lists I use are POKAZATEL (Polyansky et al. 2018) for H_2O and HITEMP2010 for CO_2 (Rothman et al. 2010). POKAZATEL is the water line list with the widest validity range in temperature for water in planetary atmospheres, with a maximum temperature of 5000 K, while HITEMP maximum temperature is 4000 K. Their

⁵<https://chaldene.unibe.ch/data/Opacity3/>

spectral coverage includes the visible and the IR. To obtain the k-tables, the spectral resolution of the original opacity table is $R = 5000$, which is the minimum recommended (Malik et al. 2017, 2019).

Figure 4.7 shows the opacities calculated with these line lists for a single (P, T) point, as a function of wavelength. The total opacity is dominated by the contribution of water, because it constitutes 99% of the atmosphere. The contribution of CO_2 is only noticeable at approximately $4.5 \mu\text{m}$.

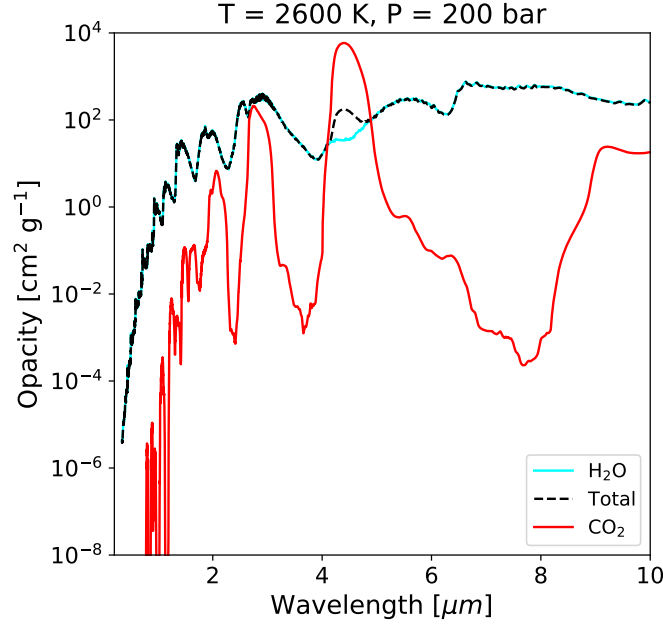


Figure 4.7.: Opacity as function of wavelength for H_2O and CO_2 . $T = 2600 \text{ K}$ and $P = 200 \text{ bar}$, which are pressure and temperature conditions easily reached in the warm atmospheres of low-mass planets. We also show the total opacity of a water-dominated gas with 1% CO_2 , which is obtained by weighting the opacities with their respective mass mixing ratios.

I implement two different versions of MSEIRADTRAN: one with a spectral resolution of $R = \lambda/\Delta\lambda = 10$, and another one with the native spectral resolution of our k-table, which is $R \approx 200$ in the IR. The former is used to obtain the atmospheric parameters necessary to couple with the interior model within the MCMC Bayesian framework, since this resolution is sufficient to produce accurate OLR and Bond albedos fast. The version with higher resolution can be used to generate emission spectra to assess the observability of planets with telescopes, such as the James Webb Space Telescope (JWST).

In the low-resolution version of MSEIRADTRAN, the spectral bands are wider than the wavelength bins of the original HELIOS-K tables. To obtain a k-table with a similar binning as Marcq et al. (2017) and Pluriel et al. (2019) (see Tables 4.1 and 4.2), I follow the procedure in Leconte (2021) to combine two spectral bins:

1. Given a (P, T) point in our table, I show the opacity as a function of G , $\kappa(G)$ (see Figure 4.8, left panel) in each bin. I exchange G in the y-axis and κ in the x-axis. Now I define the x-axis, κ_j , as a logarithmically spaced array that ranges from the minimum value of κ in both bins, to their maximum value.
2. Then I interpolate the G -points at the points of the array κ_j , to obtain $G(\kappa_j) = G_j$ for each of the bins. Note that if κ_j is greater than the κ_{max} of one of the bands, $G = 1$, since $G(\kappa)$ is the

cumulative distribution of the opacity within each bin. Similarly, if κ_j is less than the κ_{min} of one of the bands, $G = 0$.

3. I calculate the G -points of the combined band as:

$$G_{j, combined} = \frac{\Delta\nu_{band 1} G_{j,band 1}}{\Delta\nu_{combined}} + \frac{\Delta\nu_{band 2} G_{j,band 2}}{\Delta\nu_{combined}} \quad (4.29)$$

$G_{j, combined}$ as a function of κ_j can be seen in Figure 4.8 (left panel, black circles). Eq. 4.29 is the particular case of equation 7 in Leconte (2021) for two bins. This procedure can be extended with as many bins as necessary (N , in Eq. 4.30) to fill the spectral range of the superbin of the atmospheric model:

$$G_{j, combined} = \sum_{i=1}^N \frac{\Delta\nu_{band i} G_{j,band i}}{\Delta\nu_{combined}} \quad (4.30)$$

4. Finally, I interpolate κ_j as a function of G , and sample the opacity at the G -points I wish to use in the atmospheric model.

Steps 1 to 4 are performed for each (P, T) point in the opacity table.

The red triangles in Fig. 4.8 indicate the G -points at which the opacity is evaluated to obtain the binning in G -space of the approach of Marcq et al. (2017). To obtain the k-table for the low-resolution, k-uncorrelated model, I evaluate the opacity at the G -points that correspond to the roots of the 16th order Legendre polynomial (see Eq. 4.26). Table 4.3 shows the values of the G -points in both approaches.

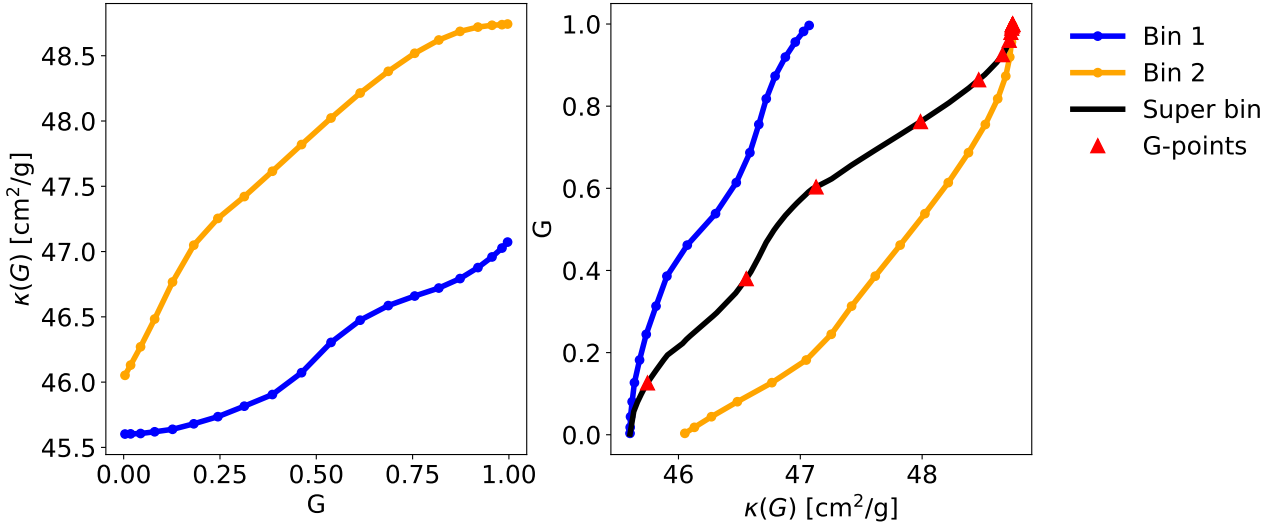


Figure 4.8.: Right: Opacity as a function of G for the two bins that need to be combined. Left: G as a function of the opacity for the combined bin (black). Red triangles mark the $\{G, \kappa(G)\}$ points where the combined opacities are sampled to have the same G -points as Marcq et al. (2017). In this example, bin 1 has a spectral coverage from 1.887 to 1.890 μm , while bin 2 covers 1.890 to 1.894 μm . Consequently, the spectral coverage of the super bin is 1.887 to 1.894 μm .

# Band	ν_{ini} [cm ⁻¹]	ν_{fin} [cm ⁻¹]	λ_{ini} [μm]	λ_{fin} [μm]
1	0	40	∞	250
2	40	160	250	62.5
3	160	280	62.5	35.7
4	280	380	35.7	26.3
5	380	500	26.3	20
6	500	582	20	17.2
7	582	600	17.2	16.7
8	600	720	16.7	13.9
9	720	752	13.9	13.3
10	752	800	13.3	12.5
11	800	900	12.5	11.1
12	900	1000	11.1	10.0
13	1000	1200	10.0	8.3
14	1200	1350	8.3	7.4
15	1350	1450	7.4	6.9
16	1450	1550	6.9	6.5
17	1550	1650	6.5	6.1
18	1650	1750	6.1	5.7
19	1750	1850	5.7	5.4
20	1850	1950	5.4	5.1
21	1950	2050	5.1	4.9
22	2050	2200	4.9	4.5
23	2200	2500	4.5	4.0
24	2500	2800	4.0	3.6
25	2800	3200	3.6	3.1
26	3200	3600	3.1	2.8
27	3600	4000	2.8	2.5
28	4000	4400	2.5	2.3
29	4400	4800	2.3	2.1
30	4800	5900	2.1	1.69
31	5900	6000	1.69	1.67
32	6000	6500	1.67	1.54
33	6500	8000	1.54	1.25
34	8000	8300	1.25	1.20
35	8300	9300	1.20	1.07
36	9300	10100	1.07	0.99

Table 4.1.: Spectral bands in the IR for MSEIRADTRAN (low-resolution version). ν_{ini} and ν_{fin} are the initial and final wavenumber limits of the bin, respectively. The other two columns are the equivalent wavelengths of these two limits.

# Band	ν_{ini} [cm^{-1}]	ν_{fin} [cm^{-1}]	λ_{ini} [μm]	λ_{fin} [μm]
21	1950	2050	5.1	4.9
22	2050	2200	4.9	4.5
23	2200	2500	4.5	4.0
24	2500	2800	4.0	3.6
25	2800	3200	3.6	3.1
26	3200	3600	3.1	2.8
27	3600	4000	2.8	2.5
28	4000	4400	2.5	2.3
29	4400	4800	2.3	2.1
30	4800	5900	2.1	1.69
31	5900	6000	1.69	1.67
32	6000	6500	1.67	1.54
33	6500	8000	1.54	1.25
34	8000	8300	1.25	1.20
35	8300	9300	1.20	1.07
36	9300	10100	1.07	0.99
37	10100	10417	0.99	0.96
38	10417	11236	0.96	0.89
39	11236	11905	0.89	0.84
40	11905	12821	0.84	0.78
41	12821	13333	0.78	0.75
42	13333	14493	0.75	0.69
43	14493	14925	0.69	0.67
44	14925	15384	0.67	0.65
45	15384	16667	0.65	0.60
46	16667	18182	0.60	0.55
47	18182	22222	0.55	0.45
48	22222	25000	0.45	0.40
49	25000	28571	0.40	0.35
50	28571	33333	0.35	0.30
51	33333	34482	0.30	0.29

Table 4.2.: Spectral bands in the nIR and the optical for MSEIRADTRAN for the calculation of the Bond albedo. ν_{ini} and ν_{fin} are the initial and final wavenumber limits of the bin, respectively. The other two columns are the equivalent wavelengths of these two limits.

Marcq et al. (2017)	MSEIRADTRAN
0.12618	0.00530
0.37996	0.02771
0.60333	0.06718
0.76231	0.12230
0.86421	0.19106
0.92533	0.27099
0.96029	0.35920
0.97952	0.45249
0.98976	0.54751
0.99504	0.64080
0.99767	0.72901
0.99897	0.80894
0.99957	0.87770
0.99983	0.93282
0.99995	0.97229
0.99999	0.99470

Table 4.3.: G-points used in Marcq et al. (2017) and MSEIRADTRAN.

4.4.3. Gas mixtures

There are three methods to compute the k-tables of a mixture of several gases: pre-mixing, correlated, and uncorrelated mixing. Pre-mixing, or pre-mixed k-coefficients (Goody et al. 1989; Amundsen et al. 2017), consists on calculating directly the k-coefficients from the line-by-line opacities. Then the k-coefficients can be interpolated in composition between the different k-tables (Marcq 2012; Marcq et al. 2017). If line-by-line calculations are not available, and we have the k-tables of the individual species that need to be mixed, the k-coefficients of these individual k-tables can be used to calculate the k-coefficients of the final mixture. This can be done by assuming that the spectral features of the individual gases are correlated (Malik et al. 2017), in which case the mixing of the individual species is:

$$\kappa_{mix,i} = \sum_{j=1}^{N_{gases}} \chi_j \kappa_{j,i} \quad (4.31)$$

where χ_j is the mixing ratio by mass of the j th gas, $\kappa_{j,i}$ is the k-coefficient of the j th gas evaluated at the G_i point, and $\kappa_{mix,i}$ is the k-coefficient of the final mixture of gases that corresponds to G_i . The mixing ratio by mass is defined as:

$$\chi_j = \frac{X_j MW_j}{\mu} \quad (4.32)$$

where $X_j = \frac{P_j}{P}$ is the volume mixing ratio of the j th species, MW_j is its molecular weight and μ is the mean molecular weight of the mixture.

Nonetheless, the wavelengths at which the spectral lines of a gas are placed are mostly dependent on the molecular structure of this gas, making it uncorrelated to the spectral lines of another species with which it is being mixed. Assuming that the spectral lines of the species are uncorrelated is

equivalent to say that the lines overlap randomly, which is known as Goody's random overlap approximation. In reality, line-by-line k-coefficients have intermediate values between the correlated and the uncorrelated approximations (Pierrehumbert 2010). Therefore, I implement both approximations in the atmospheric model and compare their final OLR for a mixture of water and CO₂ (see Sect. 4.6.1).

Mollière (2017) and Amundsen et al. (2017) describe the procedure to calculate the k-coefficients of the gas mixture with the uncorrelated method. Firstly, we calculate a matrix that contains the k-coefficients of the mixture:

$$\kappa_{mix,ij} = \chi_{H_2O} \kappa_{H_2O,i} + \chi_{CO_2,j} \kappa_{CO_2,j} \quad (4.33)$$

where $\kappa_{H_2O,i}$ is the k-coefficient of water evaluated at $G = G_i$, and $\kappa_{CO_2,j}$ is the k-coefficient of carbon dioxide evaluated at $G = G_j$. Since we use the 16 G -points listed in Table 4.3, the dimension of the matrix is $16 \times 16 = 256$ elements in total. Similarly, we calculate a second matrix that contains the weights that correspond to the k-coefficients in the first matrix:

$$\Delta G_{ij} = \Delta G_i \Delta G_j \quad (4.34)$$

T = 2600 K, P = 200 bar, $\lambda = 2 \mu\text{m}$

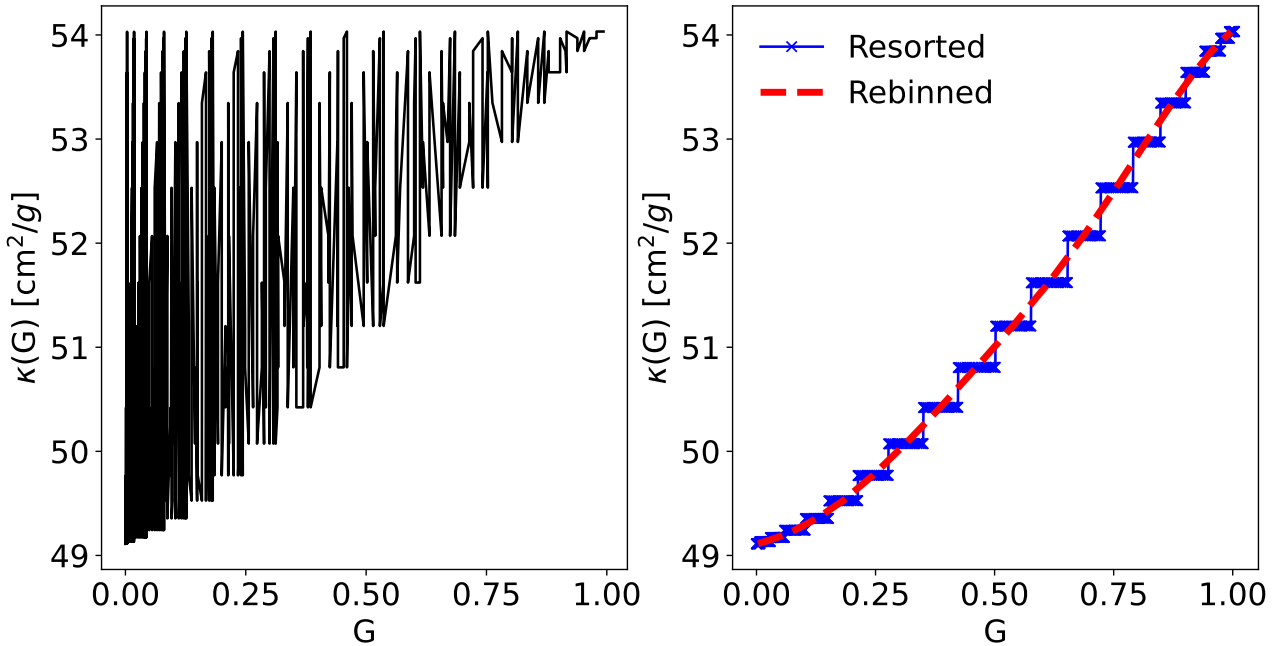


Figure 4.9.: Right: K-coefficients as a function of G for the unsorted and unbinned random overlap method. Left: K-coefficients as a function of G after sorting (blue), and both sorting and rebinning (red). Both panels correspond to a 99% water and 1% CO₂ gas mixture, a pressure of $P = 200$ bar, a temperature of $T = 2600$ K and a wavelength of $\lambda = 2 \mu\text{m}$.

Then I re-compute the G values that correspond to the weights in the matrix ΔG_{ij} by sorting them and using:

$$G_k = \sum_{l=1}^{k-1} \Delta G_l + \frac{\Delta G_k}{2} \quad (4.35)$$

$\kappa(G)$ presents a scaling problem, because it is not monotonic (see Fig. 4.9, right panel). [Lacis and Oinas \(1991\)](#) and [Amundsen et al. \(2017\)](#) suggest resorting and rebinning the k-coefficients to solve this. Instead of resorting the weights, we resort the k-coefficients in increasing order and plot it with their respective weights. This yields the blue line in the right panel of Figure 4.9. Finally, to rebin the k-coefficients to the G points I use in the atmospheric model (Table 4.3), I interpolate $\kappa(G)$ along G . These final k-coefficients are shown in Fig. 4.9 in red. The resorted κ function shown in Fig. 4.9 is similar to that displayed in figure 1 of [Amundsen et al. \(2017\)](#).

4.5. Cloud and Rayleigh scattering

4.5.1. Opacities

In Eq. 4.2, the optical depth due to Rayleigh scattering contributes to the total optical depth. In the clear atmospheric layers, Rayleigh is the only source of scattering, with $\kappa_{clouds} = 0$. The Rayleigh scattering opacity can be parameterized as presented in [Pluriel et al. \(2019\)](#):

$$\kappa_{Rayleigh}(\lambda) = \kappa_0 \left(\frac{\lambda_0}{\lambda} \right)^4 \quad (4.36)$$

where κ_0 and λ_0 are tabulated values obtained by [Kopparapu et al. \(2013\)](#) for H₂O, and [Sneep and Ubachs \(2005\)](#) for CO₂. In this case, the opacity does not depend on the pressure or the temperature, but it depends on the wavelength. These values are tabulated in Table 4.4.

	H ₂ O	CO ₂
σ_0 [cm ² /molecule]	2.5×10^{-27}	1.24×10^{-26}
λ_0 [μm]	0.6	0.532

Table 4.4.: Rayleigh scattering opacity, σ_0 , and reference wavelength, λ_0 . The references for these values are [Kopparapu et al. \(2013\)](#) and [Sneep and Ubachs \(2005\)](#) for water and CO₂, respectively.

In addition, in the atmospheric layers where the temperature reaches the condensation temperature of water, water starts to condense out and form clouds. These water clouds also contribute to the total optical depth. The opacity for clouds is calculated as:

$$\kappa_{clouds}(\lambda) [m^2/kg] = 130 Q_{ext}(\lambda) \quad (4.37)$$

where Q_{ext} is the extinction efficiency of the clouds. For Earth-like water clouds, we can follow the parametrization presented in [Marcq et al. \(2017\)](#). The density of the clouds is computed with the mass loading in [Kasting \(1988\)](#):

$$\rho_{clouds} = 4 \times 10^{-4} \rho_{gas} \quad (4.38)$$

where $\rho_{gas} = \rho_v + \rho_c$.

Q_{ext} depends on the wavelength as:

$$Q_{ext} = \begin{cases} 1 & \lambda \leq 20 \mu m \\ 3.26 \cdot \lambda^{-0.4} & \lambda > 20 \mu m \end{cases} \quad (4.39)$$

4.5.2. Single scattering albedo

The single scattering albedo is defined as the ratio of scattering efficiency to total extinction efficiency. The total extinction is a sum of both extinction by scattering and extinction by absorption. Therefore, a single scattering albedo of 1 indicates that all extinction is due to scattering, whereas a value of zero means that absorption dominates. In the clear atmospheric layers, the only source of scattering is Rayleigh scattering. Hence, the single scattering albedo can be calculated as:

$$\omega_{0, Rayleigh} = \frac{\tau_{Rayleigh}}{\tau_{clear} + \tau_{Rayleigh}} \quad (4.40)$$

where $\tau_{clear} = \tau_{CIA} + \tau_{line}$. Then when clouds are present, we need to calculate a weighted average of the total scattering:

$$\omega_{0, total} = \frac{\tau_{Rayleigh}}{\tau_{cloud} + \tau_{Rayleigh}} \times \omega_{0, Rayleigh} + \frac{\tau_{cloud}}{\tau_{cloud} + \tau_{Rayleigh}} \times \omega_{0, cloud} \quad (4.41)$$

where $\omega_{0, cloud}$ is the contribution of clouds to the total single scattering albedo. This is computed as the percentage of scattering extinction due to clouds with respect to absorption times the single scattering albedo of the cloud:

$$\omega_{0, cloud} = \frac{\tau_{cloud}}{\tau_{clear} + \tau_{cloud}} \omega_0 \quad (4.42)$$

The cloud single scattering albedo depends on wavelength as (Kasting 1988; Marcq 2012; Marcq et al. 2017):

$$\omega_0 = \begin{cases} 1 & \lambda \leq 2 \mu m \\ 1.24 \cdot \lambda^{-0.32} & \lambda > 2 \mu m \end{cases} \quad (4.43)$$

4.5.3. Phase function

The phase function that describes the scattering due to water clouds is the Henyey-Greenstein phase function (Liou 1980; Marcq et al. 2017):

$$P_{HG}(\mu) = \frac{1 - \tilde{g}^2}{(1 + \tilde{g}^2 - 2\tilde{g}\mu)^{3/2}} \quad (4.44)$$

The Henyey-Greenstein phase function, P_{HG} , is only dependent on the free variable μ . This variable is defined as $\mu = \cos(\theta)$, where θ is the angle the incoming radiation forms with the direction of the scattered light, i.e forward scattering corresponds to $\theta = 0^\circ$; backward scattering to $\theta = 180^\circ$ and scattering to the sides to $\theta = 90^\circ$ and $\theta = 270^\circ$ (see Figure 4.10). \tilde{g} is a free parameter of the Henyey-Greenstein phase function, known as the asymmetry factor, which is dependent on the composition of the cloud and the wavelength. For water clouds, Kasting (1988) and Marcq et al. (2017) use the following parametrization:

$$\tilde{g} = \begin{cases} 0.85 & \lambda \leq 10 \mu\text{m} \\ 1.40 \cdot \lambda^{-0.22} & \lambda > 10 \mu\text{m} \end{cases} \quad (4.45)$$

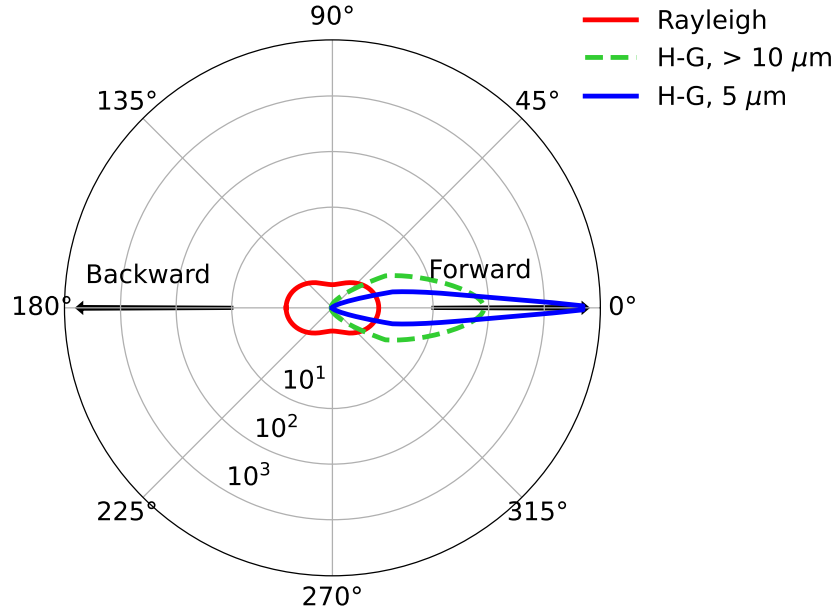


Figure 4.10.: Polar representation of the Henyey-Greenstein (at $\lambda = 5 \mu\text{m}$ and $\lambda > 10 \mu\text{m}$) and Rayleigh phase functions. Black arrows indicate the direction of forward and backward scattering.

To solve the RT equation, I need to expand the phase function in Legendre polynomials. The Legendre expansion of a function, $f(\mu)$, is defined as:

$$f(\mu) = \sum_{n=0}^{\infty} (2n+1) \psi_n P_n(\mu) \quad (4.46)$$

where P_n is the Legendre polynomial of n th order, and ψ_n is the n th order coefficient of the Legendre polynomial expansion. These coefficients are defined as the Legendre moments of the function we want to expand (Liou 1980; Boucher 1998), calculated as:

$$\psi_n = \frac{1}{2} \int_{-1}^1 f(\mu) P_n(\mu) d\mu \quad (4.47)$$

To calculate the Legendre moments of the Henyey-Greenstein function, Liou (1980) and Boucher (1998) substitute Eq. 4.44 in Eq. 4.47 and integrate. The zeroth moment of the Henyey-Greenstein phase function is zero, $\psi_0 = 0$. For the other orders, the n th moment is proven to be $\psi_n = \tilde{g}^n$. n is the number of streams considered in the RT solver DISORT, which in our case is $n = 4$ streams.

Clear atmospheric layers without clouds present Rayleigh scattering, whose phase function is (Liou 1980):

$$P_{\text{Rayleigh}}(\mu) = \frac{3}{4} (1 + \mu^2) \quad (4.48)$$

Similarly to the Henyey-Greenstein phase function, I calculate the moments of the Rayleigh phase function by substituting its analytic expression (Eq. 4.48) in Eq. 4.47. The demonstration of the

$n = 0$ to $n = 4$ moments, which are necessary for the input of the RT solver DISORT, can be found in the following subsection (Sect. 4.5.3.1). Its final values are $\psi_0 = 1$, $\psi_1 = 0$, $\psi_2 = 0.1$, $\psi_3 = 0$ and $\psi_4 = 0$.

Finally, when we have both contributions to scattering from clouds and Rayleigh scattering, we can calculate a weighted average of the moments of the phase functions:

$$\psi_{n, total} = \frac{\tau_{Rayleigh}}{\tau_{cloud} + \tau_{Rayleigh}} \times \psi_{n, Rayleigh} + \frac{\tau_{cloud}}{\tau_{cloud} + \tau_{Rayleigh}} \times \psi_{n, cloud} \quad (4.49)$$

4.5.3.1. Rayleigh phase function moments

To calculate the moments (or coefficients in Legendre expansion) of the Rayleigh phase function, I integrate for $n = 0$ to $n = 4$:

$$\psi_n = \frac{1}{2} \int_{-1}^1 P_{Rayleigh}(\mu) P_n(\mu) d\mu \quad (4.50)$$

$P_{Rayleigh}(\mu)$ is the Rayleigh scattering phase function, $P_{Rayleigh}(\mu) = \frac{3}{4}(1 + \mu^2)$, and $P_n(\mu)$ is the n th order Legendre polynomial, with the general formula (Liou 1980):

$$P_n(\mu) = \frac{1}{2^n \cdot n!} \frac{d^n}{d\mu^n} (\mu^2 - 1)^n \quad (4.51)$$

For $n = 0$, we have $P_0 = 1$. Therefore Eq. 4.50 yields:

$$\psi_0 = \frac{1}{2} \int_{-1}^1 \frac{3}{4} (1 + \mu^2) d\mu = \frac{3}{8} \left[\int_{-1}^1 d\mu + \int_{-1}^1 \mu^2 d\mu \right] = \frac{3}{8} \left[\mu \Big|_{-1}^1 + \frac{\mu^3}{3} \Big|_{-1}^1 \right] = 1$$

For $n = 1$, we have $P_1 = \mu$:

$$\psi_1 = \frac{1}{2} \int_{-1}^1 \mu \frac{3}{4} (1 + \mu^2) d\mu = \frac{3}{8} \left[\int_{-1}^1 \mu d\mu + \int_{-1}^1 \mu^3 d\mu \right] = \frac{3}{8} \left[\frac{\mu^2}{2} \Big|_{-1}^1 + \frac{\mu^4}{4} \Big|_{-1}^1 \right] = 0$$

For $n = 2$, $P_2 = \frac{1}{2}(3\mu^2 - 1)$:

$$\begin{aligned} \psi_2 &= \frac{1}{2} \int_{-1}^1 \frac{1}{2} (3\mu^2 - 1) \frac{3}{4} (1 + \mu^2) d\mu = \frac{3}{16} \left[\int_{-1}^1 2\mu^2 d\mu - \int_{-1}^1 d\mu + \int_{-1}^1 3\mu^4 d\mu \right] = \\ &= \frac{3}{16} \left[\frac{2}{3} \mu^3 \Big|_{-1}^1 - \mu \Big|_{-1}^1 + \frac{3}{5} \mu^5 \Big|_{-1}^1 \right] = \frac{1}{10} \end{aligned}$$

For $n = 3$, we obtain $P_3 = \frac{1}{2}(5\mu^3 - 3\mu)$:

$$\begin{aligned} \psi_3 &= \frac{1}{2} \int_{-1}^1 \frac{1}{2} (5\mu^3 - 3\mu) \frac{3}{4} (1 + \mu^2) d\mu = \frac{3}{16} \left[- \int_{-1}^1 3\mu d\mu + \int_{-1}^1 2\mu^3 d\mu + \int_{-1}^1 5\mu^5 d\mu \right] = \\ &= \frac{3}{16} \left[-\frac{3}{2} \mu^2 \Big|_{-1}^1 + \frac{1}{2} \mu^4 \Big|_{-1}^1 + \frac{5}{6} \mu^6 \Big|_{-1}^1 \right] = 0 \end{aligned}$$

Finally, for $n = 4$, the Legendre polynomial yields $P_4 = \frac{1}{8}(35\mu^4 - 30\mu^2 + 3)$:

$$\psi_4 = \frac{1}{2} \int_{-1}^1 \frac{1}{8} (35\mu^4 - 30\mu^2 + 3) \frac{3}{4} (1 + \mu^2) d\mu =$$

$$\begin{aligned}
&= \frac{3}{64} \left[-\int_{-1}^1 27\mu^2 d\mu + \int_{-1}^1 5\mu^4 d\mu + \int_{-1}^1 3 d\mu + \int_{-1}^1 35\mu^6 d\mu \right] = \\
&= \frac{3}{64} \left[(-9)\mu^3 \Big|_{-1}^1 + \mu^5 \Big|_{-1}^1 + 3\mu \Big|_{-1}^1 + 5\mu^7 \Big|_{-1}^1 \right] = 0
\end{aligned}$$

4.6. Atmospheric model validation

In this section, I compare the atmospheric model developed in this thesis, MSEIRADTRAN, to other well-documented atmospheric models. These models are those developed by [Marcq et al. \(2017\)](#), and [Pluriel et al. \(2019\)](#). In addition, the emission spectra and OLR generated with our model are compared with the publicly available atmospheric model petitRADTRANS (pRT) ([Mollière et al. 2019](#)), whose use is widely extended in the astrophysical community to compute emission and transmission spectra of exoplanets with clear and cloudy atmospheres.

In subsection 4.6.1, I compare the emission spectra, while in subsection 4.6.2, I show the reflection spectra and Bond albedo. Finally, in subsection 4.6.3 I re-calculate mass-radius relationships and adiabats in the water phase diagram for supercritical water planets by coupling MSEIRADTRAN with our interior structure model.

4.6.1. Emission spectra

I compare the outgoing fluxes and optical depths calculated with the approach from [Marcq et al. \(2017\)](#), with two different opacity k-tables: the one I use in this thesis (see Sect. 4.4.2), and the original opacity table of [Marcq et al. \(2017\)](#).

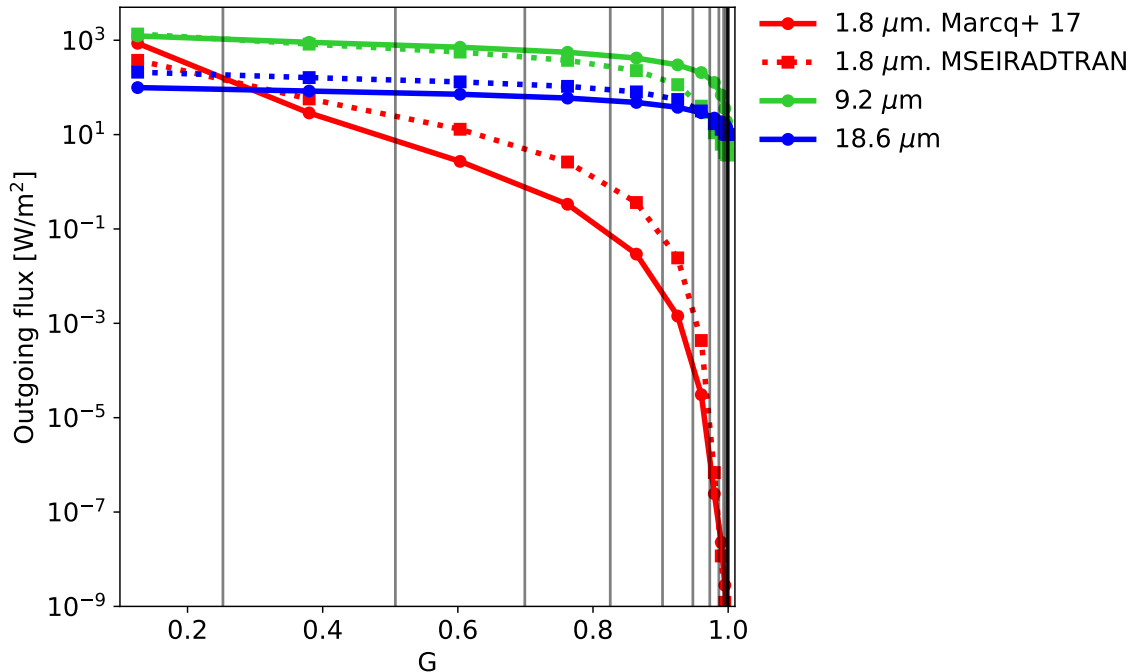


Figure 4.11.: Upward outgoing flux, $F(G)$, as a function of G -point in the k-correlated method for three bands (see text). Squares indicate the values calculated with the opacity data I compiled, while circles correspond to the data used by [Marcq et al. \(2017\)](#). Black lines mark the limits of the bins in G -space used by [Marcq et al. \(2017\)](#).

Figure 4.11 shows this comparison for three bands. These bands comprise the wavelengths between 1.7 to 2.0 μm (red), 8.3 to 10.0 μm (blue), and 17.2 to 20.0 μm (green). The fluxes between the two models are very similar in all three bands, indicating that differences are due to the opacity data in the k-tables. The outgoing fluxes in Fig. 4.11 are calculated by inputting the total optical depth to our RT solver, DISORT. To understand in what parts of the atmosphere the opacity data differ the most, we plot the difference between optical depths calculated with the two k-tables with the same approach to the k-correlated method. Figure 4.12 plots the difference in optical depth as a function of G in all atmospheric layers for the wavelength bin centered at $\lambda = 1.8 \mu\text{m}$. It can be seen that the greatest differences come from the layers at higher pressures, which is expected since the maximum pressure of the tabulated k-tables used by Marcq et al. (2017) is 100 bar. Therefore, the optical depth between 100 and 300 bar is underestimated, since it is out of its validity range. The opacity tables provided by HELIOS-K covers pressures up to 300 bar, which is the appropriate maximum pressure of validity for our interior-atmosphere coupling.

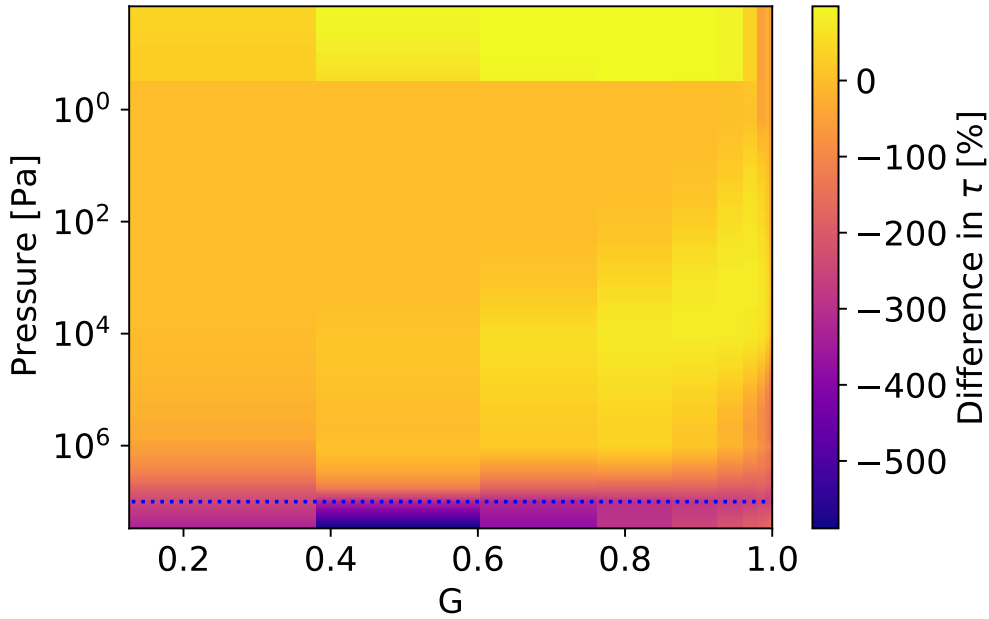


Figure 4.12.: Difference in optical depths calculated with the opacity data used in this work and that of Marcq et al. (2017). Negative values indicate that the optical depth in Marcq et al. (2017) is underestimated compared to MSEI. The atmospheric level $P = 100$ bar is indicated in dashed blue line.

After checking that the k-table is being handled consistently in our atmospheric model, the formalism and k-table data analysis described in Sect. 4.4.1 is carried out for all spectral bands in the IR. The IR emission spectrum of the same atmosphere is shown in Fig. 4.13 for different versions of the atmospheric model for the same planet. I choose the case of HD 207897 (Heidari et al. 2022), for which we perform a interior-atmosphere analysis based on mass and radius data (see Sect. 6.3.4). Its mass is approximately $M = 15 M_{\oplus}$, while its total radius is $R = 2.5 R_{\oplus}$. I approximate its bulk radius for the atmosphere model comparison to $R = 1.9 R_{\oplus}$, which is the radius that encloses the core and mantle to provide a realistic bulk density for a sub-Neptune. I recall that the atmospheric model in Marcq et al. (2017) (black dots) adopts the k-correlated approach, where the sum over G -points requires running the RT solver in the process, while the k-uncorrelated method ends the sum over G -points before reaching the RT solver. Green triangles correspond to the the k-correlated

method with the opacity data I compiled for MSEI, which includes the POKAZATEL (Polyansky et al. 2018) line list for water. The k-uncorrelated version of MSEI (red crosses) uses the same opacity data (see Sect. 4.4.1).

In the grey model (grey crosses), the line opacity is constant with wavelength, being $0.01 \text{ m}^2/\text{kg}$ for H_2O , and $10^{-4} \text{ m}^2/\text{kg}$ for CO_2 . These constant grey opacities are benchmarked with non-grey atmospheric models (Nakajima et al. 1992; Marcq et al. 2017). In the case of water, a grey opacity of $0.01 \text{ m}^2/\text{kg}$ is representative of the opacity of water in the $8\text{-}20 \mu\text{m}$ spectral window at the Standard Reference Point (Ingersoll 1969). The two versions of MSEI and Marcq et al. (2017) agree well in all spectral wavelengths. The grey line model tends to slightly underestimate the emission in the mid-IR, which is expected because the grey opacity is only valid at a limited range of wavelengths.

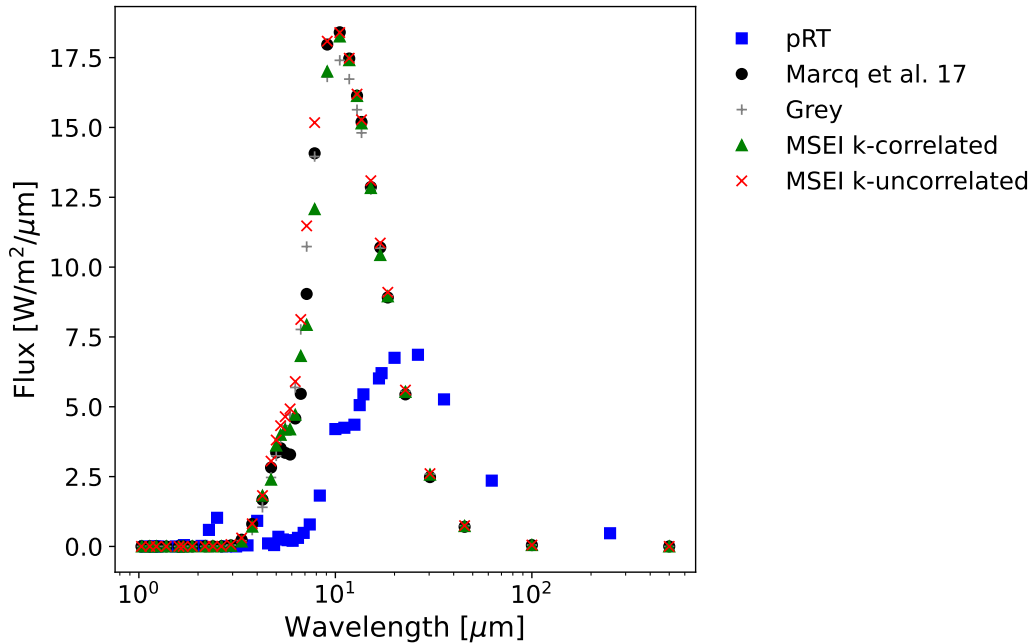


Figure 4.13.: IR emission spectrum of a planet with $T_{surf} = 1000 \text{ K}$, $P_{surf} = 300 \text{ bar}$ and 99% water with a mass of $M = 15 M_{\oplus}$ and a radius of $R = 1.875 R_{\oplus}$. The grey line model emission and the spectrum obtained by *petitRADTRANS* (pRT) are also shown.

For comparison purposes, I also show the emission spectrum calculated with *petitRADTRANS* (pRT) (Mollière et al. 2019), which is a widely used Python package to generate emission and transmission spectra of exoplanets with clear and cloudy atmospheres⁶. The k-uncorrelated version of MSEI yields slightly higher emission than Marcq et al. (2017) and the other version of MSEI at approximately $10 \mu\text{m}$. However, the difference between the k-uncorrelated version and the k-correlated models (MSEI and Marcq et al. (2017)) is smaller than that between these models and pRT, despite being a k-correlated model as well. I choose to obtain the spectrum in pRT in low-resolution ("c-k") mode to use the k-correlated method instead of the line-by-line method, so it uses the same method as the other models shown in Fig. 4.13. Differences between the MSEI models and pRT are due to the different CIA and clouds opacities. These differences can also be seen in Table 4.5, where I display the values of the bolometric OLR. The OLR of both versions of MSEI have a very small difference (11 W/m^2).

⁶<https://gitlab.com/mauricemolli/petitRADTRANS>

I maintain the mass, radius and surface pressure of the planet constant, and change the surface temperature to calculate the OLR as a function of T_{surf} (Figure 4.14). The OLR calculated by the k-uncorrelated MSEI agrees well with [Marcq et al. \(2017\)](#). The grey line model tends to underestimate the OLR compared to the non-grey models, especially in the surface temperature range of 2000 to 3000 K, when the spectrum of water presents more lines and the opacity varies very rapidly with wavelength. The inclusion of the optical spectral bands contributes to the OLR with $\approx 10^{-4} \text{ W/m}^2$, which is negligible compared to the OLR in the IR seen in Table 4.5.

Model	OLR [W/m^2]
Marcq et al. (2017)	269.91
Grey	270.89
MSEI k-correlated	268.97
MSEI k-uncorrelated	280.36
pRT	104.58

Table 4.5.: Comparison of the bolometric OLR for different models. This corresponds to an atmosphere of 99% water and a surface temperature and pressure of 1000 K and 300 bar, respectively. The bulk mass and radius of the planet (from center to surface) are $15 M_{\oplus}$ and $1.875 R_{\oplus}$.

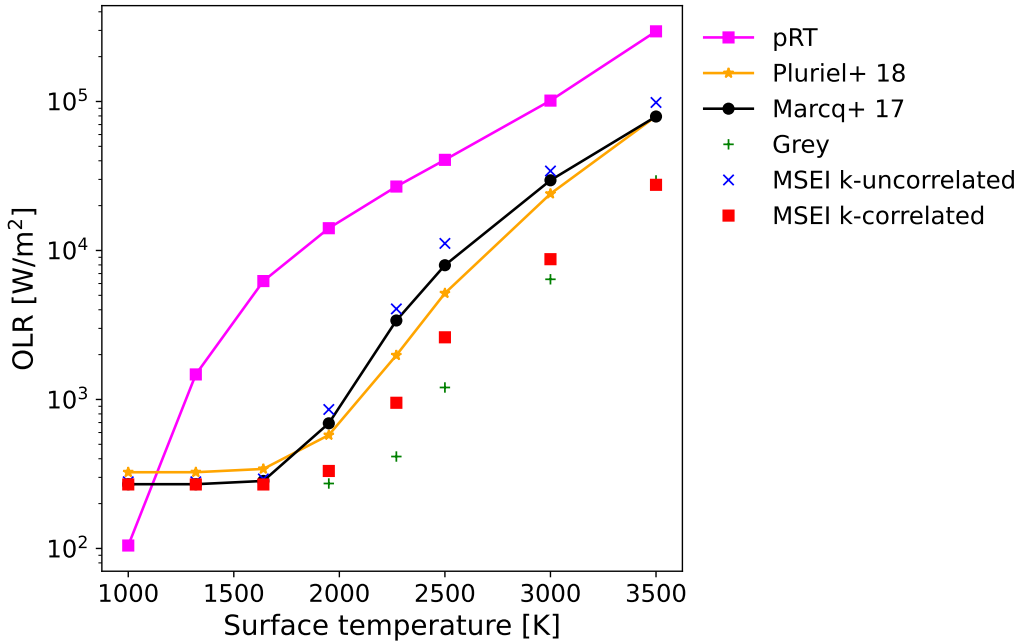


Figure 4.14.: OLR as a function of surface temperature for planets with $P_{surf} = 300 \text{ bar}$ and 99% water with a mass of $M = 15 M_{\oplus}$ and a radius of $R = 1.875 R_{\oplus}$. For comparison, the OLR estimated with the atmospheric models of [Marcq et al. \(2017\)](#), [Pluriel et al. \(2019\)](#) and pRT ([Mollière et al. 2019](#)) are also shown.

So far, I have considered the correlated method to mix the k-tables of water and CO_2 to obtain a 99% water and 1% CO_2 composition. Figure 4.15 shows a comparison of the OLR calculated with both mixing approximations (correlated and uncorrelated) for a CO_2 -dominated atmosphere. In addition, I display three reference models: [Marcq et al. \(2017\)](#), [Pluriel et al. \(2019\)](#) and pRT. The

correlated approximation agrees well with Marcq et al. (2017) and Pluriel et al. (2019) at $T_{surf} \leq 2500$ K, while at higher temperatures the OLR is within the uncertainties of these two models and pRT. The difference between pRT and the other two models could be due to the different opacity data and the use of the uncorrelated mixing method by pRT (Mollière et al. 2019). However, the uncorrelated approximation in MSEI seems to overestimate the OLR with respect to all three reference models. For this reason, I henceforth use the correlated approximation to mix the k-tables of H₂O and CO₂.

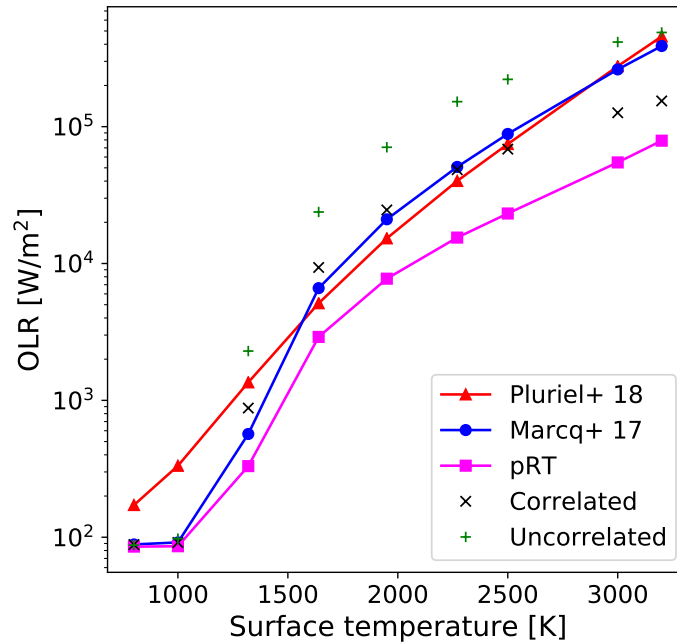


Figure 4.15.: OLR as a function of surface temperature for a 99% CO₂ and 1% water atmosphere, with surface pressure $P_{surf} = 300$ bar. The bulk mass and radius of the planet are $1 M_{\oplus}$ and $1 R_{\oplus}$, respectively.

4.6.2. Reflection spectra and Bond albedo

As explained in Sect. 4.1, DISORT obtains the reflectivity as a function of SZA and spectral band. This means that to compute the final bolometric Bond albedo, the reflectivity needs to be integrated over SZA and wavelength. Figure 4.16 shows the dependence of the reflectivity on the cosine of the SZA for a water-dominated atmosphere with clouds. At optical wavelengths, the albedo is high and approximately constant due to clouds, while at IR wavelengths the albedo is higher at low SZA. Despite clouds being present in the atmosphere, at wavelengths $\lambda > 2 \mu\text{m}$ the single scattering albedo of clouds starts to decrease (see section 4.5.2), and Rayleigh scattering dominates at these wavelengths.

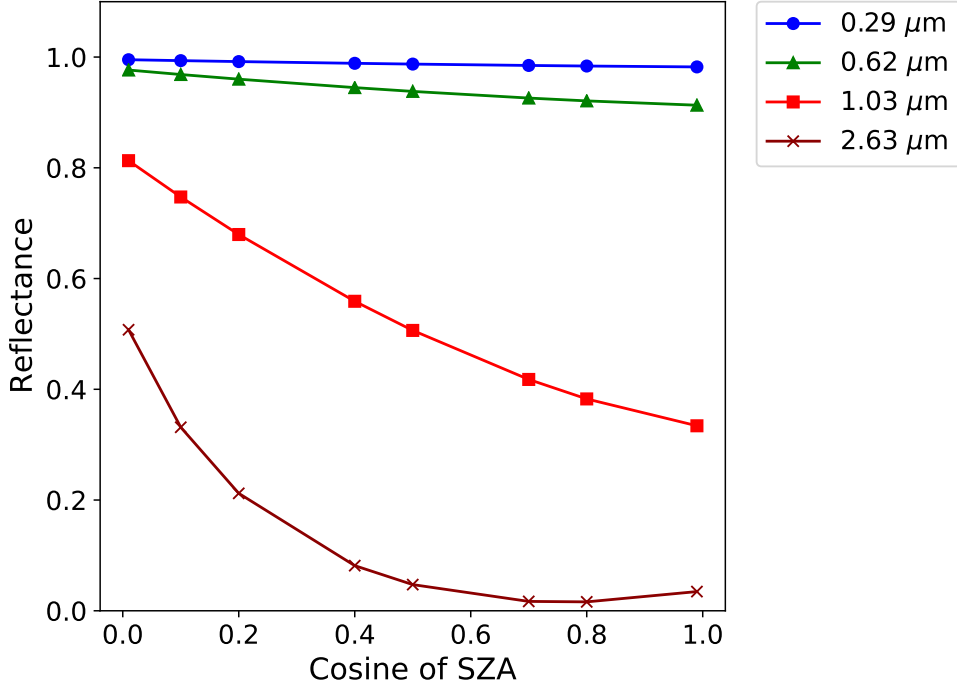


Figure 4.16.: Reflectivity as a function of solar zenith angle (SZA) for four different spectral bands. This corresponds to a 99% water atmosphere with surface temperature $T_{surf} = 1000$ K, surface pressure $P_{surf} = 300$ bar and a bulk radius and mass of $1 R_{\oplus}$ and $5 M_{\oplus}$.

Once the reflectivity is integrated as shown in Eq. 4.4, we obtain the reflection spectrum. Figure 4.17 shows the albedo as a function of wavelength for two different surface temperatures. At $T_{surf} = 1000$ K, the albedo at short wavelengths is dominated by clouds and Rayleigh scattering, yielding a very high albedo of ~ 0.92 at $\lambda < 0.5 \mu\text{m}$. At longer wavelengths, we can identify the water lines at 0.7 , 1.1 , 1.4 and $1.9 \mu\text{m}$ that are also characteristic of Earth’s reflection spectrum (Des Marais et al. 2002). We also observe the water spectral line at $2.7 \mu\text{m}$, mentioned by Pluriel et al. (2019) in their figure 3. The spectral binning used to calculate the reflection spectra is similar to that of Pluriel et al. (2019) in the optical. For $T_{surf} = 2500$ K, no clouds form. This affects the albedo at short wavelengths, with values of 0.8 - 0.6 , which are lower than the cloudy atmosphere case. Only the water line at $0.7 \mu\text{m}$ can be identified, since at high surface temperatures the atmosphere becomes very opaque.

I obtain and integrate the reflection spectra of a water-dominated atmosphere at different surface temperatures (see Eq. 4.5). I compare the final bolometric Bond albedo calculated by MSEI with that of Pluriel et al. (2019) in Fig. 4.18. Both models agree on the albedos since we use similar parametrizations of the single scattering albedo for clouds and Rayleigh scattering.

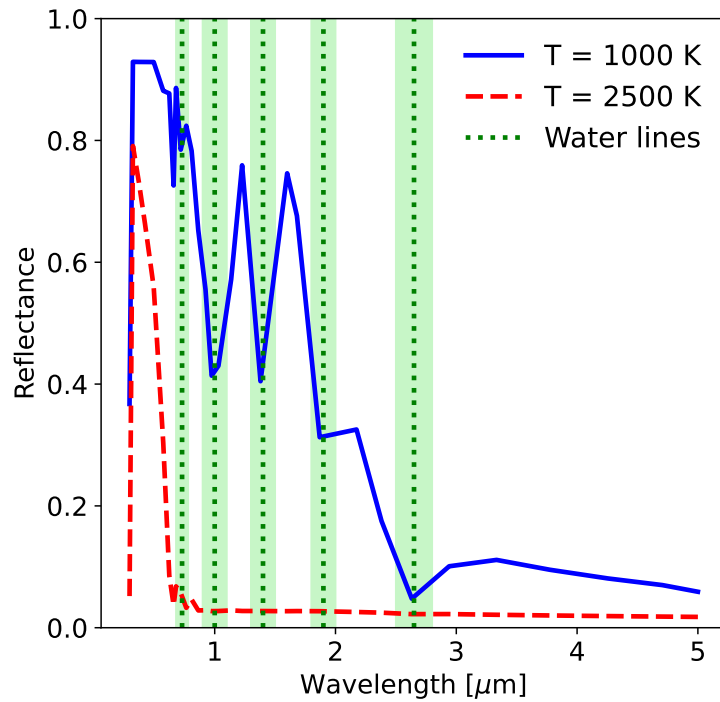


Figure 4.17.: Reflection spectrum for 99% water atmospheres with surface temperatures of 1000 K and 2500 K, surface pressures of 300 bar and bulk mass and radius of $1 R_{\oplus}$ and $5 M_{\oplus}$. Green areas indicate the water lines in the optical and near-IR.

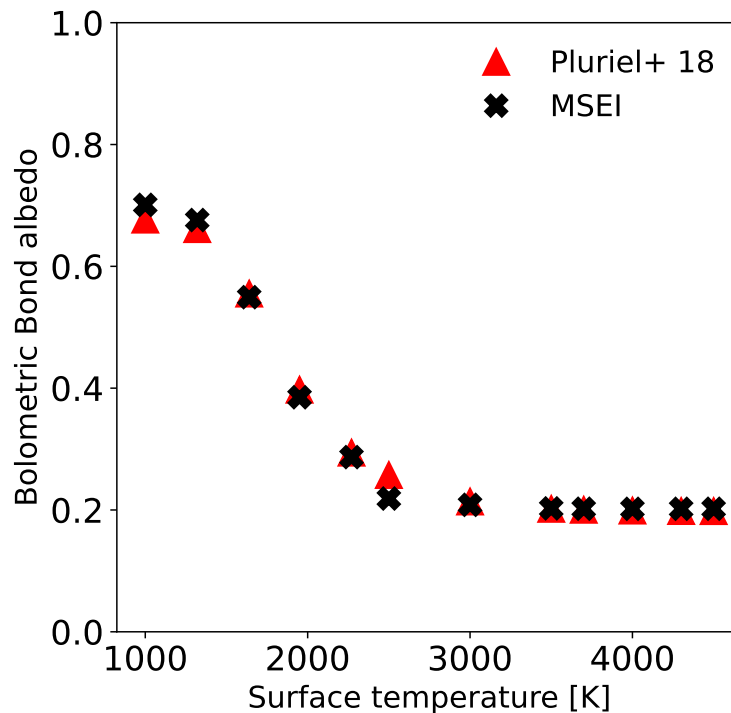


Figure 4.18.: Bolometric Bond albedo as a function of surface temperature. I assume 99% water atmospheres with surface pressures of 300 bar, and bulk mass and radius of $1 R_{\oplus}$ and $5 M_{\oplus}$. The bolometric albedo is computed considering a Sun-like star with $T_{\star} = 5777$ K.

4.6.3. Mass-radius relations and water phase diagram

I couple the MSEI atmospheric model, MSEIRADTRAN, with the interior model using the algorithm described in Chapter 3.

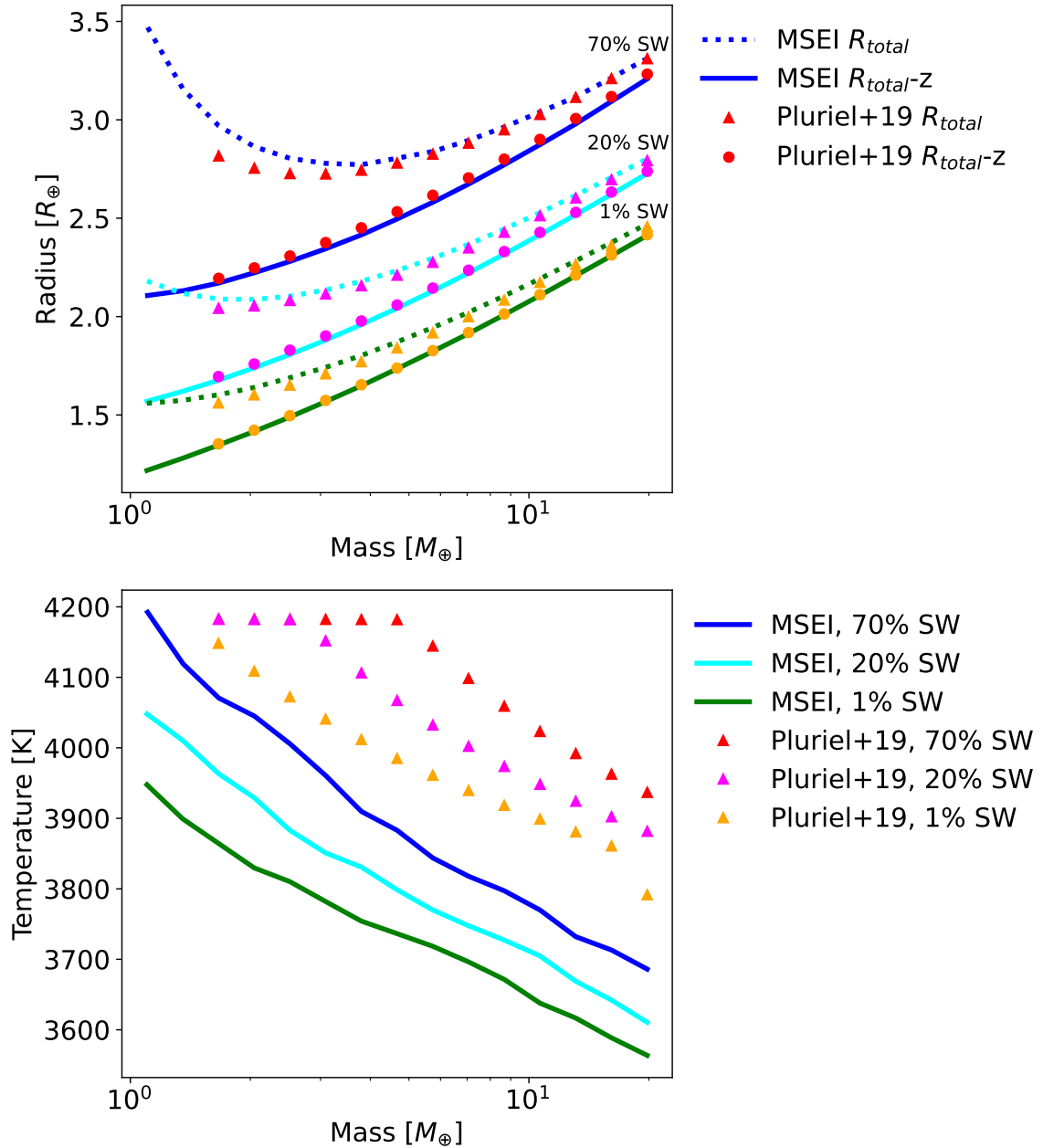


Figure 4.19.: Upper panel: Mass-radius relationships for a planet with a water-dominated atmosphere orbiting a Sun-like star at $a_d = 0.05$ AU. Dashed lines indicate the total radius calculated by MSEI (k-uncorrelated), while the solid line corresponds to the interior radius, which comprises the core, mantle, and supercritical water (SW). Triangles and circles indicate the total radius and the interior radius obtained when the interior model is coupled with the atmospheric model of [Pluriel et al. \(2019\)](#), respectively. Lower panel: Temperature at the 300 bar interface as a function of planetary mass.

I derive mass-radius relationships for a planet orbiting a Sun-like star ($T_{eff} = 5777$ K, $R_{star} = 1 R_{\odot}$)

at $a_d = 0.05$ AU, assuming a water-dominated atmosphere (99% water). This corresponds to an equilibrium temperature $T_{eq} (A_B = 0) = 1246$ K under the assumption of a black body, and $T_{eq} = 1165$ K with the albedo calculated by our model for a water-dominated atmosphere, which is $A_B = 0.237$. The pressure at the base of the supercritical hydrosphere exceeds the critical point of water for these compositions ($WMF > 0.01$), therefore the pressure at the base of the atmosphere is $P = 300$ bar, which serves as the interface between the atmosphere and the interior.

Figure 4.19 shows the mass-radius relations for three compositions: 70% supercritical water (SW) and 30% mantle, 20% SW and 80% mantle and 1% SW and 99% mantle. In other words, I adopt a constant core mass fraction of zero, for three values of the water mass fraction, $WMF = 0.70, 0.20$ and 0.01 . The contributions to the total radius are two: the interior, which is dependent on the expansion of the supercritical layer (solid lines and circles), and the atmosphere thickness (dotted lines and crosses). The mass-radius relations derived with our atmospheric model and Pluriel et al. (2019) are in good agreement, with differences below 1% in radius. The slight differences in the interior radius, $R_{total} - z$, are due to the different temperatures at the bottom of the atmosphere calculated by the two atmospheric models. Similarly, the differences in total radius, especially at low masses and high water mass fractions, can be explained by the different bottom temperatures and the EOSs used in the atmospheric models, which yield slightly different atmospheric thicknesses. I compare the temperatures at the bottom of the atmosphere ($P = 300$ bar) provided by the two atmospheric models. Maximum differences are approximately 230 K (see Fig. 4.19, lower panel). Note that the constant surface temperature at masses below $5 M_{\oplus}$ in the model by Pluriel et al. (2019) are due to the limit in the data grid I used for this plot, they do not have a physical interpretation.

Fig. 4.20 shows the OLR and Bond albedo as a function of bottom temperature for our atmospheric model and that of Pluriel et al. (2019). The trends of the Bond albedo are very similar for both models. The slight discrepancy in bottom temperature is caused by the small difference in OLR between the two models. The inaccuracy in OLR in the k-uncorrelated version of MSEI causes differences of less than 1% in radius compared to the k-correlated model of Pluriel et al. (2019). Therefore, the k-uncorrelated MSEI may be used to calculate the radius of water-rich planets within our MCMC retrieval framework, since fast computations of the OLR and Bond albedo are necessary. The k-uncorrelated approach is an alternative to grey models, whose difference in OLR with k-correlated models is greater than that obtained with the k-uncorrelated model. However, the k-uncorrelated approximation should not be used to compute spectra at higher resolution since the planetary emission differs by a factor of 2 compared to a k-correlated method, which yields a significant difference in resolutions of $R = 200$ to 300 (see Sect. 6.4).

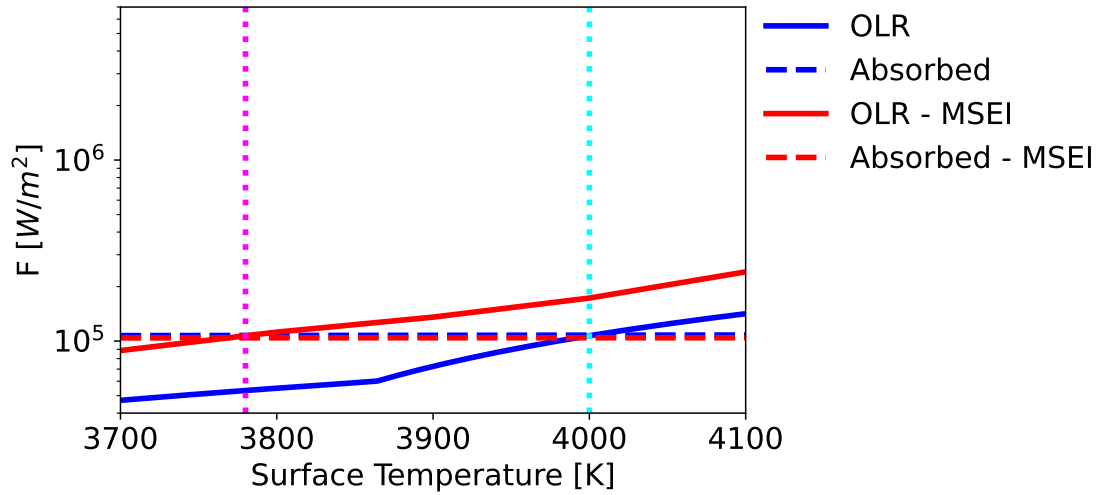


Figure 4.20.: Outgoing Longwave Radiation (OLR) and absorbed flux as a function of bottom atmospheric temperature, calculated with our atmospheric model (red) and [Pluriel et al. \(2019\)](#) (blue). Dotted lines indicate the temperature at which the OLR and the absorbed flux are equal, meaning the atmosphere is in radiative-convective equilibrium.

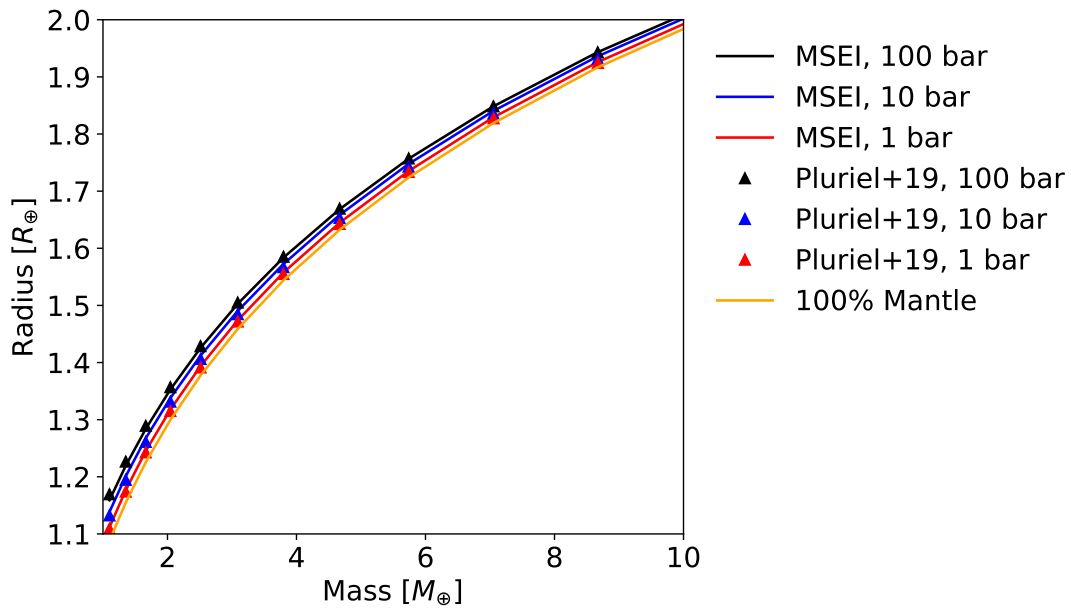


Figure 4.21.: Mass-radius relationships for a planet with water-dominated atmospheres whose stellar host is similar to TRAPPIST-1 ($R_{\star} = 0.117 R_{\odot}$, $T_{\star} = 2560$ K) with a semi-major axis $a_d = 0.01$ AU.

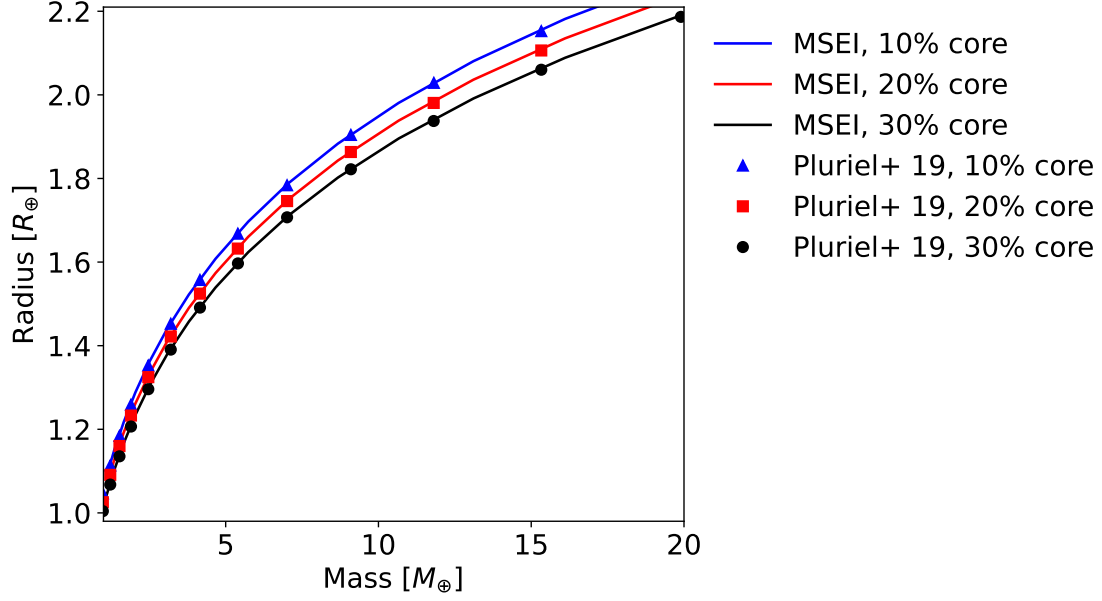


Figure 4.22.: Mass-radius relationships for a planet with CO_2 whose stellar host is similar to TRAPPIST-1 ($R_\star = 0.117 R_\odot$, $T_\star = 2560 \text{ K}$) with a semi-major axis $a_d = 2.27 \times 10^{-2} \text{ AU}$ (TRAPPIST-1 d).

Similarly to the mass-radius relationships for atmospheres with surface pressure of 300 bar, I derive the MR relationships for planets with water-dominated atmospheres and surface pressures $P_{surf} = 100 \text{ bar}$, 10 bar and 1 bar. These can be seen in Fig. 4.21, which correspond to a planet orbiting at 0.01 AU an M dwarf similar to TRAPPIST-1. In Fig. 4.22, I derive the MR relationships of a planet with a CO_2 -dominated atmosphere with $P_{surf} = 300 \text{ bar}$. The stellar host and the semi-major axis corresponds to those of TRAPPIST-1 and TRAPPIST-1 d, respectively. It can be seen that these MR relationships agree well between our model and [Pluriel et al. \(2019\)](#).

Finally, to test the consistency between the interior model (Sect. 2) and our atmospheric model, I derive the pressure-temperature profile for a planet of $M = 5 M_\oplus$, $R = 1.5 R_\oplus$, and a composition of 50% water and 50% mantle. I assume a Sun-like star for the calculation of the Bond albedo, and different semi-major axes, which yields different surface temperatures for an atmosphere in radiative-convective equilibrium. The adiabatic profiles in Fig. 4.23 are smooth, and the slope at the pressure at which the interior and the atmosphere are coupled (300 bar) is similar between the interior and the atmosphere. In the case of the interior, the adiabatic gradient is calculated using the EOS from [Mazevet et al. \(2019\)](#), whereas in the atmosphere the adiabatic gradient is computed as described in Section 4.2. Note that the smooth transition in the adiabat between the interior and the atmosphere is due to assuming an isentropic profile in both. This means that we consider that heat is transported in the interior and in the atmosphere by an adiabatic and reversible process, $dS/dr = 0$, where S corresponds to the entropy, and r to the planetary radius.

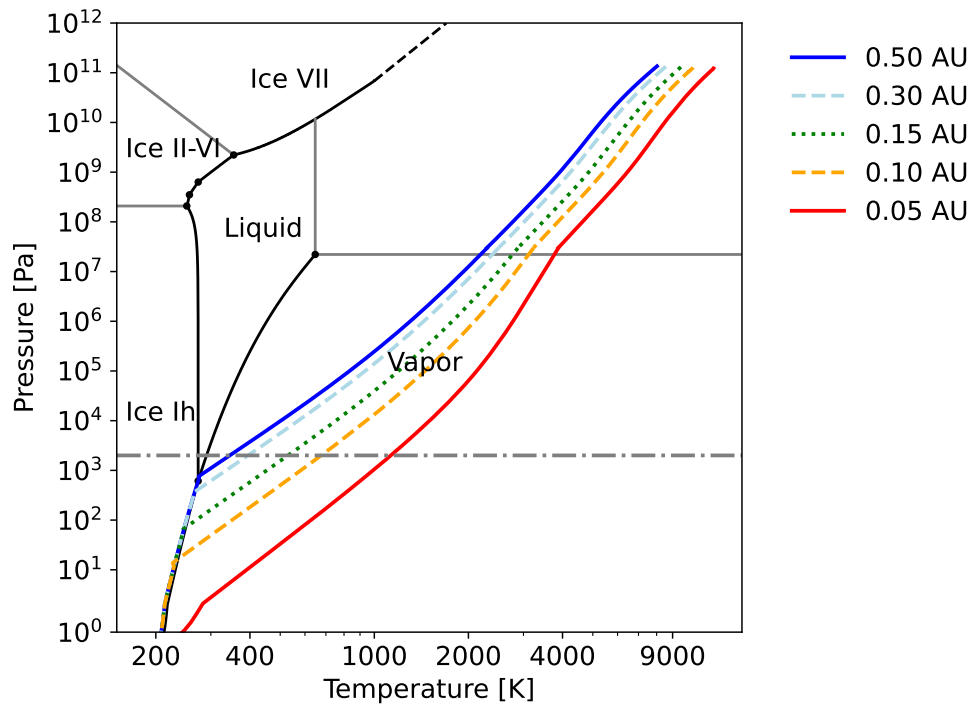


Figure 4.23.: Adiabatic pressure-temperature profiles for the complete hydrospheres (interior and atmosphere) of water-rich planets. I assume a planet with $M = 5 M_{\oplus}$, $R = 1.5 R_{\oplus}$, a composition of 50% water and 50% mantle, and a Sun-like star. The dotted-dashed grey line indicates the transit pressure level $P = 20$ mbar.

5. Bayesian statistics

The interior-atmosphere model described in Chapters 2 to 4 is a forward model. This means that given a certain set of planetary variables, such as the CMF, WMF and total mass, the forward model obtains the total radius that should be observed. Since we currently have available mass and radius observations for many low-mass planets, we invert the forward interior-atmosphere model using the Markov chain Monte Carlo (MCMC) method. The aim of the MCMC method is to infer the posterior distribution functions, mean and uncertainties of the compositional parameters of super-Earths and sub-Neptunes. In contrast to comparing the mass and radius with the mass-radius relations, the MCMC provides an estimate of the uncertainties of the core mass fraction and water mass fraction based on the error bars of the observables.

In Sect. 5.1 we describe the basic MCMC Bayesian algorithm and sampling, while in Sect. 5.2 we explain how the sampling of the compositional parameters (CMF and WMF) is improved with an adaptive step size. Finally, in Sect. 5.3 we validate our MCMC implementation with two cases: the super-Earth CoRoT-7 b, and the sub-Neptune TOI-220 b.

5.1. Algorithm

I use the Bayesian MCMC algorithm described in [Dorn et al. \(2015\)](#) to compute the posterior PDF of the non-observables from data and prior information. Our model parameters are the planetary mass, M ; the CMF, x_{core} ; and the WMF, x_{H2O} . Therefore, one single model is determined by these three parameters as $\mathbf{m} = \{M, x_{core}, x_{H2O}\}$. The available data are the total mass M , the total radius R , and the Fe/Si abundance, $\mathbf{d} = \{M_{obs}, R_{obs}, Fe/Si_{obs}\}$, whose observational errors are $\sigma(M_{obs}), \sigma(R_{obs}), \sigma(Fe/Si_{obs})$, respectively. When the Fe/Si mole ratio is not considered in the inverse problem, the data is reduced to only the total planetary mass and radius, $\mathbf{d} = \{M_{obs}, R_{obs}\}$. The prior information consists on a Gaussian distribution centered on the mean value of the observed mass, with a standard deviation equal to the observational uncertainty. For the CMF, I use a uniform distribution between 0 and 1. In some cases, we further constrain the CMF setting a maximum value in its prior distribution. This maximum value is 0.75, which corresponds to the highest possible estimated Fe/Si ratio of the proto-Sun ([Lodders et al. 2009](#)). We only assume this maximum limit for planets whose density is low enough to suggest that they have not experienced post-formation events that could increase their Fe content, such as mantle evaporation or giant impacts that can strip away the mantle. In the case of the WMF, I consider a uniform distribution between 0 and 0.8. This maximum limit is the maximum volatile content found in comets of the Solar System ([McKay et al. 2019](#)).

The MCMC scheme first starts by drawing a value for each of the model parameters from their prior distributions, which we denote as $\mathbf{m}_{old} = \{M_{old}, x_{core,old}, x_{H2O,old}\}$. The interior model calculates the planetary radius and Fe/Si abundance that corresponds to these model parameters, which is $\mathbf{g}(\mathbf{m}_{old}) = \{R_{old}, M_{old}, Fe/Si_{old}\}$. The likelihood of this set of model parameters is calculated with equation 6 in [Dorn et al. \(2015\)](#), which in our case is:

$$L(\mathbf{m}_{old} | \mathbf{d}) = C \exp \left(-\frac{1}{2} \left[\left(\frac{(R_{old} - R_{obs})}{\sigma(R_{obs})} \right)^2 + \left(\frac{(M_{old} - M_{obs})}{\sigma(M_{obs})} \right)^2 + \left(\frac{(Fe/Si_{old} - Fe/Si_{obs})}{\sigma(Fe/Si_{obs})} \right)^2 \right] \right), \quad (5.1)$$

C is the normalization constant of the likelihood function, defined as:

$$C = \frac{1}{(2\pi)^{3/2} [\sigma^2(M_{obs}) \cdot \sigma^2(R_{obs}) \cdot \sigma^2(Fe/Si_{obs})]^{1/2}} \quad (5.2)$$

Then a new model is proposed, $\mathbf{m}_{new,1} = \{M_{new,1}, x_{core,new,1}, x_{H2O,new,1}\}$, by adding a perturbation to the previous model parameters, \mathbf{m}_{old} . The absolute difference between the old and the new model is lower than a fixed step, which is the maximum size of the perturbation. By limiting the size of the perturbation, we are certain that the new state is uniformly bounded and centered around the state of the old model. The maximum step size in the walker regulates the acceptance rate of the MCMC. The acceptance rate is defined as n_{acc}/n_{prop} , which is the number of accepted models over the number of total proposed models. We choose our maximum step size so that our acceptance rate is above 20%. The output of the interior model for this new set of model parameters is obtained as $\mathbf{g}(\mathbf{m}_{new,1}) = \{R_{new,1}, M_{new,1}, Fe/Si_{new,1}\}$. The likelihood function for this new model is then computed as:

$$L(\mathbf{m}_{new,i} | \mathbf{d}) = C \exp \left(-\frac{1}{2} \left[\left(\frac{(R_{new,i} - R_{obs})}{\sigma(R_{obs})} \right)^2 + \left(\frac{(M_{new,i} - M_{obs})}{\sigma(M_{obs})} \right)^2 + \left(\frac{(Fe/Si_{new,i} - Fe/Si_{obs})}{\sigma(Fe/Si_{obs})} \right)^2 \right] \right), \quad (5.3)$$

The normalization constant is given by Eq. 5.2, similarly to the likelihood expression for the old set of model parameters.

We then compute the log-likelihoods of the old and the new model as $l(\mathbf{m}_{old} | \mathbf{d}) = \log(L(\mathbf{m}_{old} | \mathbf{d}))$, and $l(\mathbf{m}_{new,i} | \mathbf{d}) = \log(L(\mathbf{m}_{new,i} | \mathbf{d}))$, respectively. The probability with which $\mathbf{m}_{new,i}$ is accepted is:

$$P_{accept} = \min \left\{ 1, e^{(l(\mathbf{m}_{new,i} | \mathbf{d}) - l(\mathbf{m}_{old} | \mathbf{d}))} \right\} \quad (5.4)$$

Consecutively, a random number is drawn from a uniform distribution between 0 and 1. If P_{accept} is greater than this random number, $\mathbf{m}_{new,1}$ is accepted and the chain moves to this set of model parameters, starting the next chain $n + 1$ with $\mathbf{m}_{old} = \mathbf{m}_{new,1}$. Otherwise, the chain remains in \mathbf{m}_{old} and a different set of model parameters is proposed, $\mathbf{m}_{new,2}$. The accepted models are stored, and values of their parameters conform the PDF that will enable us to estimate their mean and uncertainties.

5.2. Adaptive MCMC

In the MCMC algorithm described in the previous section, the random walker uses a uniform distribution to choose the next state where it is going to move in the parameter space of the CMF

and WMF. This is called a naive walk (Mosegaard and Tarantola 1995), in which all points in the parameter space have a probability of being chosen proportional to their number of neighbours. This poses the following problem: for the states whose CMF or WMF are close to 0 or 1, they are less likely to be sampled in the random walk, because they have less neighbours than the central values (i.e the central range 0.2-0.7). In Sect. 5.1, I commented that planets with compositions of CMF or WMF equal to 1 are unlikely. However, many highly-irradiated rocky planets might present low-mass atmospheres, which correspond to a WMF close to 0.

To compensate the lower probability of being chosen in the limiting states of the prior, I propose to use an adaptive step size in the walker. This consists on adapting the maximum size of the perturbation used to generate a new model instead of using a fixed value everywhere in the parameter space. This adaptive step size will tend to decrease in the limiting areas of the prior (i.e low WMF states) and have its greatest value in the centre of the prior (WMF = 0.5). The self-adjusting step size is usually carried by a transformation of the parameter space, which ranges from exponential to spherical transforms (Chaudhry et al. 2021). In this work, we choose to implement the self-adjusting logit transform (SALT), proposed by Director et al. (2017). This is based on the element-wise transformation known as stick-breaking process, where a stick of length 1 is broken into pieces of different lengths. The more pieces the stick is broken into, the more concentrated are the distributions. The SALT transform is publicly available in the *SALTSampler* R package¹, which eases its implementation in Python for our own forward model.

In our particular problem, we consider a simplex of $k = 3$ categories. The simplex is defined as the mathematical space where each point represents a probability distribution between a finite number of mutually exclusive events. Each of these events is called a category. In our case, the 3 categories are the CMF, the WMF, and the mantle mass fraction (MMF), which fulfill the requirement $x_{Fe} + x_{H2O} + x_{mantle} = 1$. Thus, the simplex is represented by an object of $k-1 = 2$ dimensions, which is the ternary diagram (see Fig. 2.4 for an example). For clarity, in this section I redefine the model parameter notation as $\mathbf{m}_{new} = \{M_{new}, \mathbf{x}'\}$, where $\mathbf{x}' = \{x'_{core}, x'_{H2O}, x'_{mantle}\}$. Similarly, $\mathbf{m}_{old} = \{M_{old}, \mathbf{x}\}$, where $\mathbf{x} = \{x_{core}, x_{H2O}, x_{mantle}\}$.

I first start the calculation of the new proposal, \mathbf{x}' , by choosing one of the categories randomly. Then the proposal of its new value is obtained as:

$$x'_i = \text{ilogit}[\text{logit}(x_i) + h_i \times Z] \quad (5.5)$$

where h_i is a real positive constant chosen arbitrarily, and Z is a value sampled from the normal distribution, $\mathcal{N}(0, 1)$. The logit transformation of a variable consists of:

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) \quad (5.6)$$

And its inverse transformation:

$$\text{ilogit}(x) = \frac{e^x}{1 + e^x} \quad (5.7)$$

Once I have the proposal of the first category, I select the second category randomly from the two that were left, and calculate its proposal by rescaling it with a randomly uniform proportional constant:

¹<https://rdrr.io/cran/SALTSampler/man/SALTSampler-package.html>

$$x'_j = (1 - x'_i) \left(\frac{x_j}{1 - x_i} + U_j \right) \quad (5.8)$$

where U_j is the constant sampled from the uniform distribution $U(-\epsilon, \epsilon)$, and x'_i is the proposal of the first category (Eq. 5.5). The constant ϵ is selected arbitrarily close to zero, which restricts the proposals to our simplex. The proposal for the remaining category is calculated as:

$$x'_l = 1 - x'_i - x'_j \quad (5.9)$$

In addition, the acceptance probability is modified with respect to Eq. 5.4. In the most general case of the Metropolis rule, the acceptance probability is $P_{accept} = \min\{1, r_{accept}\}$, where r_{accept} is the product of the posterior density ratio and the transition probability ratio (Mosegaard and Tarantola 2002; Director et al. 2017):

$$r_{accept} = \frac{p(\mathbf{x}'|\mathbf{d}) q(\mathbf{x}|\mathbf{x}')}{p(\mathbf{x}|\mathbf{d}) q(\mathbf{x}'|\mathbf{x})} \quad (5.10)$$

where p is the posterior density, and q is the transition probability.

If in the proposal calculation, there is no transformation, we retrieve Eq. 5.4, because $q(\mathbf{x}|\mathbf{x}') = q(\mathbf{x}'|\mathbf{x})$, and $\frac{p(\mathbf{x}'|\mathbf{d})}{p(\mathbf{x}|\mathbf{d})} = e^{l(\mathbf{x}'|\mathbf{d}) - l(\mathbf{x}|\mathbf{d})}$. For the logit transform, the transition term is computed as (Director et al. 2017):

$$\frac{q(\mathbf{x}|\mathbf{x}')}{q(\mathbf{x}'|\mathbf{x})} = \frac{x'_i}{x_i} \left[\frac{1 - x'_i}{1 - x_i} \right]^{k-1} \quad (5.11)$$

where $k-1 = 2$ for our two-dimensional simplex. We can summarize the Metropolis-Hastings algorithm in the SALT MCMC as:

1. Calculate a new proposal for one of the categories, randomly chosen (Eq. 5.5)
2. Obtain the acceptance probability with Eqs. 5.10 and 5.11
3. If accepted, compute the proposals of the remaining categories with Eqs. 5.8 and 5.9. If rejected, a new category is chosen randomly to obtain its proposal in the next iteration.

5.3. MCMC Validation

Dorn et al. (2017) study the case of CoRoT-7 b with a non-adaptive MCMC. I test the non-adaptive MCMC (see Sect. 5.1) for CoRoT-7 b to compare with the analysis of Dorn et al. (2017), by using their mass and radius data (see Table 5.1).

M (data) [M_\oplus]	4.386 ± 0.985
R (data) [R_\oplus]	1.614 ± 0.102
M [M_\oplus]	4.328 ± 0.572
R [R_\oplus]	1.632 ± 0.076
x_{core}	0.23 ± 0.18
x_{H2O}	0.14 ± 0.12

Table 5.1.: CoRoT-7 b MCMC input (data), and output mean values and 1σ uncertainties.

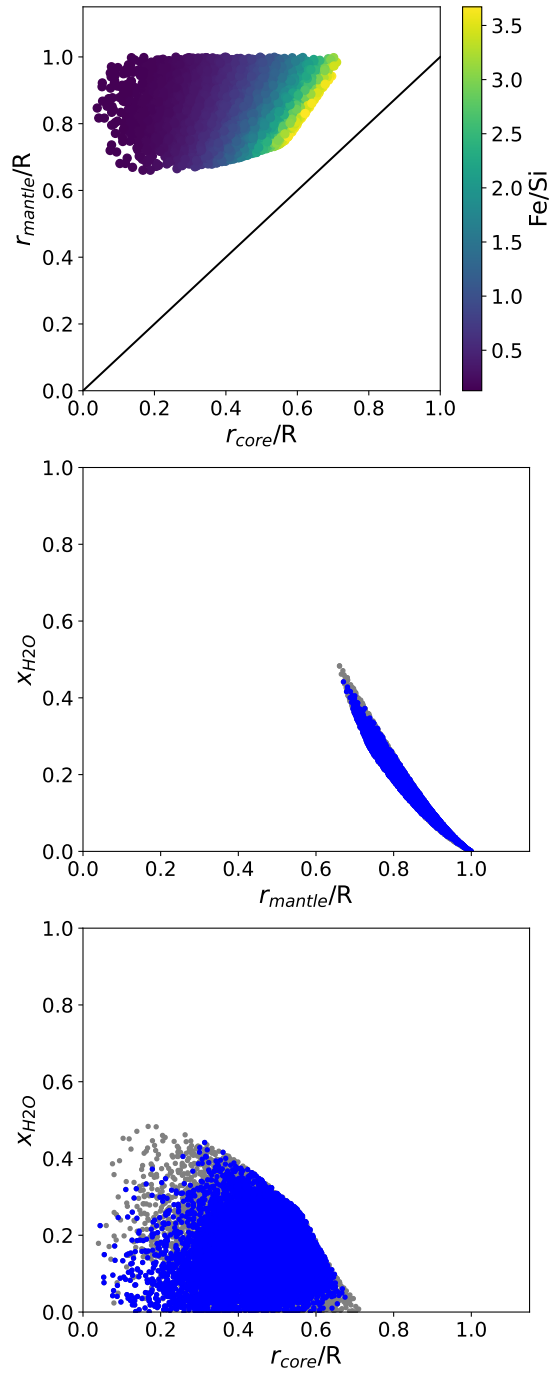


Figure 5.1.: Sampled 2D PDFs for CoRoT-7 b showing correlations between core and mantle size, and WMF. Upper panel: black solid line indicated the lower limit of the mantle size, where $r_{mantle} = r_{core}$. Middle and lower panels: blue points indicate simulations whose mass and radius are within their respective 1σ uncertainties.

The equilibrium temperature of CoRoT-7 b is 1756 K, which is too warm to hold liquid water in its surface. Nonetheless, to simplify this comparison between our MCMC implementation and that of [Dorn et al. \(2017\)](#), I assume that the hydrosphere is in condensed form in the interior model (see Sect. 2.1). I adopt Earth’s surface temperature and pressure values, since [Dorn et al. \(2017\)](#) include a liquid water layer with a gas layer on top. We consider a maximum limit for the CME, $x_{core} < 0.65$,

based on the maximum CMF of Mercury. The maximum limit on the WMF is 50%, which is the maximum value estimated by interior modelling of the moons in the Solar System (i.e Titan).

Fig. 5.1 shows the correlation between the mantle radius, r_{mantle} ; the core radius, r_{core} ; and the WMF, x_{H_2O} , of our sampled PDFs. Both analyses agree in that the size of the core is below $0.7 R$ (see figure 4 in Dorn et al. 2017), and that r_{mantle} is constrained in the interval $0.6-1 R$. Dorn et al. (2017) consider an extra gas layer in their interior model that we do not include in this validation test. If we compare figures 5 and 6 in Dorn et al. (2017) with Fig. 5.1, we can see that the WMF in our case extends to higher values. This is because our water layers are accounting for all the volatile content, whereas in Dorn et al. (2017) the volatile content is distributed between the water and the gas layer. We conclude that with our implementation of the algorithm described in Sect. 5.1, we have been able to reproduce the results of the core and mantle layers in Dorn et al. (2017).

Once the non-adaptive MCMC is validated, I compare the non-adaptive (Sect. 5.1) and the adaptive (Sect. 5.2) versions of the MCMC for TOI-220 b (Hoyer et al. 2021). I consider as data the total mass and radius, as well as the Fe/Si mole ratio, which has been calculated with the stellar abundances of the host star. No maximum limits have been established for the CMF or the WMF. TOI-220 b has an equilibrium temperature of 806 K, which means that it is strongly irradiated and could present steam and supercritical phases. Thus, I apply the interior-atmosphere model described in Chapters 3 and 4, implemented within both versions of the MCMC (adaptive and non-adaptive).

Table 5.2 shows a comparison of the input data, and the retrieved parameters of the non-adaptive and adaptive MCMCs. All three agree within uncertainties for mass, radius and Fe/Si. The uncertainties of the mass and radius in the non-adaptive MCMC are smaller than the input data. This difference in uncertainties is significant in the case of the total mass, and it is also observed in the case of Corot-7 b (Table 5.1). This indicates that the non-adaptive MCMC is not as effective as the adaptive MCMC at sampling all possible $\{x_{core}, x_{H_2O}\}$ pairs that could reproduce the mass and radius data. As a consequence, the uncertainties of the WMF are being underestimated in the non-adaptive MCMC, while the adaptive MCMC produces a greater confidence interval for the WMF, and retrieves the exact uncertainties of the mass and radius.

	Data	Non-adaptive	Adaptive
$M [M_{\oplus}]$	13.8 ± 1.0	13.8 ± 0.7	13.7 ± 1.0
$R [R_{\oplus}]$	3.03 ± 0.15	3.06 ± 0.12	2.98 ± 0.15
Fe/Si	0.65 ± 0.09	0.64 ± 0.11	0.64 ± 0.10
x_{core}		0.08 ± 0.03	0.09 ± 0.03
x_{H_2O}		0.62 ± 0.10	0.58 ± 0.14

Table 5.2.: TOI-220 b MCMC input (Data), and output mean values and 1σ uncertainties for the non-adaptive and adaptive MCMCs.

In Fig. 5.2, I show the sampled 2D PDFs for the CMF and the WMF in the ternary diagram. In addition to the same area of the ternary diagram as the non-adaptive algorithm, the adaptive MCMC explores an area at lower WMF along the Fe/Si = 0.65 isoline, going down to WMF = 0.10 in the driest simulations. This is a consequence of the ability of the adaptive MCMC to sample better the extremes of the prior distribution of the WMF, in comparison to the non-adaptive MCMC. Furthermore, the acceptance rate is also improved in the adaptive case, being able to accept 2110 simulations (53% acceptance rate) in comparison to 1400 simulations (35% acceptance rate) of the non-adaptive case within the same time. For a fixed number of accepted simulations $n_{accept} \simeq 8000$,

the total time is reduced from 7 days in the non-adaptive MCMC, to 4-5 days in the adaptive MCMC.

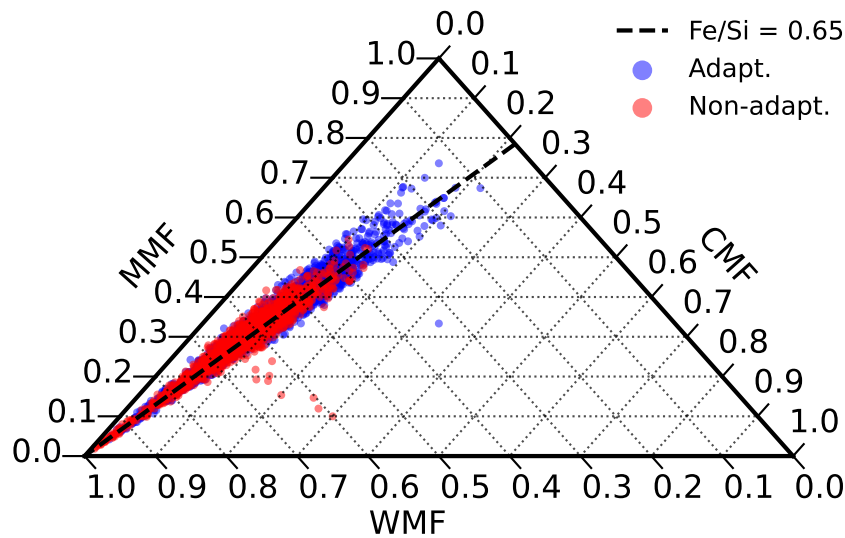


Figure 5.2.: Sampled 2D PDFs of the CMF and WMF in the ternary diagram for TOI-220 b for non-adaptive (red) and adaptive (blue) MCMCs. The mean value of the input Fe/Si mole ratio is indicated in a dashed black line.

5.4. Optimization and Python interface

The adaptive step size for MCMC methods for concentrations was written by [Director et al. \(2017\)](#) in the SALT Sampler package in R, which is a programming language specific for statistics. The interior and atmosphere models are written in Fortran 99. A first approach to put together the adaptive MCMC with the interior and atmosphere models could be translating the SALT Sampler package from R into Fortran 99, since Fortran has a few statistics libraries. However, these are not as well-documented and extensive as the statistical libraries already developed in Python. In fact, other forward models, such as petitRADTRANS, have their core code (input/output, radiative transfer calculations) in Fortran. Then, an interface built between Fortran and Python calls the Fortran subroutines to obtain the output and handle the plotting and statistics in Python. An example of this is the implementation of atmospheric retrieval of petitRADTRANS with emcee ([Foreman-Mackey et al. 2013](#)) and pyMultiNest ([Buchner et al. 2014](#)), which are MCMC samplers written in Python. Therefore, I build an interface for the interior and atmosphere models between Fortran and Python, so I can call the interior and atmosphere models from Python. Later on, I translate the original SALT Sampler package from R to Python.

The interface between Fortran and Python is built with a Python package named *f2py*² (Fortran to Python interface generator). Before building the interface with this Python package, the code in Fortran needed to be adapted so it would have the format of a subroutine, such as:

```
SUBROUTINE MSEI_interior(input_variable1, input_variable2, ...,
output_variable1, output_variable2,...)
```

²<https://numpy.org/doc/stable/f2py/>

This was the format of the atmospheric model initially, although this was not the case for the interior model. In addition, I optimized the interior model by separating the code that reads the input files and loads the data of the materials of each layer from the code that solves the differential equations presented in Sect. 2.1. This allows us to load the data just one, at the beginning of the interior model run, instead of each time the interior profiles are integrated, as it was done in the previous code implementation. Furthermore, *f2py* gives the possibility to generate the signature files with an optimization flag, such as `--opt='-03'`. The signature files are the files that contain the wrapper functions for the Fortran subroutines to Python. After this optimization, one complete run for one single exoplanet in the forward interior model was reduced from 1 min 45 seconds, to 30 seconds. This means that the computing time was reduced to 30% of its original value. Now both the interior and atmosphere models can be called from Python in a user-friendly way, where the model is a module, and its functions are contained in a class. In the following example, the module MSEI is the package that contains the class `int_planet`. The function `calc_radius` in this class calculates the radius of the planet from its center up to the upper interface of the supercritical layer:

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
Created on Mon Aug 29 13:10:09 2022

@author: lacuna
"""

# Import module
import MSEI

# Class
myplanet = MSEI.int_planet()

# Function to load layer material data
myplanet.setup_parameters()

# Input parameters
M_Pt = 1.6628      # Mass in Earth mass units
x_coret = 0.0     # Core mass fraction
x_H2Ot = 0.7      # Water mass fraction
T_surft = 4183    # Surface temperature in K
P_surft = 3e7     # Surface pressure in Pa

# Function that calculates the radius
myplanet.calc_radius(M_Pt,x_coret,x_H2Ot,T_surft,P_surft)

# Print output
print('Radius [R_E] = ', myplanet.R_P)
print('Fe/Si = ', myplanet.FeSi)
print('Density [g/m-3] = ', myplanet.rho_p)
```

6. Analysis of planetary systems

This chapter shows the results I obtained on different planetary systems I analysed with MSEI. These planetary systems host a wide variety of exoplanets, ranging from rocky Earth-sized planets, to sub-Neptunes.

In Sect. 6.1, I explore the interior structure and composition of TRAPPIST-1. This planetary system is well-known for hosting seven Earth-sized planets around a cool M-dwarf star (Gillon et al. 2016, 2017). Their equilibrium temperatures range from 400 K to 170 K, making this system ideal to apply interior models with different volatile phases, from ice and liquid water to supercritical. The content of this section led to my first-author publication, Acuña et al. (2021), which is attached to this manuscript in Appendix A.1.

I continue the application of our model to other multiplanetary systems (Sect. 6.2), to perform a homogeneous analysis of a sample of planets that formed in similar environments. I select systems that host only low-mass planets ($M < 20 M_{\oplus}$), with a 5 or more planets whose masses and radii are available. The results displayed in this section led to my second first-author publication, Acuña et al. (2022), which can be found in Appendix A.2.

In Sect. 6.3, I apply the model to several planetary systems that present less than three planets. The results exposed in this section have been gathered from eight publications in which I am co-author. These are enumerated in Appendix B.

In Sect. 6.4 I show how the interior-atmosphere model can be used to assess the observability of the atmospheres of rocky planets with the James Webb Space Telescope (JWST). I explore the particular cases of TRAPPIST-1 c and 55 Cancri e, which have been proposed for observations in emission spectroscopy. Finally, in Sect. 6.5, I discuss the results of our complete sample of planetary systems in the context of the low-mass planet population.

6.1. TRAPPIST-1

For the analysis of TRAPPIST-1, I adopt the observed mass, radius and semi-major axis values obtained by Agol et al. (2021) from transit timing variations (TTVs). I present two compositional scenarios: in scenario 1, I only consider as data the mass and the radius, whereas in scenario 2, I assume that the planet has a Fe/Si mole ratio equal to the stellar value.

TRAPPIST-1 is a cool M-dwarf, which poses a challenge when deriving its chemical stellar abundances from stellar spectroscopy due to its low brightness. Therefore, I estimate the Fe/Si ratio of TRAPPIST-1 from a sample of stars with similar metallicity and age, as proposed by Unterborn et al. (2018). We obtain the sample of stars from the Hypatia Catalogue (Hinkel et al. 2014, 2016, 2017), and select them based on $C/O < 0.8$, and $-0.04 \leq [Fe/H] \leq 0.12$, which is the 1σ confidence interval obtained by Gillon et al. (2017) for TRAPPIST-1. Fig. 6.1 shows the metallicity and Fe/Si of the thin disk star sample with their respective probability density functions. The Fe/Si mole ratio presents a mean of 0.76, with a standard deviation of 0.12, which are the values we use as data in scenario 2 of our analysis.

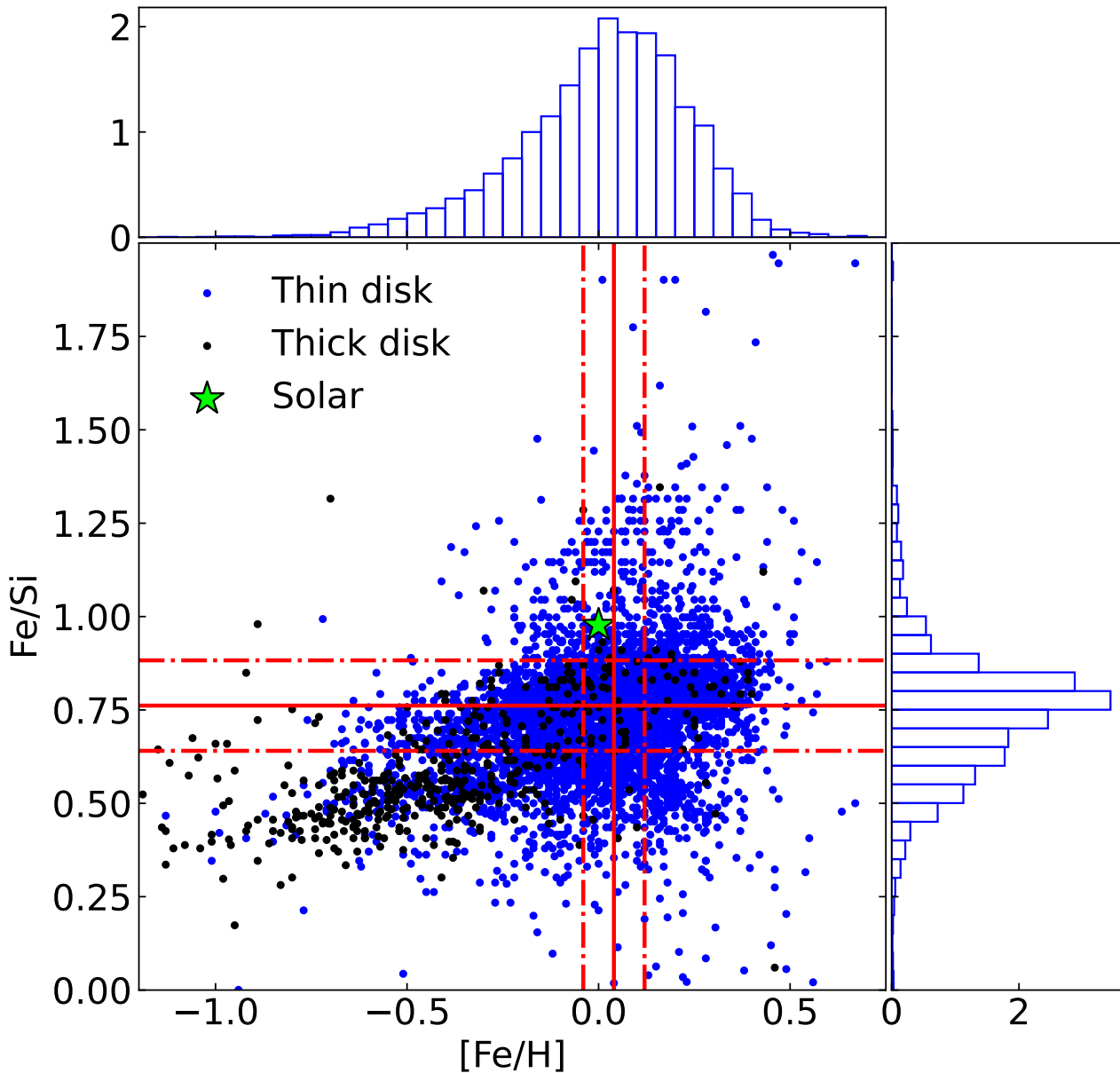


Figure 6.1.: Fe/Si and metallicity scatter plot and probability density functions of stars similar to TRAPPIST-1 (see text) from the Hypatia Catalogue. Red lines indicate the mean and standard deviation limits of the thin disk sample used to estimate the Fe/Si mole ratio of TRAPPIST-1.

6.1.1. Results

I apply our interior-atmosphere model (Chapters 2 and 3) to the most irradiated planets in TRAPPIST-1, which are planets b, c and d. In Fig. 6.2, we can see that for planets b and c the atmospheric model can find a surface temperature at which the OLR equals the absorbed flux, meaning that the atmosphere is in radiative-convective equilibrium. These surface temperatures are approximately 2450 K and 2250 K for TRAPPIST-1 b and c, respectively. On the contrary, the absorbed radiation does not reach the OLR for any surface temperature for the atmosphere of TRAPPIST-1 c. This suggests that planet d would be decreasing its global temperature by emitting radiation to space. In addition, an internal flux of 33 W/m^2 would be necessary to establish radiative equilibrium. [Barr et al. \(2018\)](#)

estimate an internal heat flux for TRAPPIST-1 c induced by tidal heating of 0.16 W/m^2 , which is significantly lower than what is needed to bring the atmosphere to radiative equilibrium. Thus, under the assumption of a water-dominated atmosphere with 1% CO_2 , the surface of TRAPPIST-1 d could be cold enough to maintain condensed phases.

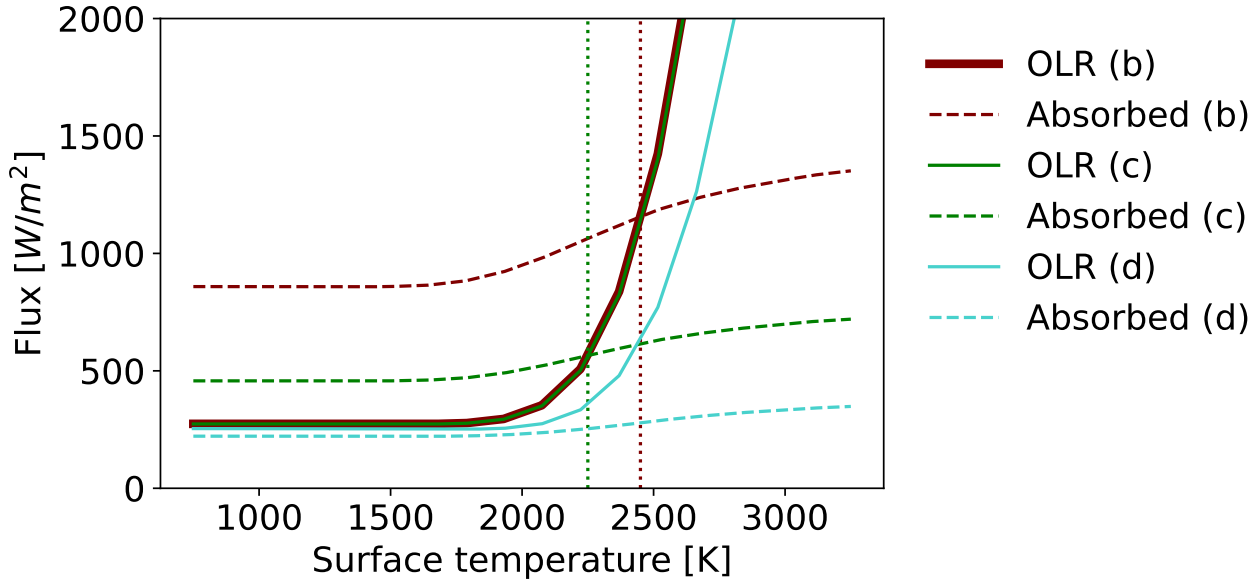


Figure 6.2.: Absorbed and emitted flux as a function of surface temperature for TRAPPIST-1 b, c and d. Vertical dotted lines specify the bottom atmospheric temperature at which the atmospheres of planet b and c are in radiative-convective equilibrium.

For planets d to h, which are temperate enough to have condensed phases of water, we consider their surface temperature for our interior model equal to their equilibrium temperatures at a null Bond albedo. The thickness of the mantle and condensed water layers has little dependence on the surface temperature. An Earth mass planet whose surface temperature is changed from 100 K to 360 K experiences a variation in total radius of $0.002 R_{\oplus}$ in our interior model. For the temperate planets in TRAPPIST-1, the contribution of the atmosphere to the total radius is expected to be negligible, with altitudes of 80 km (Lincowski et al. 2018). Therefore, we use the versions of the interior model with condensed phases for these planets. For TRAPPIST-1 d, which would have liquid surface conditions, we use the interior model described in Sect. 2.1.

For planets e to h, we use the version of the interior model described in section 2.2 of Acuña et al. (2021), and in Levesque (2019), which includes ice phases Ih, II, III, IV, V, VI and VII.

I estimate CMF and WMF for all planets in TRAPPIST-1 under the two scenarios (Tables 6.1 and 6.2). We obtain masses and radii that agree within uncertainties with the observed values in both scenarios, while in scenario 2 we also retrieve the Fe/Si we considered as input. In scenario 1, we are able to constrain a common Fe/Si mole ratio for the planetary system by considering the overlap of the 1σ confidence intervals of the individual planets. This interval is $\text{Fe/Si} = 0.45 - 0.97$, which is compatible with the interval estimated from the stellar sample (Unterborn et al. 2018) that we use in scenario 2. To visualize this overlap, we display in Fig. 6.3 the 1σ -confidence areas obtained from the 2D posterior distributions of the CMF and WMF for scenario 1 (top panel). The common CMF has a minimum value that is delimited by the lower limit of the confidence region of TRAPPIST-1 g. This value is 0.23, while the maximum common CMF is 0.4. The range $\text{CMF} = 0.23 - 0.4$ is in

agreement with the CMFs found in scenario 2 (Fig. 6.3, bottom panel), with values ranging from 0.2 to 0.3. These ranges of CMFs also comprise Earth’s value, 0.32.

Planet	Data		Scenario 1				
	M [M_{\oplus}]	R [R_{\oplus}]	M [M_{\oplus}]	R [R_{\oplus}]	CMF	WMF	Fe/Si
b	1.374±0.069	1.116 ^{+0.014} _{-0.012}	1.375±0.041	1.116±0.013	0.261±0.146	(3.1 ^{+5.0} _{-3.1}) × 10 ⁻⁵	1.00±0.56
c	1.308±0.056	1.097 ^{+0.014} _{-0.012}	1.300±0.036	1.103±0.015	0.239±0.084	(0.0 ^{+4.4} _{-0.0}) × 10 ⁻⁶	0.71±0.26
d	0.388±0.012	0.788 ^{+0.011} _{-0.010}	0.388±0.007	0.790±0.010	0.409±0.167	0.084±0.071	1.22 ^{+1.30} _{-1.22}
e	0.692±0.022	0.920 ^{+0.013} _{-0.012}	0.699±0.013	0.922±0.015	0.447±0.123	0.094±0.067	1.75±1.17
f	1.039±0.031	1.045 ^{+0.013} _{-0.012}	1.043±0.019	1.047±0.015	0.409±0.140	0.105±0.073	1.44±1.14
g	1.321±0.038	1.129 ^{+0.015} _{-0.013}	1.327±0.024	1.130±0.016	0.399±0.144	0.119±0.080	1.33±1.29
h	0.326±0.020	0.755±0.014	0.327±0.012	0.758±0.013	0.341±0.192	0.081 ^{+0.089} _{-0.081}	0.13 ^{+1.80} _{-0.13}

Table 6.1.: Masses, radii and compositional parameters obtained by the MCMC analysis for TRAPPIST-1 (scenario 1). For comparison, the second and third columns show the observed mass and radius of each planet (Agol et al. 2021), which are the input for the MCMC.

Planet	Data		Scenario 2				
	M [M_{\oplus}]	R [R_{\oplus}]	M [M_{\oplus}]	R [R_{\oplus}]	CMF	WMF	Fe/Si
b	1.374±0.069	1.116 ^{+0.014} _{-0.012}	1.359±0.043	1.124±0.016	0.259±0.032	(0.0 ^{+3.4} _{-0.0}) × 10 ⁻⁶	0.79±0.10
c	1.308±0.056	1.097 ^{+0.014} _{-0.012}	1.299±0.034	1.103±0.014	0.257±0.031	(0.0 ^{+2.7} _{-0.0}) × 10 ⁻⁶	0.79±0.11
d	0.388±0.012	0.788 ^{+0.011} _{-0.010}	0.387±0.007	0.792±0.010	0.241±0.032	0.036±0.028	0.76±0.12
e	0.692±0.022	0.920 ^{+0.013} _{-0.012}	0.695±0.012	0.926±0.012	0.249±0.031	0.024 ^{+0.031} _{-0.024}	0.78±0.12
f	1.039±0.031	1.045 ^{+0.013} _{-0.012}	1.041±0.019	1.048±0.013	0.240±0.031	0.037±0.026	0.76±0.12
g	1.321±0.038	1.129 ^{+0.015} _{-0.013}	1.331±0.023	1.131±0.015	0.235±0.031	0.047±0.028	0.75±0.12
h	0.326±0.020	0.755±0.014	0.326±0.011	0.758±0.013	0.232±0.032	0.055±0.037	0.75±0.12

Table 6.2.: Masses, radii and compositional parameters obtained by the MCMC analysis for TRAPPIST-1 (scenario 2). For comparison, the second and third columns show the observed mass and radius of each planet (Agol et al. 2021), which are the input for the MCMC in addition to the Fe/Si mole ratio estimated for TRAPPIST-1, Fe/Si = 0.76±0.12.

I compare our CMF estimates to those of Barr et al. (2018), who use an interior model with a Fe core, a silicate rock mantle and an ice and liquid water layer, while also computing the tidal heat flux with a thermal model. In addition, Barr et al. (2018) do not make use of an estimate of the stellar Fe/Si ratio to constrain the core mass fraction, and use mass estimates for TRAPPIST-1 from Wang et al. (2017), which are shifted to lower values compared to the mass estimates we use as input from Agol et al. (2021). Despite this, Barr et al. (2018) obtain CMF ranges compatible with our scenario 1. For planets b, d and e, Barr et al. (2018) obtain a maximum CMF of 0.40, which is approximately the maximum CMF we obtain in scenario 1 for planets b and c.

The WMFs we obtain for TRAPPIST-1 planets are well below 20% in the most general case (scenario 1). The maximum WMF is even lower for scenario 2, where planet h could have up to 9.2% of its mass in ices. We compare our WMF estimates to those of Agol et al. (2021), who assume different cases with constant CMF (18%, 25%, 32.5% and 50%) to calculate the WMF. Their interior structure model considers a steam atmosphere with N₂ as a background gas for planets b, c and d (Turbet et al. 2020), while planets e to h are modelled with condensed phases (Dorn et al. 2018). The WMF is strongly dependent on the assumed CMF, therefore we compare our scenario 2, where the CMF had values between 0.2 and 0.3, with the case CMF = 0.25 in Agol et al. (2021). For the inner planets of the TRAPPIST-1 system, we agree that they are compatible with a dry composition. Agol et al.

(2021) estimate a maximum WMF = 10^{-5} for planets b, c and d, while our maximum WMFs are 3.4×10^{-6} and 2.7×10^{-6} for planets b and c, respectively. In the case of planet d, we obtain a maximum WMF = 0.064, which is several orders of magnitude higher than the estimate in Agol et al. (2021). This is because the hydrosphere in our interior model is in condensed phases, whereas Agol et al. (2021) consider that TRAPPIST-1 d would have a steam atmosphere in a runaway greenhouse state. Planet d is at the very edge of the habitable zone, meaning that different atmospheric compositions could enable the presence of surface liquid water, whereas compositions with warmer surface temperatures would evaporate all surface water. Agol et al. (2021) assume a water-dominated atmosphere with N_2 as a background gas. In contrast, we adopt CO_2 as background gas, which affects radiative balance since it is a strong IR absorber in comparison to N_2 . For planets e to h, our WMF estimates are in agreement within uncertainties with Agol et al. (2021). Nonetheless, their mean values are lower, which could be due to the differences in EOS within the interior models. The EOS used by Agol et al. (2021) for condensed phases are in agreement with SESAME and ANEOS (Baraffe et al. 2008), which overestimate the density at pressures above 70 GPa (Mazevet et al. 2019). This causes the WMF to be underestimated for similar mass, radius and CMF.

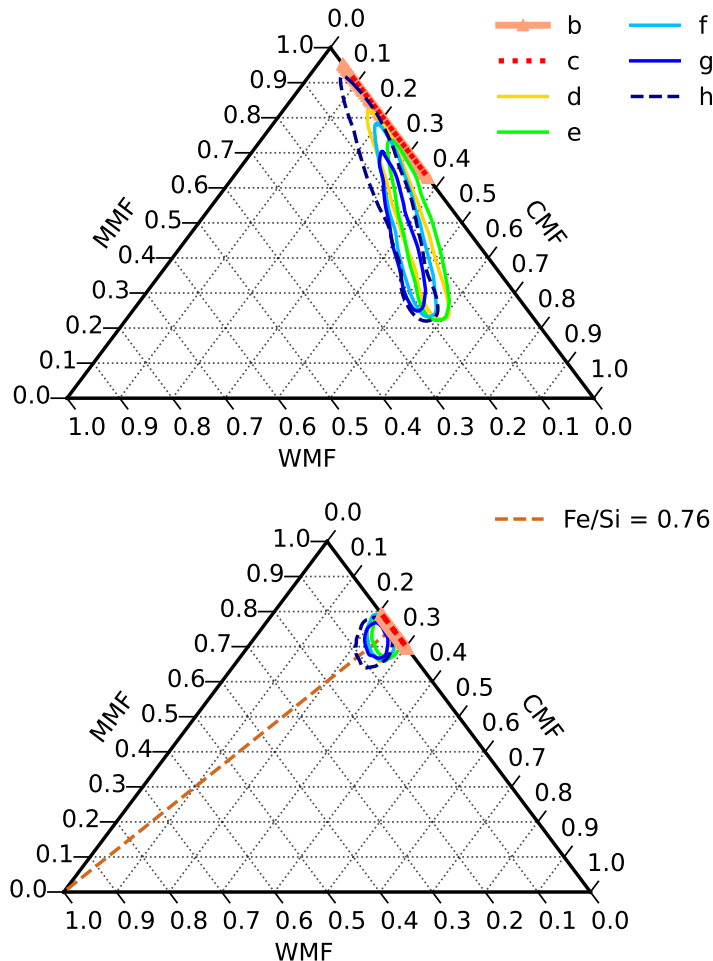


Figure 6.3.: Top panel: 2D 1σ -confidence intervals of the CMF and WMF for scenario 1, where mass and radius are the data. Bottom panel: 2D 1σ -confidence areas for scenario 2, where the Fe/Si abundance ratio estimated by Unterborn et al. (2018) is considered in addition to the masses and radii. The axes correspond to the CMF, the WMF, and the mantle mass fraction $MMF = 1 - CMF - WMF$.

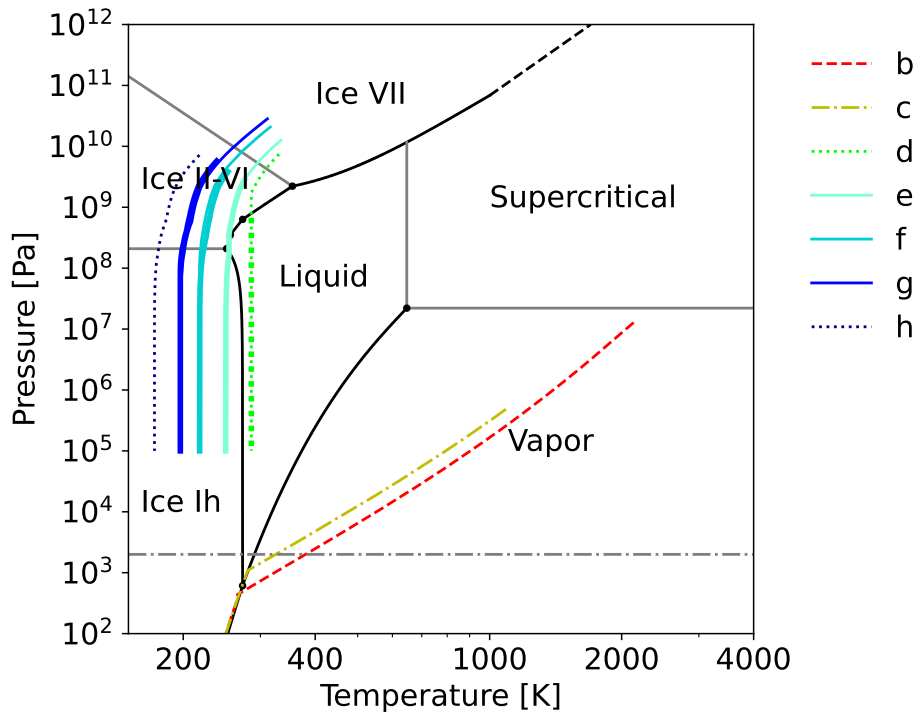


Figure 6.4.: Pressure-temperature adiabats of the water layers of TRAPPIST-1 planets for their maximum limits. For planets d to g, thicker lines indicate the PT profiles for their lower estimates, which are non-zero. Grey line marks the pressure level $P = 20$ mbar.

Fig. 6.4 shows the adiabatic profiles for the maximum and minimum WMF estimates. The atmospheres of planets b and c are not massive enough to reach the supercritical regime, having maximum surface pressures of 128.9 and 4.9 bar, respectively. We revisit the surface pressure of TRAPPIST-1 c with the adaptive MCMC in Sect. 6.4. The atmospheres of TRAPPIST-1 b and c are mostly comprised of a dry convective layer, although a small, wet convective area is present at 10 mbar. Therefore, clouds would be present above the 20 mbar pressure level, which is the clear transit radius (Mousis et al. 2020; Grimm et al. 2018). This could flatten the spectral features of water or CO_2 in transmission spectroscopy (Turbet et al. 2019; Katyal et al. 2020). Planets d to h present condensed phases, with the surfaces being liquid (planet d) or covered in ice Ih. Planets d to g have hydrospheres massive enough to reach 100 GPa at their base, meaning that they reach high-pressure ices (ice VII). Noack et al. (2016) discuss that having a high-pressure ice layer in the hydrosphere-mantle boundary is necessary to make a sub-surface ocean habitable in a water-rich planet, since the heat from the mantle would melt the base of the ice layer. However, for $\text{WMF} \geq 0.14$, the pressures at the bottom of the ice layer would be too high to form a liquid sub-surface ocean at high temperatures, rendering the hydrosphere uninhabitable. This is approximately the maximum value of our WMF within the confidence intervals. Nonetheless, the minimum values are 0 to 0.03, which are within the habitable sub-surface ocean range.

6.1.2. TRAPPIST-1 formation scenarios

Fig. 6.5 shows the WMF trend of both scenarios with semi-major axis. We observe that, in general, the inner planets are water-poor, while the outer planets seem more volatile-rich. This difference in composition between the inner and the outer parts of a system is also observed in the Galilean

moons: Io is likely a dry moon, followed by Europa with a WMF of 8%, and the volatile-rich Callisto and Ganymede (Ronnet et al. 2017). Regarding the planet formation mechanism that could produce the WMFs we retrieve, both pebble accretion and planetesimal-driven formation could be possible. The former can form planets with WMFs up to 5% within the ice line (Coleman et al. 2019; Liu et al. 2020), which is in agreement with the mean value of the outermost planet, TRAPPIST-1 h, in scenario 2. The maximum WMF estimated in the most general case (scenario 1) is 20%, which is far below the typical WMFs produced by planetesimal accretion according to Miguel et al. (2020). However, Kimura and Ikoma (2020, 2022) use a planet synthesis model that considers planetesimal-driven accretion, similar to Miguel et al. (2020), but include more processes that affect the final water content of the planet. These processes are the movement of the snow line due to the thermal evolution of the disk, the accumulation of primordial atmospheres and the production of water in the primordial atmosphere via oxidisation. When these mechanisms are considered, Kimura and Ikoma (2020) obtain that 25% of their simulated planets present thick oceans (WMF > 10%) which could be the case of the outer TRAPPIST-1 planets (e, f, g, h), while the other 75% are water poor, which could be representative of the inner planets in TRAPPIST-1 (b, c, d). In addition, they also find that planets interior to the snow line can accrete small amounts of water due to the sublimation of ice in its vicinity, in agreement with Mousis et al. (2020).

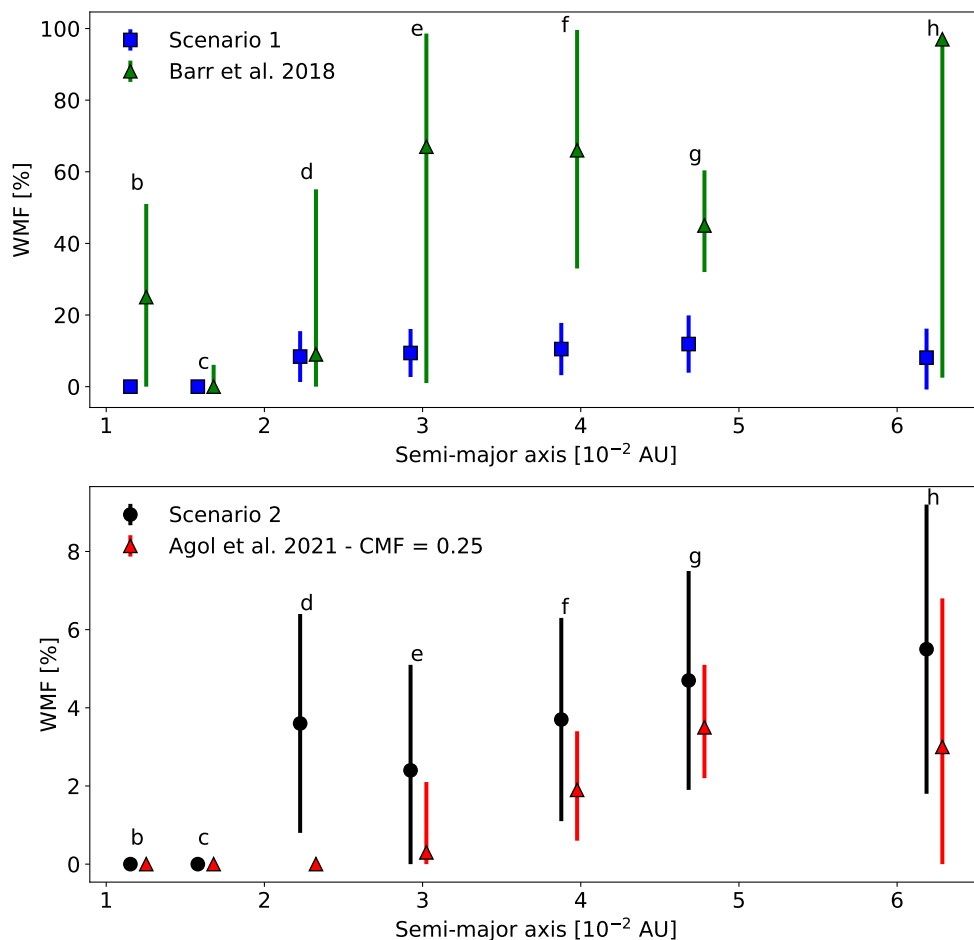


Figure 6.5.: Water mass fraction trend with semi-major axis for scenario 1 (top panel) and scenario 2 (bottom panel). Estimates from previous works are also shown for comparison.

Furthermore, in Fig. 6.5 we can recognise an increase of WMF as the distance from the host star

increases, especially in scenario 2, where the same Fe/Si mole ratio is considered for all planets in the system. TRAPPIST-1 d seems to present a deviation from this trend, since its mean WMF is slightly higher than that of planet e. This can be explained from the point of view of planet formation by a cold finger, pebble ablation or migration. A cold finger consists of a region in the protoplanetary disk where the gas is more enriched in volatiles (Stevenson and Lunine 1988; Cyr et al. 1998), meaning that if planet d formed in such an area, it would have accreted more ices than other planets outside its orbit. Similarly, the vicinity of the water ice line can be more enriched in ices thanks to the backward diffusion of vapour combined with the inward drift of icy particles (Mousis et al. 2019). Migration could also explain the high WMF of planet d (compared to planet e), if TRAPPIST-1 d formed beyond the snow line prior to its inward migration (Raymond et al. 2018). Finally, the accretion of volatiles, including ices, could have been less efficient for planet e than for planet d due to pebble ablation and ice recycling back into the disk (Coleman et al. 2019). In addition to these formation mechanisms, the increasing WMF trend with distance from the star could have been shaped by atmospheric escape. TRAPPIST-1 is a cool M-dwarf that has a strong emission in the XUV. If we take into account the water mass lost due to XUV photoevaporation (Bolmont et al. 2017), the initial WMF of planets b, c and d would be 2.37×10^{-3} , 2.50×10^{-3} , and 0.085, respectively. Thus, the individual atmospheric loss decreases the WMF with time for each planet, but the general trend we observe in Fig. 6.5 remains unaffected by atmospheric evaporation. Water may be difficult to probe in the outer planets of TRAPPIST-1 with atmospheric characterisation observations. Krissansen-Totton and Fortney (2022) use an interior-atmosphere model to simulate the thermal evolution of all TRAPPIST-1 planets, and they find that the outer planets (f and g) are likely to have anoxic, cold CO₂ atmospheres with a very low H₂O partial pressure due to the condensation of water on the surface. This explains the absence of spectral features in the transmission spectrum of TRAPPIST-1 h (Gressier et al. 2022). In our interior analysis, we did not include atmospheric modelling for the coldest planets of the system since their atmospheres have a negligible contribution to the total radius (similar to Earth, which is 50 km or less).

We have considered so far that the water layer in TRAPPIST-1 d is in condensed phases. Nonetheless, its volatile layer could be composed of other atmospheric species, which would produce a thick atmosphere. Atmospheric characterisation data has discarded atmospheres dominated by H/He and CH₄ (de Wit et al. 2016, 2018; Ducrot et al. 2020). CO₂ can be produced in small-sized exoplanets where tidal heating is a source of internal heat that favours plate tectonics and volcanic outgassing (Papaloizou et al. 2018; Ortenzi et al. 2020). Therefore, we compare the mass and radius of TRAPPIST-1 d with mass-radius relations for a planet with a CO₂-dominated atmosphere of $P_{surf} = 300$ bar. We find that TRAPPIST-1 d is compatible with such atmosphere and a CMF between 0.2 and 0.3, which is the range derived for the CMF in scenario 2 (Fig. 6.6).

Krissansen-Totton and Fortney (2022) adapt an interior-atmosphere evolution model for Venus to simulate the evolution of the atmospheres in the TRAPPIST-1 planets. They find that the most likely atmospheric composition of TRAPPIST-1 b and c (if there is an atmosphere) is CO₂-dominated, with a high concentration of O₂. This concentration depends on the efficiency of O sinks and the initial water inventory. In the case of Venus, the CO₂-rich atmosphere is anoxic. Both the wet, temperate scenario and the dry one after the runaway greenhouse phase of Venus leave behind a significant amount of oxygen. Therefore, in Venus, oxidation of the mantle is required to remove atmospheric oxygen and reduce oxygen levels to those we observe today. This sink could be provided by explosive volcanism (Gillmann et al. 2022). In the case of the TRAPPIST-1 planets, concentrations of at least 1 bar in oxygen could be detectable with JWST (Krissansen-Totton and Fortney 2022). If oxygen is low in their atmospheres, it would mean that a similar sink to that of Venus exists in exoplanets. The surface temperature in TRAPPIST-1 b and c is higher than in Venus, so dehydration reactions of the

mantle would be more efficient. However, in the inner planets of TRAPPIST-1 atmospheric escape is stronger than in Venus. So while in Venus we find a hydrated mantle with a water-poor atmosphere, TRAPPIST-1 b and c may present a dehydrated mantle with no water in the atmosphere.

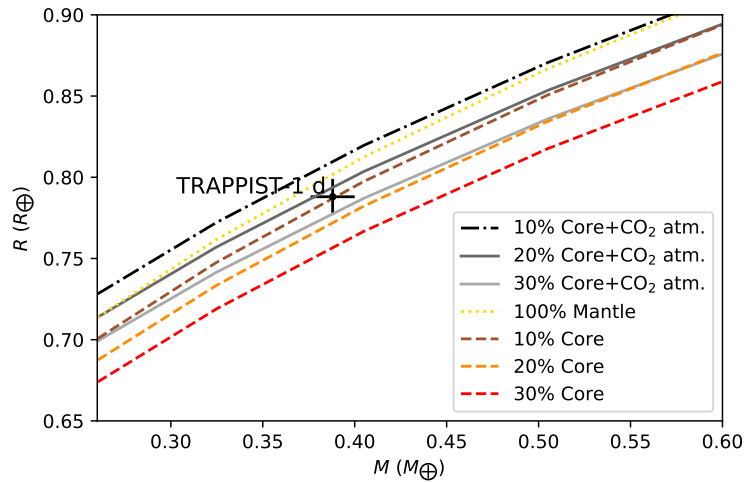


Figure 6.6.: Mass-radius relations for a planet with a CO_2 -dominated atmosphere with a surface temperature of 300 bar, and an equilibrium temperature equal to that of TRAPPIST-1 d. The mass-radius relations of a planet with no atmosphere with different CMF values are also shown for comparison.

6.2. Multiplanetary systems

In Sect. 6.1 we found that in the TRAPPIST-1 system, the WMF presents a trend with semi-major axis, which could be due to formation and evolution processes. To explore the compositional trends in other multiplanetary systems, we apply the interior-atmosphere model to a sample of planetary systems with five or more planets that have their mass and radius well-constrained. These are five systems whose input planetary parameters are shown in Table 6.3. All planets in these systems have irradiation temperatures greater than 500 K, which means that their surface water layer is in steam or supercritical phases. In addition, for this analysis we consider an upper limit for the WMF = 0.8. The application of the same interior structure model to a sample of multiplanetary systems constitutes an homogeneous derivation of their compositions. This minimizes the differences when comparing compositional parameters (Fe and volatile contents, core radius, etc) between models that take distinct assumptions on the interiors, such as the presence of a H/He gaseous layer (Lopez and Fortney 2014), a condensed water layer (Dorn et al. 2015; Zeng et al. 2019) or a steam atmosphere (Mousis et al. 2020). Therefore, we will be able to compare potential volatile and core mass fraction trends between different systems, and assess how their general composition constrains planet formation and evolution.

System	Planet	M [M_{\oplus}]	R [R_{\oplus}]	a_d [AU]	T_{irr} [K]	Fe/Si
K2-138 ¹	b	2.80 ^{+0.94} _{-0.96}	1.442 ^{+0.071} _{-0.063}	0.033	1291	0.77±0.07
	c	5.95 ^{+1.17} _{-1.12}	2.198 ^{+0.066} _{-0.054}	0.044	1125	
	d	7.20 ^{+1.39} _{-1.40}	2.310 ^{+0.077} _{-0.068}	0.058	978	
	e	11.28 ^{+2.78} _{-2.72}	3.276 ^{+0.095} _{-0.082}	0.077	850	
	f	2.43 ^{+3.05} _{-1.75}	2.787 ^{+0.093} _{-0.085}	0.103	735	
	g	2.45 ^{+2.92} _{-1.74}	2.911 ^{+0.305} _{-0.230}	0.227	494	
	TOI-178 ²	b	1.5 ^{+0.39} _{-0.44}	1.152 ^{+0.073} _{-0.070}	0.026	
c		4.77 ^{+0.55} _{-0.68}	1.669 ^{+0.114} _{-0.099}	0.037	873	
d		3.01 ^{+0.80} _{-1.03}	2.572 ^{+0.075} _{-0.078}	0.059	691	
e		3.86 ^{+1.25} _{-0.94}	2.207 ^{+0.088} _{-0.090}	0.078	600	
f		7.72 ^{+1.67} _{-1.52}	2.287 ^{+0.108} _{-0.110}	0.104	521	
g		3.94 ^{+1.31} _{-1.62}	2.87 ^{+0.14} _{-0.13}	0.128	471	
Kepler-11 ^{3,4}		b	4.3 ^{+2.2} _{-2.0}	1.97±0.19	0.091	953
	c	13.5 ^{+4.8} _{-6.1}	3.15±0.30	0.106	883	
	d	6.1 ^{+3.1} _{-1.7}	3.43±0.32	0.159	721	
	e	8.4 ^{+2.5} _{-1.9}	4.52±0.43	0.194	653	
	f	2.3 ^{+2.2} _{-1.2}	2.61±0.25	0.250	575	
	Kepler-102 ^{5,6}	b	0.41±1.6	0.47±0.02	0.055	868
c		-1.58±2.0	0.58±0.02	0.067	786	
d		3.80±1.8	1.18±0.04	0.086	597	
e		8.93±2.0	2.22±0.07	0.117	694	
f		0.62±3.3	0.88±0.03	0.165	501	
Kepler-80 ^{7,8}	d	5.95 ^{+0.65} _{-0.60}	1.309 ^{+0.036} _{-0.032}	0.033	990	0.97±0.08
	e	2.97 ^{+0.76} _{-0.65}	1.330 ^{+0.039} _{-0.038}	0.044	863	
	b	3.50 ^{+0.63} _{-0.57}	2.367 ^{+0.055} _{-0.052}	0.058	750	
	c	3.49 ^{+0.63} _{-0.57}	2.507 ^{+0.061} _{-0.058}	0.071	679	
	g	0.065 ^{+0.044} _{-0.038}	1.05 ^{+0.22} _{-0.24}	0.094	588	

Table 6.3.: Input data for our MCMC analysis: masses, radii, semi-major axis, irradiation temperature ($A_B = 0$), and Fe/Si mole ratio for the multiplanetary systems K2-138, TOI-178, Kepler-11, Kepler-102, and Kepler-80. **References:** (1) [Acuña et al. \(2022\)](#); (2) [Leleu et al. \(2021\)](#); (3) [Lissauer et al. \(2011\)](#); (4) [Brewer et al. \(2016\)](#) (5) [Marcy et al. \(2014\)](#); (6) [Brewer and Fischer \(2018\)](#); (7) [MacDonald et al. \(2016\)](#); (8) [MacDonald et al. \(2021\)](#)

I define the variable $d_{obs-ret} = \max\{|R_{data} - R|, |M_{data} - M|\}$, which quantifies the difference between the observed data and the retrieved value for the mass and radius of a planet. If the retrieved and the observed value have a difference greater than 1σ , the model does not agree within uncertainties with the observation, and we can conclude that the assumption of a hydrosphere on top of a rocky core is not compatible with the observations.

System	Planet	CMF	WMF	$d_{obs-ret}$	$\Delta M_{H_2} [M_{\oplus}]$	$\Delta M_{XUV} [M_{\oplus}]$
K2-138	b	0.27±0.02	0.000 ^{+0.007} _{-0.000}	1.5 σ	0.132	0.40
	c	0.23±0.02	0.13±0.04	<1 σ	< 0.01	< 0.01
	d	0.22±0.03	0.17±0.05	<1 σ	< 0.01	< 0.01
	e	0.11±0.02	0.57±0.08	<1 σ	< 0.01	< 0.01
	f	0.11±0.02	0.60±0.07	<1 σ	< 0.01	< 0.01
	g	0.12±0.05	0.55±0.18	1.3 σ	< 0.01	< 0.01
TOI-178	b	0.21±0.30	0	<1 σ	0.83	0.45
	c	0.30±0.02	0.02 ^{+0.04} _{-0.02}	<1 σ	< 0.01	0.21
	d	0.10±0.01	0.69±0.05	1.3 σ	0.16	0.48
	e	0.18±0.02	0.40±0.06	<1 σ	< 0.01	0.13
	f	0.22±0.03	0.28±0.10	<1 σ	< 0.01	0.04
	g	0.10±0.01	0.58±0.16	3.0 σ	< 0.01	0.11
Kepler-11	b	0.20±0.04	0.27±0.10	<1 σ	< 0.01	0.10
	c	0.18±0.01	0.33±0.04	1.7 σ	< 0.01	0.10
	d	0.10±0.02	0.65±0.05	2.4 σ	< 0.01	0.13
	e	0.12±0.01	0.55±0.04	4.4 σ	< 0.01	0.14
	f	0.14±0.06	0.47±0.10	1.9 σ	0.56	0.06
Kepler-102	b	0.91 ^{+0.09} _{-0.16}	0	<1 σ	0.13	0.03
	c	0.95 ^{+0.05} _{-0.30}	0	<1 σ	0.10	0.03
	d	0.80±0.14	0	<1 σ	< 0.01	0.03
	e	0.22±0.02	0.17±0.07	<1 σ	0.01	0.03
	f	0.27±0.09	0.04±0.04	<1 σ	0.02	0.01
Kepler-80	d	0.97 ^{+0.03} _{-0.05}	0	<1 σ	< 0.01	0.35
	e	0.43±0.18	0	<1 σ	< 0.01	0.29
	b	0.13±0.02	0.58±0.07	<1 σ	< 0.01	0.11
	c	0.09±0.01	0.70±0.04	<1 σ	< 0.01	0.13
	g	0.31±0.02	< 1.5 × 10 ⁻³	<1 σ	140	0.60

Table 6.4.: Retrieved compositional parameters for our sample of multiplanetary systems. $d_{obs-ret}$ quantifies how well the assumption of a water-dominated atmosphere fits the data for each planet (see text). ΔM_{H_2} and ΔM_{XUV} are the maximum atmospheric masses lost due to H₂ Jeans escape and XUV photoevaporation, respectively.

Table 6.4 displays the mean and 1 σ uncertainties of the CMF and WMF, as well as the parameter $d_{obs-ret}$ for all planets in the multiplanetary systems I analysed. We can see that, in some cases, $d_{obs-ret} > 1\sigma$ simultaneously to a high WMF (more than 30%), and a retrieved radius lower than the observed one. This indicates that despite exploring very volatile-rich compositions, the observed radius is greater than that of a planet with a water-dominated atmosphere, which suggests that the atmosphere is made of more volatile species, such as H and He. For planets whose density is higher than the density of a bare rocky planet with 100% mantle composition, we set the WMF equal to zero and let the CMF be the only free parameter, while the mass and radius are the data.

In addition to the interior composition, we also compute the mass lost in the atmosphere due to H₂ Jeans escape and XUV photoevaporation (Table 6.4). The former process occurs when the gravitational pull of the bulk of the planet is not enough to retain the atmosphere particles whose thermal energy is greater than the kinetic energy associated to the escape velocity (Jeans 1925). In contrast, if the star emits strong radiation in the X-ray and ultraviolet (UV) wavelengths, this

radiation provides extra momentum and ionizes the molecules, generating collisions between them and raising the temperature, which leads to atmospheric evaporation (Sanz-Forcada, J. et al. 2011). We estimate mass loss rates as described in Aguichine et al. (2021), where an energy-limited approach is taken. We also calculate mass loss of H_2O due to Jeans escape, where we obtain lost atmospheric masses less than $0.01 M_\oplus$ for all planets in the multiplanetary systems we analyse, except for Kepler-80 g, which presents $\Delta M_{H_2O} = 3.23 M_\oplus$. We assume that the XUV flux is constant during the saturation regime, followed by a power-law evolution in time. Then the mass loss rate (Owen and Jackson 2012; Aguichine et al. 2021) is integrated over time under the assumption that the total mass and radius are constant. We use as input the masses, radii and equilibrium temperatures in Table 6.3. The inclusion of these mass loss estimates in our analysis should help discern whether the composition of volatile-poor planets is due to their original ice content during planet formation, or due to post-formation atmospheric escape.

6.2.1. K2-138

In Fig. 6.7, we can observe that the 1σ confidence areas of all planets in the system are aligned along the Fe/Si iso-line that corresponds to the mean input Fe/Si value we derived from the host stellar abundances, which is Fe/Si = 0.77. This means that the densities of all planets in the system can be accounted for with a rocky core and mantle with a Fe/Si ratio similar to the stellar value with a volatile layer on top. For K2-138 b, an atmosphere of 300 bar or more reproduces a radius 1.5σ larger (see Table 6.4) than the observed radius. If we assume the mean observed mass of planet b, with a CMF = 0.27, being compatible with the CMF we obtain for planet b in Fig. 6.7, a steam atmosphere with $P_{surf} = 300$ bar produces a radius of $1.461 R_\oplus$. This radius is compatible within uncertainties with the observed radius. Since this atmospheric pressure corresponds to a WMF = 0.01%, planet b is likely to have a very light atmosphere ($P_{surf} < 300$ bar), or no atmosphere at all.

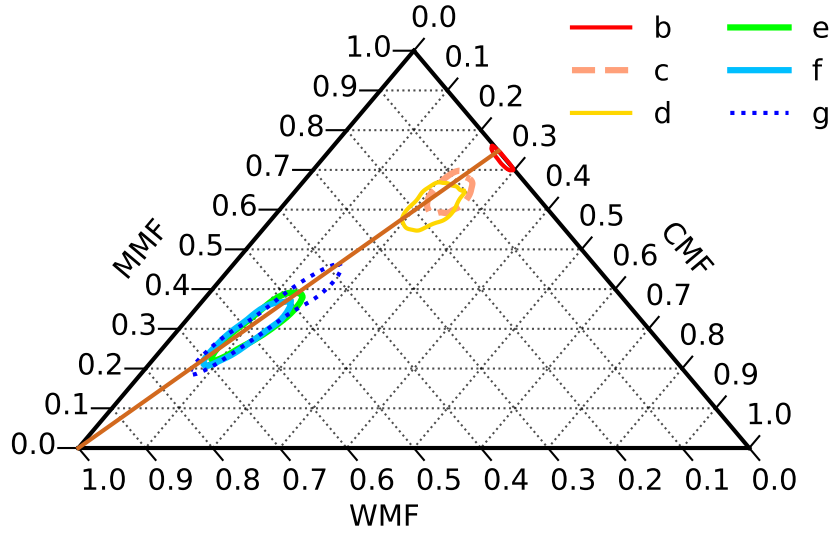


Figure 6.7.: 2D 1σ confidence regions of the compositional parameters in the ternary diagram for the multiplanetary system K2-138. The brown solid line indicates the iso-line where Fe/Si = 0.77, which is the mean value of the input Fe/Si mole ratio.

In contrast, planets c to g are very volatile-rich, with WMFs greater than 10%. Furthermore, K2-138 g presents a retrieved radius in our analysis that is 1.3σ greater than the observed value,

with a high WMF (up to 80%). This indicates that the interior models, in which we assume that the volatile layer is dominated by water, cannot reproduce the extended radius of planet g, suggesting that its atmosphere contains H and/or He as dominant species. The mass and radius of K2-138 g is compatible with a rocky core with a 0.1-5% H/He mass fraction (Lopez and Fortney 2014; Zeng et al. 2019).

Finally, of all planets in the K2-138 system, the planet that has had a significant atmospheric mass loss due to Jeans escape and photoevaporation is planet b. Most of the atmosphere could be lost during the saturation regime due to hydrodynamic escape, yielding an integrated mass loss that constitutes up to 14% of K2-138 b total mass. This relative lost atmospheric mass of 14% is similar to the WMF of planets c and d, which are approximately 10 to 20%. It is therefore possible that planet b formed with an ice mass fraction similar to the current WMF of planet c, and lost its massive H₂O envelope due to XUV photoevaporation.

6.2.2. TOI-178

TOI-178 presents a clear dichotomy between the inner and the outer planets in terms of their volatile mass fractions (see Table 6.4). The inner planets b and c are water-poor, with WMF < 6%, whereas the outer planets d to g have high WMF with a minimum value of 30%. The WMF values of the outer planets do not follow a clear increase with semi-major axis as in the case of K2-138 or TRAPPIST-1, which could be due to different abundances of H/He and water mixed in the volatile layers of each planet. Furthermore, the retrieved radii of TOI-178 d and g in our MCMC are significantly higher than their observed radii ($d_{obs-ret} > 1\sigma$). This, combined with the high WMFs (see Table 6.4), suggests that planets d and g have H/He-dominated atmospheres. According to our mass loss estimates, TOI-178 b could have undergone the loss of a maximum of $0.83 M_{\oplus}$ in H₂ from Jeans escape, while planet c could have evaporated $0.21 M_{\oplus}$. Under these maximum atmospheric loss estimates and the assumption of an hydrosphere, the initial volatile mass fractions of TOI-178 b and c are 0.36 and 0.10, respectively. This means that the inner, dry planets in TOI-178 could have been water-rich at the time of their formation.

6.2.3. Kepler-11

For Kepler-11 b, the retrieved mass and radius are in agreement within uncertainties with the observed values, which means the hypothesis of a water-dominated atmosphere is compatible with the data, with a maximum WMF of $\simeq 40\%$. For planet c to e, their retrieved radii are significantly lower than the observed radius, with differences of 1.7σ , 2.4σ and 4.4σ , respectively. These, together with high WMF values, indicate that the outer planets of Kepler-11 probably have H/He envelopes. They might present increasing volatile mass fractions with semi-major axis since their radii differ more from that of a planet with a hydrosphere as we move further away from the star. Nonetheless, the outermost planet has a $d_{obs-ret} = 1.9\sigma$, which is smaller than planet d's value. This could be due to the low mass of Kepler-11 f, which could have not retained a substantial part of its atmospheric mass ($0.56 M_{\oplus}$) due to H₂ Jeans escape.

6.2.4. Kepler-102

The three inner planets of Kepler-102 are positioned below the 100% mantle line in the mass-radius diagram, which points to dry, rocky compositions. Thus, we assume a WMF = 0 in the MCMC Bayesian analysis of Kepler-102 b, c, and d, and let the CMF as the free parameter, while the mass

and radius are the observable data. We do not set the Fe/Si mole ratio as a constraint, since these planets could have CMFs higher than would be expected from the Fe and Si content of the star due to their high densities. We obtain $\text{CMF} = 0.91_{-0.16}^{+0.09}$, $0.95_{-0.30}^{+0.05}$, and 0.80 ± 0.14 for Kepler-102 b, c and d, respectively. These high CMFs indicate that the inner planets of this system have interiors similar to that of Mercury in the Solar System. During planet formation, the planetary embryos could have accreted pebbles that are rich in Fe, which are found in the inner part of the protoplanetary disk, close to the rock-lines (Aguichine et al. 2020; Scora et al. 2020). Alternatively, formation beyond these Fe-rich regions of the disk would produce planets with average CMFs (0.2-0.4), but post-formation processes could destroy part of the mantle, increasing the core-to-mantle ratio and the CMF. These mechanisms include mantle evaporation (Cameron 1985) and impacts (Benz et al. 1988; Asphaug and Reufer 2014).

In contrast to the Mercury-like, dry inner planet, Kepler-102 e is compatible with a water-rich planet, with a WMF $\approx 20\%$. The outermost planet has a WMF that ranges from zero to WMF = 0.08. In addition, $0.02 M_{\oplus}$ could have been lost in H_2 due to Jeans atmospheric escape in Kepler-102 f, which would yield an initial WMF = 0.07 in the case that planet f is dry currently. Therefore, Kepler-102 f could have formed in an ice-abundant region of the protoplanetary disk similar to Kepler-102 e, making Kepler-102 show a dichotomy between inner, dry planets and outer, volatile-rich planets that other multiplanetary systems in our sample present.

6.2.5. Kepler-80

Similarly to Kepler-102 b, c and d, Kepler-80 d is a Mercury-like planet with a CMF = 0.97. Kepler-80 e is also a dry, rocky planet, but with a lower CMF, being compatible with an Earth-like interior (CMF = 0.32). The outer planets, which are Kepler-80 b, c and g, have non-zero WMF. The radius of Kepler-80 g is 2.7σ higher than the radius of a bare, 100% mantle planet with similar mass. This means that planet g presents a low-mass atmosphere ($P_{surf} < 300$ bar), since the total planetary mass of $M = 0.065 M_{\oplus}$ is too low to retain a more massive atmosphere. Furthermore, based on our Jeans and XUV mass loss estimates, both atmospheric escape mechanisms could have removed completely a H/He envelope. Kepler-80 g is also the only planet in our sample that could have experienced significant Jeans escape of H_2O , with $\Delta M_{\text{H}_2\text{O}} = 3.26 \times 10^{-3}$ to $3.24 M_{\oplus}$. This total atmospheric loss in addition to the current one adds up to an original WMF similar to that of planet b.

6.2.6. Compositional trends

In all the multiplanetary systems I have analysed, with the exception of Kepler-11 where all planets are volatile-rich, there is a clear division between at least one inner, dry planet, and the outer planets that present volatiles. Moreover, in some of the systems, it is possible to unveil a trend of the WMF with semi-major axis. To see this more clearly, we display in Fig. 6.8 the WMF trend with absolute incident stellar irradiation (top panel), and with the irradiation normalised to the flux of the innermost planet in each system (bottom panel). K2-138 presents a very clear volatile trend: the inner planets have an increasing WMF as we move further away from the star, while the outer planets (e to g) present an approximately constant volatile content. This gradient-plus-plateau trend can also be hinted in TRAPPIST-1, with the exception of planet d. The slight deviation of TRAPPIST-1 d from this trend could be due to the assumption of a condensed water layer instead of a CO_2 -dominated atmosphere, as discussed in Sect. 6.1.2. It would be necessary to use atmospheric characterization techniques to confirm that the volatiles of TRAPPIST-1 d are in gaseous phase, as

well as the composition of the envelope. The system Kepler-102 does not show the gradient part of the trend in the inner planets, since the three of them have a volatile mass fraction of zero, but it could potentially present the WMF plateau within uncertainties if Kepler-102 e and f have WMF \approx 8-10%. Similarly, Kepler-80 shows dry inner planets in addition to outer volatile-rich planet with a comparable WMF, with the exception of Kepler-80 g, which might have lost most of its envelope due to photoevaporation or Jeans escape. Finally, the volatile mass fraction trend is difficult to discern given that some of the planets in these two systems are likely to have H/He envelopes, and thus they should be analysed with a self-consistent interior model that include H/He as a gaseous species. For the rest of the planets, there exists a degeneracy between light H/He atmospheres and envelopes with a higher molecular weight, such as H₂O, as we consider in our analysis. If we were to re-analyse these planets assuming a H/He envelope instead of a water-dominated one, the volatile mass fraction of each individual planet would decrease, but the overall trend within each system would be preserved. To break this degeneracy between atmospheric composition and atmospheric mass, it is necessary to acquire atmospheric characterisation data. Moreover, transmission spectroscopy in sub-Neptunes has shown that their atmospheres are not pure H/He or water, but they are combination of both (Tsiaras et al. 2019; Benneke et al. 2019; Madhusudhan et al. 2020). This is further supported by meteorite outgassing experiments (Thompson et al. 2021) and atmospheric disequilibrium chemistry models (Moses et al. 2013; Guzmán-Mesa et al. 2022).

The gradient-plus-plateau trend is likely to be yielded by a combination of accretion of ice-rich pebbles, atmospheric loss and migration in multiplanetary systems. Ice-rich pebbles could be accreted in the vicinity of the water ice line (Mousis et al. 2019, 2021), followed by inward migration, where the planet loses part or all of its envelope due to XUV evaporation as it gets closer to the star, or Jeans escape if its bulk is not enough to retain the atmosphere. For K2-138 and TOI-178, their Laplace resonances are indicative of inner planetary migration (Terquem and Papaloizou 2007; Izidoro et al. 2017; Ramos et al. 2017). For systems whose typical water mass fractions are \approx 10%, such as TRAPPIST-1 and Kepler-102, water could be simply obtained from chondritic material within the snow line, without the enhancement produced by the radial drift of ice-rich planetesimals. Furthermore, 10% in volatile content is the typical water mass fraction of the asteroids in the Main Belt (Vernazza et al. 2015). This suggest that these TRAPPIST-1 and Kepler-102 could have formed similarly to the Main Belt, where silicate material and ice were accreted within the snow line (Rivkin et al. 2002). Therefore, the most volatile-rich systems (K2-138, TOI-178 and Kepler-11) could have had longer migration distances than TRAPPIST-1 and Kepler-102.

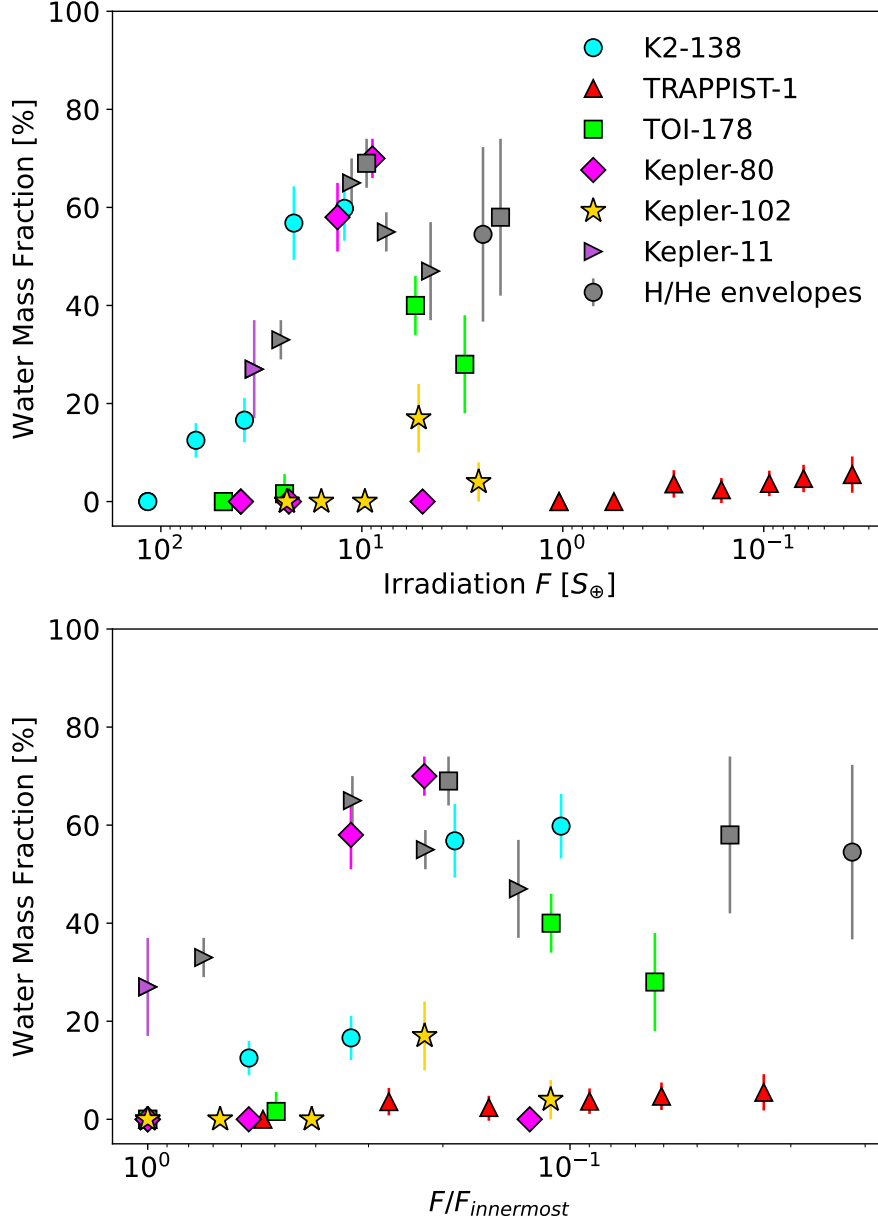


Figure 6.8.: WMF as a function of absolute irradiation (top panel), and irradiation relative to the innermost planet (bottom panel) for multiplanetary systems. The absolute stellar incident flux is expressed in Earth irradiation units, $S_{\oplus} = 1361 \text{ W/m}^2$, whereas the relative irradiation is normalised to the incident flux of the innermost planet in each system, $F_{innermost}$. Planets with $d_{obs-ret} > 1\sigma$ and high WMF values in our analysis are indicated in grey.

6.3. Low multiplicity systems

In this Section, I present a compilation of planetary systems with a multiplicity of less than five planets, for which I have run our MCMC interior-atmosphere analysis. I also introduce one system, TOI-1233 (Hoyer et al. 2022), which hosts five confirmed planets, but only two of them have masses available. In the following subsections, I describe system by system, the interior composition of their planets, and the implications for their formation history. The publications to which each

planetary system corresponds is indicated at the beginning of each subsection.

Planet	T_{irr} [K]	Scenario	CMF	VMF	Fe/Si	Section
K2-32 e	1066	Dry	0.62-0.98	0	0.80-34.2	6.3.1
K2-233 b	1121	Dry	0.20-0.70	0	0.0-5.5	6.3.1
K2-233 c	790	Dry	$0.90^{+0.03}_{-0.20}$	0	9.35 ± 7.05	6.3.1
LHS1140 b	379	No Fe/Si	0.49 ± 0.07	0.03 ± 0.07	2.17 ± 0.83	6.3.2
LHS1140 c	709	No Fe/Si	0.59 ± 0.05	$(0.0^{+6.0}_{-0.0}) \times 10^{-3}$	2.71 ± 0.59	6.3.2
TOI-220 b	806	With Fe/Si	0.09 ± 0.03	0.58 ± 0.14	0.64 ± 0.10	6.3.3
HD 207897 b	632	No Fe/Si	0.25 ± 0.18	0.24 ± 0.13	1.54 ± 1.81	6.3.4
		With Fe/Si	0.19 ± 0.03	0.20 ± 0.07	0.74 ± 0.09	
TOI-1233 b	1202	With Fe/Si	0.24 ± 0.06	0.05 ± 0.05	0.79 ± 0.08	6.3.5
TOI-1233 c	1021	With Fe/Si	0.20 ± 0.02	0.21 ± 0.05	0.79 ± 0.08	6.3.5
TOI-969 b	1223	No Fe/Si	0.19 ± 0.16	0.50 ± 0.09	$0.55^{+1.07}_{-0.55}$	6.3.6
		With Fe/Si	0.12 ± 0.04	0.47 ± 0.09	0.71 ± 0.26	
		No Fe/Si	0.70 ± 0.30	$(3.8 \pm 2.0) \times 10^{-3} *$	$17.0^{+25.8}_{-17.0}$	
TOI-2196 b	1856	With Fe/Si	$0.28^{+0.05}_{-0.09}$	$(1.4^{+0.2}_{-0.3}) \times 10^{-3} *$	$0.76^{+0.21}_{-0.26}$	6.3.7
		No Fe/Si	$0.0-1.0$	$(7.7^{+9.4}_{-3.2}) \times 10^{-3} *$	$6.26^{+11.54}_{-6.26}$	
K2-106 b	2275	Dry	$0.49^{+0.16}_{-0.22}$	0	2.05 ± 1.52	6.3.8
		No Fe/Si	$0.63^{+0.19}_{-0.17}$	$(5.8^{+8.6}_{-5.8}) \times 10^{-5}$	$4.75^{+5.27}_{-4.75}$	
		With Fe/Si	0.24 ± 0.03	$(9.8 \pm 7.8) \times 10^{-5} \dagger$	0.71 ± 0.11	

Table 6.5.: Mean and uncertainties of core (CMF) and volatile (VMF) mass fractions; irradiation temperatures, T_{irr} ; and Fe/Si mole ratio for planetary systems with low multiplicity. * We consider a H/He atmosphere as volatile layer (see text). † The retrieved radius is $d_{obs-ret} = 1.7 \sigma$ larger than the observed value. The last column indicates the subsection in which their analysis is detailed.

Table 6.5 shows the mean and 1σ uncertainties of the CMF, volatile mass fraction (VMF), and Fe/Si mole ratio retrieved by our MCMC analysis. For planets positioned below the 100% mantle composition in the mass-radius diagram, we set the WMF = 0 in our simulations, leaving the CMF as the only free parameter, and the mass and radius as the observable data. For planets with volatiles, we may consider two scenarios, depending on the system: the data are the mass and radius only; and, if the host chemical abundances are available, the mass, radius and Fe/Si mole ratio are the observables.

For some of these planets, the VMF is calculated assuming a H/He atmosphere, since their density is compatible with a water-dominated envelope or a H/He layer (see for example, TOI-969 b, Sect. 6.3.6), or their radius is too inflated to be compatible with a WMF < 0.8. (see TOI-2196 b, Sect. 6.3.7). To estimate the VMF assuming a H/He envelope, I combine the core and mantle of our interior structure model with the mass-radius relations of Zeng et al. (2019) for rocky planets with gaseous H/He atmospheres. I obtain the atmospheric thickness for these planets by subtracting the radius of a rocky planet with a H/He atmosphere minus the radius of the same planet (same core and mass), with both radii obtained by Zeng et al. (2019). We define the surface gravity, $g_0 = GM/R^2$, where G is the gravitational constant, and M and R are the mass and radius of the core and the mantle together, respectively. Therefore, we can express the atmospheric thickness, z_{atm} , as a function of the surface gravity, g_0 ; and the H/He mass fraction, $x_{H/He}$. Both TOI-969 b and TOI-2196 b are

very irradiated, with $T_{irr} > 1000$ K, thus the most adequate mass-radius relations that [Zeng et al. \(2019\)](#) provide for these planets are those with a constant atmospheric temperature of 2000 K. To be consistent, in our interior model we establish a surface temperature of 2000 K and a surface pressure of 1 bar, although these parameters have a negligible effect in the total radius of a dry mantle and core.

6.3.1. K2-32 and K2-233

K2-32 and K2-233 are two multiplanetary systems, of four and three planets, respectively ([Heller et al. 2019](#); [David et al. 2018](#)). [Lillo-Box et al. \(2020\)](#) obtained new RV data and refined the planetary masses of these two systems. In this subsection, we analyse the inner planets of these systems, K2-32 e, K2-233 b and K2-233 c.

Their densities indicate that these are likely to be rocky planets without any volatiles. Thus, we perform the MCMC simulations setting a constant WMF = 0, while leaving the CMF as the only free compositional parameter. For K2-32 e and K2-233 b, the posterior distributions of the CMF are not Gaussian, so we express its 1σ confidence intervals as uniform intervals instead of a mean and a standard deviation (see Table 6.5). K2-32 e and K2-233 c are consistent with being super-Mercuries, while K2-233 b is slightly less Fe-rich, with a maximum CMF = 0.70. In contrast with the outer planets of their respective systems, they are rocky, dry planets, which suggests that these two systems present the dichotomy we observed in Sect. 6.2 in other multiplanetary systems between inner, rocky planets and outer, volatile-rich ones.

6.3.2. LHS1140

LHS1140 is an M dwarf star that hosts two low-mass planets. Their densities are compatible with rocky planets, discarding the presence of a H/He atmosphere. The inner planet, LHS1140 c, is highly irradiated, whereas LHS1140 b has an irradiation temperature that places it in the habitable zone. Thus, for LHS1140 c we use our interior-atmosphere model for planets that could present steam and supercritical phases, while for the outermost planet we assume that the hydrosphere is in liquid and high-pressure states. LHS1140 is too dim to obtain its chemical abundances, so we perform our MCMC analysis by using the masses and radii from [Lillo-Box et al. \(2020\)](#) as data.

In Table 6.5 we can see that the retrieved CMF for LHS1140 b is in the 0.40-0.50 range. If we assume a condensed water layer, LHS1140 c presents a $\text{CMF} = 0.45 \pm 0.10$, placing the inner planet of the system in the CMF range of 0.35-0.55. This means that if we were to assume that both planets have a CMF below 0.35, such as the Earth's value (0.32), they would be unlikely to present volatiles. Moreover, LHS1140 b is compatible with a liquid water content of zero to only 100 times the water content of Earth. This maximum water content translates into an ocean of a maximum thickness of ≈ 700 km, which favours a partially liquid-covered surface (eyeball state), or a surface partially covered in ice, known as snowball state ([Yang et al. 2020](#)). The presence of a water layer is further supported by the tentative detection of water vapour in transmission spectroscopy of LHS1140 b ([Edwards et al. 2021](#)). Nonetheless, the detection relies on stellar contamination models, and more atmospheric characterization observations are necessary to confirm it. For LHS1140 c, the WMF distribution is centered in zero, with a 1σ maximum value $\text{WMF} = 6 \times 10^{-3}$, which is equivalent to a surface pressure of 300 bar. This means that if LHS1140 had a high CMF (>0.4), it could present a secondary atmosphere. Otherwise, for CMFs in the 0.0-0.3 range, LHS1140 c mass and radius are compatible with a dry planet.

6.3.3. TOI-220

TOI-220 b is a warm sub-Neptune recently discovered by the Transiting Exoplanet Survey Satellite (TESS), for which [Hoyer et al. \(2021\)](#) derived precise mass, radius and stellar abundances. In Sect. 5.3 I show a comparison of the interior structure analysis with the adaptive and the non-adaptive MCMC for this planet. Henceforth, I discuss the CMF and WMF values obtained in the adaptive MCMC (see Table 6.5), although the difference between the two MCMC methods does not change our main conclusions on this system.

The CMF of TOI-220 b is substantially lower than Earth's value (0.32), since we use the Fe/Si mole ratio of its host star to constrain it in our analysis, and it is less enriched in Fe ($\text{Fe/Si} = 0.93$) than the Sun ($\text{Fe/Si}_\odot = 0.96$). We also employ the Mg/Si mole ratio to obtain a PDF in our MCMC analysis for the Mg number, Mg\# , which indicates the level of differentiation of the mantle (see Sect. 2.1). We obtain a uniform distribution that ranges between $\text{Mg\#} = 0.85$ and $\text{Mg\#} = 1.0$, which points to a highly differentiated interior. The WMF of TOI-220 b is compatible with the water content obtained for Solar System moons, such as Titan, while its maximum limit, $\text{WMF} = 0.72$, is below the mean cometary WMF ([McKay et al. 2019](#)). Nonetheless, given the degeneracy between envelope composition and mass, TOI-220 b could also have a less massive atmosphere dominated by H/He. This is further supported by atmospheric mass loss estimates due to Jeans escape and XUV photoevaporation ([Aguichine et al. 2021](#); [Hoyer et al. 2021](#)), which are low enough to suggest that TOI-220 b has retained most of its original gas reservoir.

6.3.4. HD207897

[Heidari et al. \(2022\)](#) report the discovery of another warm sub-Neptune orbiting a main sequence star, HD 207897 (TOI-1611), with TESS, providing precise mass, radius and stellar chemical abundance estimates. Two sets of mean and 1σ uncertainties are calculated for the mass of HD 207897 b with equal probabilities. We find that these masses are close enough to yield similar compositional and atmospheric parameters in our interior structure analysis. Therefore, in Table 6.5, we present the mean CMF, WMF and Fe/Si mole ratio obtained from both mass estimates.

In scenario 1, where only the mass and radius are considered as data, HD 207897 b can present a CMF close to zero (100% mantle) up to almost 0.45, which is the wide range of CMFs found in bare rocky planets via interior structure models ([Plotnykov and Valencia 2020](#); [Wang et al. 2022](#)). The CMF distribution is significantly narrower in scenario 2 due to the use of the stellar Fe/Si mole ratio in the MCMC analysis, which yields a $\text{CMF} \approx 20\%$. This is lower than Earth's value, because HD 207897 is less enriched in Fe than the Sun. However, the WMF distribution is very similar in both scenarios, showing that for volatile-rich planets, the total density and radius are the two observable parameters that constrain the volatile mass fraction. This is also the case for interior structure models that consider a H/He-dominated envelope ([Otegi et al. 2020](#)). We conclude that HD 207897 b has a likely $\text{WMF} \approx 0.20$, with a maximum $\text{WMF} = 0.30$, which makes it an interesting case, together with TOI-220 b, to study the survival of gaseous envelopes in highly irradiated environments.

6.3.5. TOI-1233

TOI-1233 hosts five low-mass planets ($M < 25 M_\oplus$), with irradiation temperatures ranging from 1200 to 600 K. This makes it an interesting system to be analysed in our sample to discuss its possible formation and evolution scenarios, similar to the multiplanetary systems studied in Sect. 6.2. [Hoyer et al. 2022](#) provide refined radii values for the complete system from TESS and CHEOPS data.

However, the masses of only the two innermost planets are precise enough to provide constraints on the interior composition. Therefore, we only perform our MCMC interior analysis for TOI-1233 b and c (see Table 6.5), adopting the masses, radii and stellar abundances provided by Hoyer et al. 2022.

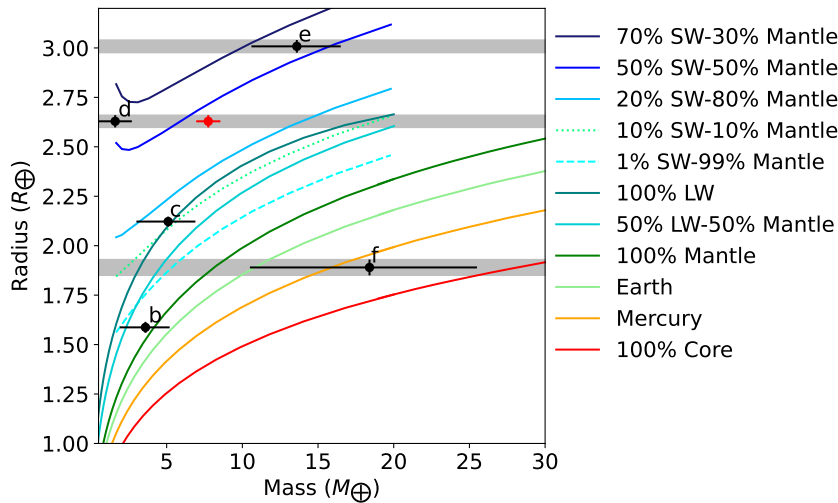


Figure 6.9.: TOI-1233 planets positions in the mass-radius diagram for planets with liquid (LW) and supercritical (SW) hydrospheres. Grey shaded areas indicate the 1σ region of the radius estimated by Hoyer et al. 2022 for planets d, e and f. Black dots indicate both the mass and radius obtained by Hoyer et al. 2022 for all planets, while the red dot marks the mass limits for planet d established by previous work (Bonfanti et al. 2021).

Under the assumption that the planetary Fe/Si mole ratio reflects the stellar Fe/Si, TOI-1233 b could be a dry planet or have a water-dominated atmosphere that constitutes up to 10% of its total mass. We explore the possibility of a dry rocky interior for the innermost planet, setting the WMF = 0 and leaving the CMF as the only free parameter in our MCMC analysis. We obtain a $\text{CMF} = 0.04^{+0.22}_{-0.04}$, which is compatible within uncertainties with the CMF derived when considering the stellar abundances. Therefore, the planets in TOI-1233 are likely to have CMFs of approximately 20%, below Earth's value (0.32). In addition, planet c presents a WMF between 0.16 and 0.26, being more volatile-rich than planet b. To have an idea of the possible volatile contents of the outer planets d, e and f, we plot in Fig. 6.9 the masses and radii obtained by Hoyer et al. 2022 for all planets. The radii of planets d and e suggest that these planets have more volatiles than planet c, no matter the assumed masses. However, the composition of planet f ranges from a super-Mercury, down to a rocky planet with a thin envelope. Hence, planets b to e seem to present an increase in volatile mass fraction with semi-major axes (and therefore, irradiation fluxes too), which we already observed in Sect. 6.2 in other multiplanetary systems. It is interesting to note that TOI-1233 seems to present 3:2 motion resonances (Hoyer et al. 2022), indicating inward migration, which might play a role in producing the water mass fraction gradient in multiplanetary systems as in TRAPPIST-1 and K2-138. Finally, the low volatile mass fraction of TOI-1233 f compared to planets c to e could be due to Jeans escape, where the solid core of the planet is not massive enough to retain the envelope. In this case, this would be the third multiplanetary system where we observed that the outermost planet suffers Jeans escape in contrast to the inner planets that are more volatile-rich (see also Kepler-11 f and Kepler-80 g in Sect. 6.2.3 and 6.2.5, respectively). Radial velocity or transit timing variation data are needed to constrain the mass of TOI-1233 and confirm this hypothesis.

6.3.6. TOI-969

TOI-969 is a K-type star that hosts one confirmed planet, TOI-969 b, and one candidate planet. Lillo-Box et al. (In rev.) derive precise radius, mass and host chemical abundances for this system, suggesting that TOI-969 b is a warm sub-Neptune.

The results of our MCMC interior analysis are shown in Table 6.5. The CMF derived in scenario 2, with the Fe/Si mole ratio as data, is compatible within uncertainties with the CMF of scenario 1. Moreover, in scenario 1, the CMF ranges from almost zero (100% mantle) up to 0.45, similar to HD 207897 b and the general population of naked rocky planets (see Sect. 6.3.4). The WMF distribution between scenario 1 and 2 has a negligible difference, which adds to the evidence that for sub-Neptunes, the volatile mass fraction is mostly constrained by the density. In addition to our analysis for water planets, we repeat our MCMC analysis assuming a H/He atmosphere, as described at the beginning of Sect. 6.3. In scenario 1, where only the mass and radius are taken into account, the MCMC explores CMF values from 0.40 up to 1.0, which is higher than the observed CMF of rocky planets (Plotnykov and Valencia 2020; Wang et al. 2022). Therefore, the mass fraction of H/He in this scenario is considered a maximum estimate. In contrast, when we consider the stellar Fe/Si mole ratio and a H/He atmosphere, we obtain a CMF that is compatible within uncertainties with the CMF derived in scenario 1 for a supercritical water envelope. Therefore, we are able to fit the same mass and radius data with a WMF = 0.50 with a water atmosphere, and a VMF $\approx 10^{-3}$ with a H/He envelope, showcasing the degeneracy between envelope mass and composition in sub-Neptunes. Finally, by comparing scenario 2 between the water and the H/He atmosphere, we observe that the CMF increases in the H/He case. This is due to the large portion of the total planetary mass composed by the Fe core, since the water envelope is significantly more massive than the H/He atmosphere.

6.3.7. TOI-2196

TOI-2196 is a single planetary system hosted by a Sun-like star, discovered by Persson et al. (2022). TOI-2196 b is also the only warm sub-Neptune in our sample that presents a mass of more than $20 M_{\oplus}$, and whose density is compatible with a WMF = 70% or more.

Hence, we exclusively perform our interior MCMC analysis for this planet assuming a H/He atmosphere. In Table 6.5, we can see that in scenario 1, the adaptive MCMC explores all possible CMF values, obtaining a PDF for the CMF that is uniform between 0 and 1. Nonetheless, the VMF is tightly constrained, and its PDF is very similar to that obtained in scenario 2. This is because for sub-Neptune and Neptune-class planets, the density and the radius are mostly determined by the amount of volatiles, as seen in Sect. 6.3.4 and 6.3.6, in agreement with Otegi et al. (2020). Persson et al. (2022) estimate XUV atmospheric mass loss as described in Aguichine et al. (2021). They estimate that the initial VMF of TOI-2196 b was 35%, in contrast to its current VMF $\approx 0.6\%$. This makes TOI-2196 b an interesting case to study in atmospheric loss and evolution, since it has retained a significant part of its envelope despite its high irradiation conditions.

6.3.8. K2-106

K2-106 is a Sun-like star that hosts two planets: the ultra-short period super-Earth K2-106 b, and the sub-Neptune K2-106 c. K2-106 b was originally reported as a very dense rocky planet, with a CMF > 0.60 (Sinukoff et al. 2017; Guenther et al. 2017), although later radial velocity measurements decreased its density, indicating that it was not as enriched in Fe as estimated (Dai et al. 2019).

Rodríguez Martínez et al. (2022) re-analyse photometric data from Kepler/K2 and radial velocity data from HIRES to refine the density of K2-106 b.

K2-106 b is highly irradiated ($T_{eq} = 2275$ K), meaning that if volatiles (including water) are present on its surface, these are in a gaseous or supercritical phase. I consider three scenarios: in scenario 1, I assume that K2-106 b is a completely dry planet, leaving the CMF as the only free non-observable parameter in our MCMC Bayesian analysis, while the WMF is constant and equal to zero, and the mass and the radius are the observables. Scenario 2 is similar to scenario 1, but we do not take any assumptions on the WMF, leaving it as a free parameter. Finally, in scenario 3 I also consider the Fe/Si mole ratio as an observable. Its mean value and uncertainties are calculated from the host stellar abundances provided by Rodríguez Martínez et al. (2022). The resulting value is $\text{Fe/Si} = 0.672 \pm 0.094$. The retrieved compositional parameters are shown in Table 6.5. In scenario 1, the retrieved CMF is compatible to that estimated by Rodríguez Martínez et al. (2022) with the interior model of Schulze et al. (2021), $\text{CMF} = 0.45^{+0.14}_{-0.16}$. This interior model has two layers, a Fe-rich core and a silicate mantle, which combined with a MCMC framework, is able to obtain the posterior distribution function of the core mass fraction in a similar way to our model. The 1σ confidence intervals obtained by both interior models overlap with the interval of $\text{CMF}_{star} = 0.29 \pm 0.06$. I also explore if the estimated CMF by our model assuming the presence of an atmosphere is compatible with that of the star. In scenario 2, the mean value of the CMF is higher than the CMF in scenario 1, which means that both scenarios are compatible within uncertainties. In scenario 2, a higher CMF is necessary to reproduce the observed mass and radius since a more Fe-rich bulk is more dense, leaving more space for an atmosphere to expand with respect to the dry scenario (scenario 1). Although the estimated CMF in scenario 2 is compatible within uncertainties with our scenario 1 and with the estimate of Schulze et al. (2021), it is not compatible with CMF_{star} . This means that if K2-106 b was to reflect the CMF estimated from the abundances of its host star, it would be very unlikely to have an atmosphere. If an atmosphere exists, its CMF would have to be greater than 0.46, which is the lower limit of the 1σ confidence interval in scenario 2. The parameters of this thin atmosphere are $P_{surf} = 184.9 \pm 120.8$ bar, $z_{atm} = 404 \pm 82$ km, $T_{surf} = 4154 \pm 326$ K and $A_B = 0.210 \pm 0.001$ of surface pressure, atmospheric thickness, surface temperature and Bond albedo, respectively. The observed mass and radius are well-reproduced by their respective posteriors in scenarios 1 and 2. However, in scenario 3 the MCMC Bayesian analysis can reproduce the Fe/Si mole ratio derived from the host stellar abundances, but the mass and radius are not compatible with their observed mean values and uncertainties. The retrieved CMF in scenario 3 is compatible with CMF_{star} , as expected, since both estimated are calculated from the chemical abundances of the host star. In scenario 3, the retrieved mass is lower than the observed value, whereas the retrieved radius is higher than the observed one, yielding a lower planetary density. This supports the conclusion we reached in scenario 2: K2-106 b cannot have an atmosphere and present the Fe/Si mole ratio (or CMF) of its host star.

6.4. JWST rocky planets

As mentioned in Sect. 4.4.2, when I run the atmosphere model within the MCMC algorithm, I need fast computations of the OLR and Bond albedo. For this reason, in the MCMC analysis I use the version of the atmospheric model that calculates spectra with a resolution of $R = 10$, since this is accurate enough to obtain the OLR and the Bond albedo within a few seconds. However, I use the k-correlated version of the atmospheric model to generate spectra with a higher resolution to assess whether the atmosphere of a planet with a given mass, radius, and composition would be

observable. Therefore, I modify the atmospheric module to take as input the k-correlated table in its original resolution, which is $R = 200$. This is adequate to produce theoretical emission spectra that can be the input for noise generators, such as Pandexo (Batalha et al. 2020) for the James Webb Telescope (JWST).

6.4.1. TRAPPIST-1 c

TRAPPIST-1 c has been proposed to be observed in thermal emission with the Mid-infrared instrument (MIRI) in JWST Cycle 1 (Kreidberg et al. 2021). It will be observed with the F1500W filter during 4 eclipses, which is the filter centered at $\lambda = 15 \mu\text{m}$. We analyse TRAPPIST-1 c with our adaptive MCMC with a water and a CO_2 -dominated atmosphere. Fig. 6.10 shows the marginal posterior distributions (PDF) in 1D and 2D for this analysis. The CMF agrees well with the value obtained previously with the non-adaptive MCMC (see Table 6.2). In contrast, the distribution of the WMF derived by the adaptive MCMC is wider than the non-adaptive one, with mean and standard deviation values of $\text{WMF}_{\text{adapt}} = (11.0 \pm 5.6) \times 10^{-6}$, and $\text{WMF}_{\text{non-adapt}} = (0.0^{+2.7}_{-0.0}) \times 10^{-6}$, respectively. This is because the adaptive MCMC is more effective at exploring the corner region of the compositional parameter space, where the WMFs are close to zero. More models in this region are accepted in the posterior distribution, and therefore it becomes wider with larger standard deviation than the non-adaptive posterior distribution. The corresponding surface pressure derived by the adaptive MCMC is $P_{\text{surf}} = 15 \pm 7$ bar. This 1σ confidence interval must be taken carefully since the PDF of the surface pressure does not present a Gaussian distribution shape. A WMF of zero is also compatible with the density of TRAPPIST-1 c. Consequently, we can conclude that TRAPPIST-1 c could have a H_2O atmosphere of up to ≈ 25 bar of surface pressure, or no atmosphere at all. The results of the adaptive MCMC (with the k-uncorrelated atmosphere model), and the non-adaptive one with the k-correlated model of Pluriel et al. (2019), agree that a H_2O atmosphere in TRAPPIST-1 c would have a surface temperature between 1000 and 1500 K, and an atmospheric thickness of 150 to 250 km. The retrieved Bond albedos differ by 0.05 due to the use of different atmospheric opacity data, as we discussed in Sect. 4.6.2.

For a CO_2 -dominated atmosphere, we retrieve a similar CMF to the water case, although the volatile mass fraction increases by one order of magnitude, $\text{VMF} = (2.49 \pm 2.07) \times 10^{-5}$. The molecular weight of CO_2 is higher than that of water vapour, producing a more compressed atmosphere for a similar surface pressure and temperature. In addition, the radiative properties (i.e opacity) of CO_2 yields a lower surface temperature for the same irradiation conditions in comparison to a water-dominated envelope, which contributes to a lower atmospheric thickness. As a consequence, the models with a CO_2 envelope can accommodate a more massive atmosphere for TRAPPIST-1 c than the water models, retrieving a surface pressure of $P_{\text{surf}} = 35 \pm 29$ bar. The surface temperature and atmospheric thickness are $T_{\text{surf}} = 807 \pm 102$ K, and $z_{\text{atm}} = 63 \pm 12$ km, respectively.

We assume the atmospheric parameters retrieved in our adaptive MCMC analysis and generate emission spectra with their respective temperature-pressure profiles. Consecutively, we bin the emission spectrum using the response functions of each of the MIRI photometry filters¹ (Glasse et al. 2015; Piette et al. 2022). The mean flux, $\langle f_\lambda \rangle$, of an emission spectrum, $f(\lambda)$, observed with a filter with transmission function $R(\lambda)$, is defined in Eq. 6.1 (Stolker et al. 2020). We consider random Gaussian noise of 100 ppm for each filter to derive the uncertainties of the mean flux (Lustig-Yaeger et al. 2019; Piette et al. 2022).

¹<http://svo2.cab.inta-csic.es/svo/theory/fps3/index.php?id=JWST>

$$\langle f \rangle = \frac{\int f(\lambda) R(\lambda) d\lambda}{\int R(\lambda) d\lambda} \quad (6.1)$$

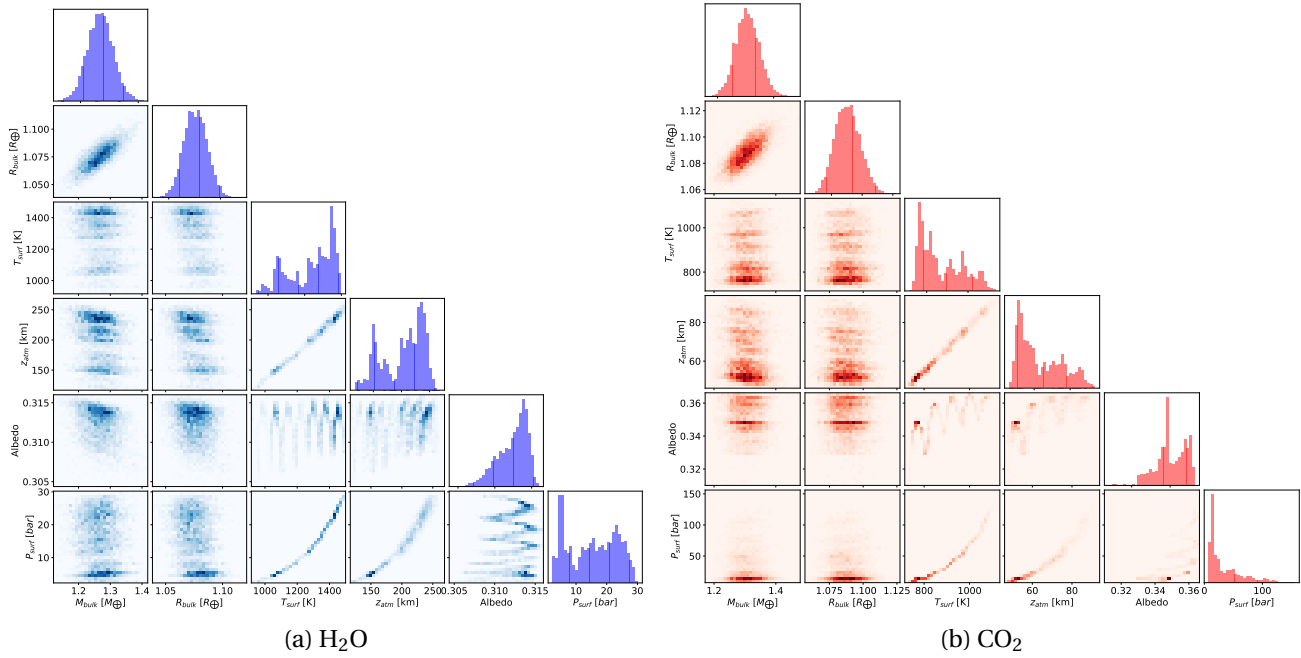


Figure 6.10.: PDF pyramid plot of the atmospheric parameters of TRAPPIST-1 c for H₂O (left panel) and CO₂ (right panel) atmospheres. The atmospheric parameters include the surface temperature, T_{surf} ; atmospheric thickness, z_{atm} ; albedo; and the surface pressure, P_{surf} . We assume the Fe/Si mole ratio estimated in Sect. 6.1.1 (Unterborn et al. 2018).

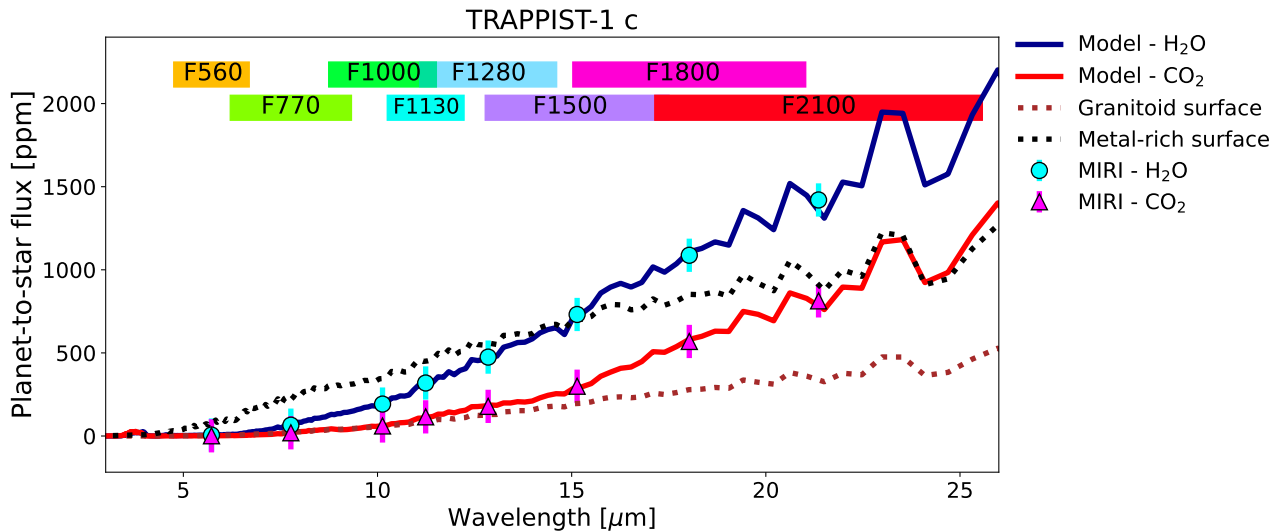


Figure 6.11.: Simulated emission spectra of TRAPPIST-1 c with MIRI photometric filter mean fluxes for water and CO₂ atmospheres, obtained with the high-resolution, k-correlated version of MSEIRADTRAN. Color boxes indicate the wavelength ranges of the different filters.

Fig. 6.11 shows the complete emission spectra and mean filter fluxes for TRAPPIST-1 c. We observe that for the nIR filters ($\lambda = 5.60$ to $11.30 \mu\text{m}$), both atmospheres have very similar fluxes that are compatible within uncertainties, which makes it not possible to distinguish between the two compositions in these wavelengths. For the filters F1500, F1800 and F2100, the mean flux uncertainties of the water and the CO_2 atmospheres do not overlap, allowing the different compositions to be distinguished. Therefore, observing TRAPPIST-1 c in emission with filter F1500, as proposed by Kreidberg et al. (2021), is well-adapted to differentiate between a water- or a CO_2 -dominated atmosphere. We also consider the possibility of a bare surface in TRAPPIST-1 c, given the high probability obtained in our MCMC analysis for a volatile mass fraction equal to zero. Hu et al. (2012) obtain the emission spectra of bare terrestrial surfaces for different minerals. I estimate the brightness temperature for the irradiation conditions of TRAPPIST-1 c for two minerals, a metal-rich surface and a granitoid one, since these are the two surfaces with the highest and lowest emission for the same irradiation conditions, respectively. I approximate the emission spectrum of these surfaces to that of a black body with a temperature equal to the estimated brightness temperature. The emission flux in the F1500 filter for a water atmosphere is slightly higher than of the metal-rich surface. If TRAPPIST-1 c emission flux in this filter is 731 ppm or higher, it is indicative of a thin, water-rich atmosphere. For very low emission fluxes (≈ 300 ppm), TRAPPIST-1 c would present a CO_2 -dominated atmosphere. For fluxes between 730 to 400 ppm, TRAPPIST-1 c would have no atmosphere, with an emission that corresponds to a bare surface. The surfaces with the lowest emission are granitoid, feldspathic, or clay (Hu et al. 2012).

6.4.2. 55 Cancri e

55 Cancri e is a super-Earth in a close orbit ($P = 0.66$ days) to a bright star, for which several interior and atmospheric hypothesis have been proposed. Madhusudhan (2012) explored a carbon-rich interior given the high C/O ratio found for its host star, showing that in this case the planetary bulk density would be lower than that of a silicate-rich mantle planet, such as Earth. They concluded that a volatile layer would not be necessary to account for its density. On the other hand, a classical Fe-rich core and a silicate mantle are compatible with a volatile envelope rich in secondary atmosphere species. Furthermore, the absence of a H/He-dominated envelope seems likely due to the lack of hydrogen and helium emission and absorption lines in the spectrum (Ehrenreich, D. et al. 2012; Zhang et al. 2021). The presence of a secondary atmosphere is supported by phase curve data from the *Spitzer* Space Telescope (Angelo and Hu 2017). A fully H_2O -dominated atmosphere has been discarded, since it will require the presence of water and hydrogen simultaneously in the atmosphere due to water dissociation. Therefore, the most likely composition for the atmosphere of 55 Cancri e is a mixture of silicate compounds (Keles et al. 2022), such as HCN, detected by Tsiaras et al. (2016), with traces of water (detected by Esteves et al. 2017), or CO_2 . Despite the water and CO_2 -dominated atmospheres being discarded by the current data, it is still interesting to carry out our interior-atmosphere analysis for 55 Cancri e, since the scale height of a silicate high-molecular weight atmosphere might be similar to that of an envelope with the two compositions we consider. The scale height influences, together with the abundances, the intensity of a line in the spectrum of an exoplanet.

	Data	H ₂ O	CO ₂
$M [M_{\oplus}]$	$7.99^{+0.32}_{-0.33}$	$7.99^{+0.29}_{-0.32}$	8.17 ± 0.29
$R [R_{\oplus}]$	1.875 ± 0.029	1.877 ± 0.020	1.850 ± 0.021
Fe/Si	0.60 ± 0.14	0.62 ± 0.14	0.48 ± 0.13
CMF		0.20 ± 0.05	0.15 ± 0.05
VMF		$(6.7^{+7.4}_{-5.9}) \times 10^{-5}$	$(5.1^{+5.9}_{-4.6}) \times 10^{-5}$
P_{surf} [bar]		209 ± 93	> 300
T_{surf} [K]		4161 ± 199	4035 ± 597 (at 300 bar)
z_{atm} [km]		522 ± 46	152 ± 33 (from 300 bar)
Bond albedo		0.191 ± 0.001	0.351 ± 0.004

Table 6.6.: MCMC retrieved mean value and 1σ uncertainties for observable (mass, radius and Fe/Si mole ratio), compositional (core and volatile mass fractions) and atmospheric (surface pressure and temperature, atmospheric thickness and Bond albedo) parameters of 55 Cancri e.

Hu et al. (2021) have proposed to observe 55 Cancri e in emission spectroscopy combining NIRCcam F444W filter (3-5 μm), and MIRI’s Low Resolution Spectrograph (MIRI LRS; 5-14 μm). Thus, we run our MCMC analysis with our interior-atmosphere model, and then use the retrieved mean surface conditions to generate a pressure-temperature profile and a emission spectrum with a resolution $R = 200$. We adopt mass and radius data for 55 Cancri e from Bourrier, V. et al. (2018) (see Table 6.6), and host stellar abundances from Luck (2016), from which we derive a Fe/Si = 0.60 ± 0.17 . In Table 6.6, we observe that a water-dominated atmosphere reproduces well the observed data, with a surface pressure higher than 200 bar, whereas the CO₂ envelope is not extended enough to match the density of 55 Cancri e, yielding a more dense interior. At temperatures higher than 4000 K, CO₂ would not be the dominant species in a C-rich atmosphere, but CO. This changes the emission of the atmosphere as CO is a different absorber from CO₂. A CO-rich atmosphere could also explain the low-density of 55 Cancri e in this scenario, since CO has a lower molecular weight than CO₂, yielding a larger atmospheric scale height. When H/He is not included in the interior modelling, water as a trace species is necessary to explain the low density of 55 Cancri e, since a purely dry silicate atmosphere will have a smaller thickness than a CO₂ atmosphere due to their heavier molecular weights under similar atmospheric surface conditions. Adding silicate absorbers decreases the total planetary radius in H/He envelopes (Misener and Schlichting 2022). However, new models are necessary to explore the effect of silicates in atmospheres that have lost their primordial H/He. A planet with no volatiles could match the low planetary density if the core and the mantle were less dense than that of an Earth-like interior, pointing to a carbon-rich mantle as suggested by Madhusudhan (2012). The emission spectrum in this scenario would be that of a bare surface, requiring terrestrial surface models such as the ones presented in Hu et al. (2012).

Henceforth, I generate an emission spectrum for the water-rich atmosphere. I input the emission spectrum calculated with MSEIRADTRAN to Pandexo (Batalha et al. 2020), JWST observation and noise simulator. I assume two eclipse observations, with a baseline of $3.2 + 1.6 = 4.8$ hours of total observing time. I select the automatic optimized number of groups per integration, as well as a saturation limit of 80%, as indicated by Hu et al. (2021). I adopt a constant minimum noise floor of 100 ppm, similarly to TRAPPIST-1 c. Fig. 6.12 shows the complete emission spectrum of 55 Cancri e from 3 to 14 μm , thanks to the combined simulated spectra of NIRCcam 4.44 μm filter and MIRI LRS. We notice that for wavelengths below $\approx 3.5 \mu\text{m}$, the noise is too high to distinguish any spectral lines. Nonetheless, the rest of the spectral coverage of the proposed observations have low noise level, which makes the spectral features of water easy to identify with JWST in the high-molecular

weight atmosphere scenario of 55 Cancri e.

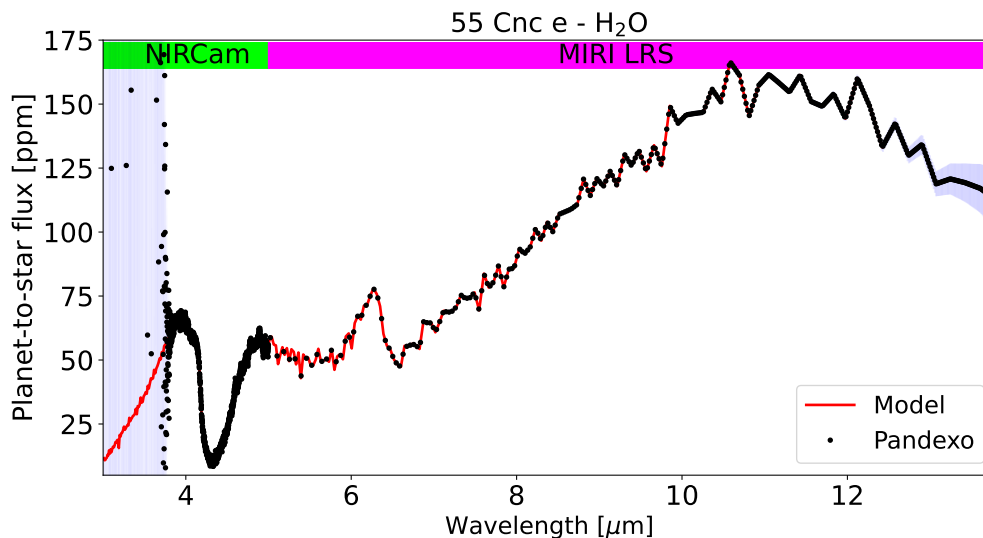


Figure 6.12.: Emission spectrum with the k-correlated, high-resolution MSEIRADTRAN model for a water-rich atmosphere in 55 Cancri e with NIRCam and MIRI LRS.

6.5. Planet composition properties

In this section, I gather together the data obtained from the individual analyses of all planetary systems I performed in Sect. 6.1 to 6.4, and discuss the properties I observe in our sample in the context of the general low-mass planet population. This population presents two characteristic peaks in the radius histogram, with a gap in between these two peaks. The first peak is located at $R = 1.3 R_{\oplus}$, where its planets (super-Earths) have densities consistent with a rocky, volatile-poor composition. The second peak is found at $R = 2.4 R_{\oplus}$, corresponding to the volatile-rich sub-Neptunes. This gap is known as Fulton gap (Fulton et al. 2017; Fulton and Petigura 2018), and it indicates that planets with radii between $1.5 R_{\oplus}$ and $2.0 R_{\oplus}$ are scarce in the low-mass planet population. Our sample has 46 exoplanets whose radii span a wide range values, from 0.5 to $4.5 R_{\oplus}$, representing the sub-populations of both super-Earths and sub-Neptunes. This can be seen in Fig. 6.13, where I show the mass-radius diagram of our sample, with their respective distribution function histograms. The grey circles indicate the sample from the NASA Exoplanet Archive². I selected this sample for planets with masses below $30 M_{\oplus}$, that also have available mass and radius measurements with their respective uncertainties.

²<https://exoplanetarchive.ipac.caltech.edu>. Updated on June 28, 2022.

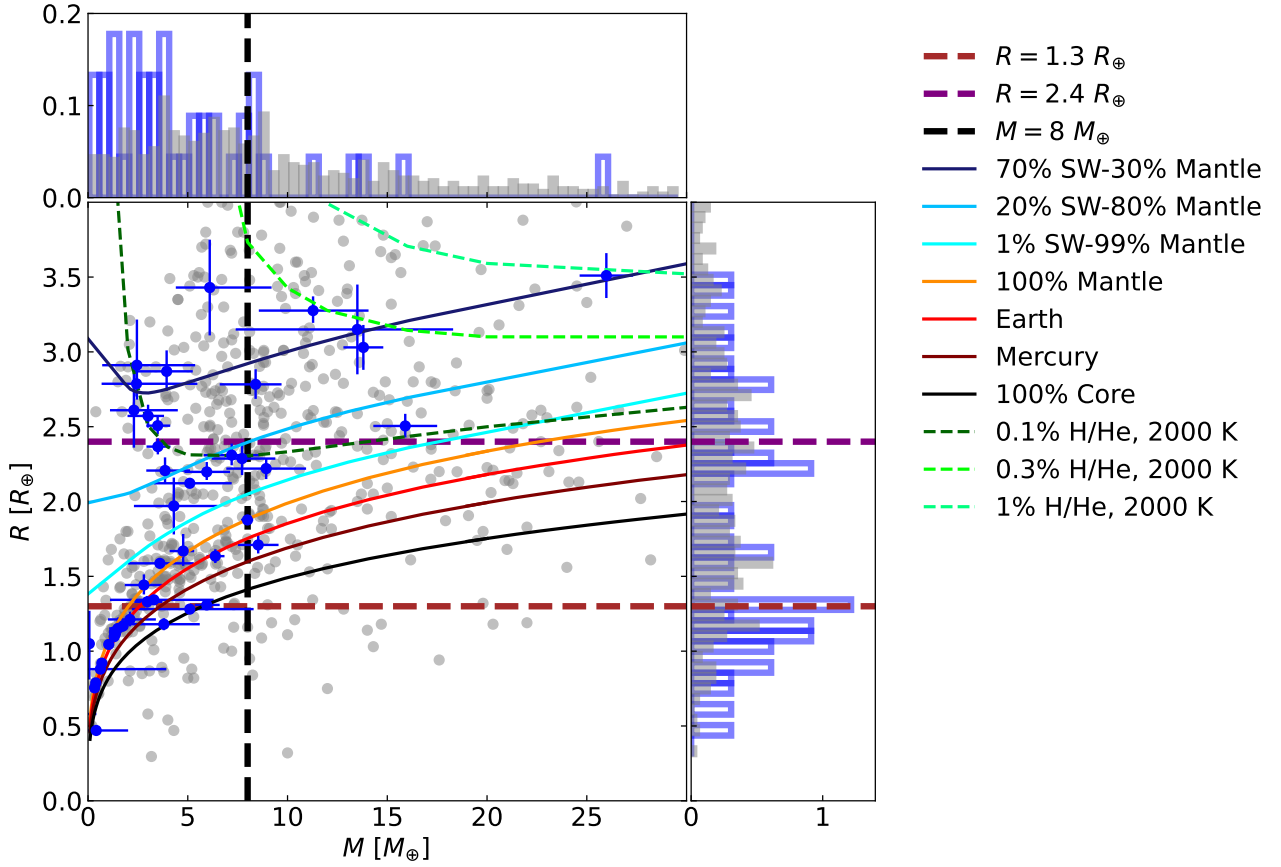


Figure 6.13.: Mass-radius diagram and their respective histograms of our sample of planets (blue circles), together with a larger sample (grey circles) from NASA Exoplanet Archive. Theoretical relationships for different CMFs are shown for dry planets, as well as for varying WMFs for water-rich planets, obtained with MSEI, our interior-atmosphere model (Brugger et al. 2017; Acuña et al. 2021). In addition, I also plot mass-radius relations for H/He-rich planets with $T_{irr} = 2000$ K from Zeng et al. (2019).

After having checked that our sample presents planets from the two low-mass sub-populations, I proceed to display the water mass fraction and the core mass fractions. I discussed in the analyses of TRAPPIST-1 (Sect. 6.1.2), HD207897 (Sect. 6.3.4), and TOI-2196 (Sect. 6.3.7), that the WMF estimates are very similar between scenario 1 and 2 for the same planet, especially for volatile-rich planets. Hence, for the WMF histogram I select only one of the two scenarios when a planet has been analysed under both, to avoid having the same planet repeated in the sample. In Fig. 6.14, I show the total radius as a function of WMF for our sample, with their respective histograms.

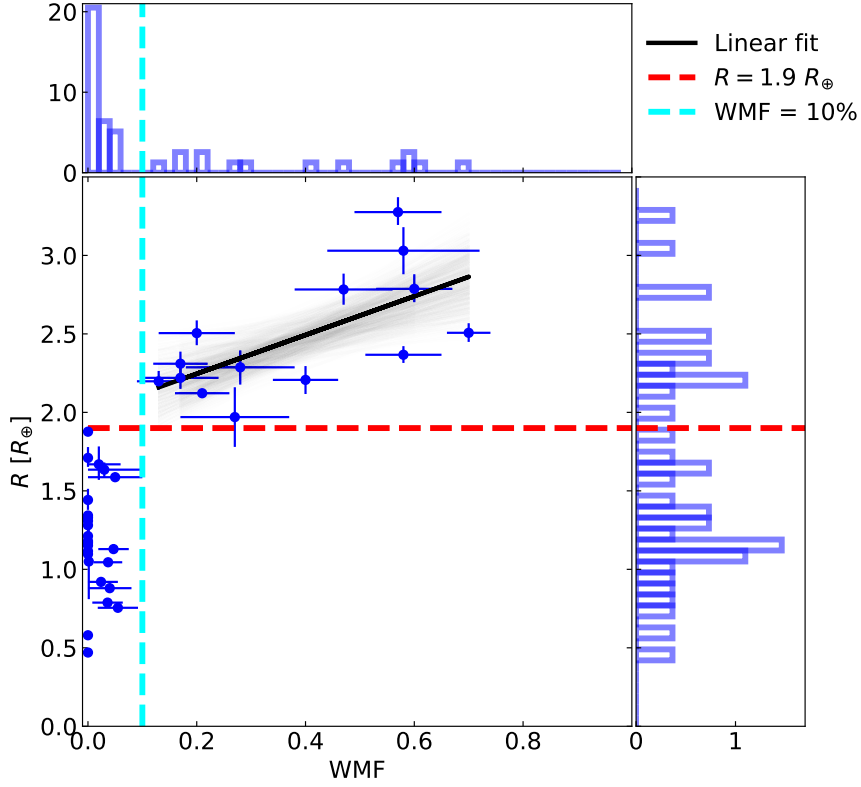


Figure 6.14.: Water mass fraction-radius diagram and their respective histograms of our sample of planets (blue circles). We show a linear fit (black) for planets with $R > 1.9 R_{\oplus}$. The grey lines indicate the realizations of the fit within the bootstrap method used to estimate the uncertainties of the slope and the intercept.

We can see that super-Earths ($R = 1.3 R_{\oplus}$) correspond to planets whose WMF range between zero and 0.1, where the WMF histogram shows that the distribution is centered between $\text{WMF} = 0$ and $\text{WMF} = 0.02$. In contrast, the sub-Neptune population ($R = 2.4 R_{\oplus}$) is constituted by planets with WMF from 0.2 up to the maximum limit we set in our MCMC simulation analysis, which is $\text{WMF} = 0.7$. The WMF distribution for sub-Neptunes seems mostly uniform, although there is a hint for a gap at $\text{WMF} = 0.3 - 0.4$. Pebble-driven formation can generate planets with ice mass fractions above 15%, although more than 30% are difficult (Liu et al. 2020). Moreover, Gupta and Schlichting (2019) find that, with their core-powered mass loss model, the Fulton gap with both the super-Earth and sub-Neptune peaks can be reproduced with planet bulks whose ice mass fraction is less than 20%. Taking into account that for these sub-Neptune planets there exists a degeneracy between envelope composition (water vs H/He) and envelope mass, the planets in our radius-WMF diagram with $\text{WMF} > 0.4$ could be rich in H/He instead of water. In this case, the gap at $\text{WMF} = 0.3-0.4$ could be a transition between the water-rich planets with a maximum $\text{WMF} \approx 0.20-0.25$, and planets with a non-negligible amount of H/He in their atmospheres. Nonetheless, the gap in the WMF distribution could also be an effect due to the size of our sample, being necessary to increase the number of planets in the sample to confirm this.

I select a sub-sample of the exoplanets shown in Fig. 6.14, with the selection criteria of $R = 1.9 R_{\oplus}$. For this sub-sample of super-Earths, the radius is uncorrelated to the WMF, with a Pearson's correlation coefficient $r_{\text{WMF}-R, \text{superEarth}} = -0.098$. This is indicative that the radius is mostly determined by the total mass and the core mass fraction for $R < 1.9 R_{\oplus}$. On the contrary, for sub-

Neptunes ($R > 1.9 R_{\oplus}$), I estimate a correlation coefficient $r_{WMF-R, subNeptune} = 0.66$, suggesting a correlation between radius and water mass fraction for these planets. I fit by regression a linear relation between radius and WMF (Fig. 6.14, black), and estimate the uncertainties of the slope and the intercept by combining the bootstrap and the Monte Carlo (MC) methods to take into account simultaneously the uncertainties and the dispersion of the data points. I obtain a slope of $a = 0.24 \pm 0.47 R_{\oplus}$, and an intercept $b = 2.00 \pm 0.20 R_{\oplus}$ for $R(x_{H_2O}) = a \times x_{H_2O} + b$. This empirical radius-WMF relation can be useful to infer the typical radius of sub-Neptunes given an ice water mass fraction, for example in planet population synthesis, where the observed distribution of the radius can be linked to that of the initial water content of the planet embryos (Mordasini 2018).

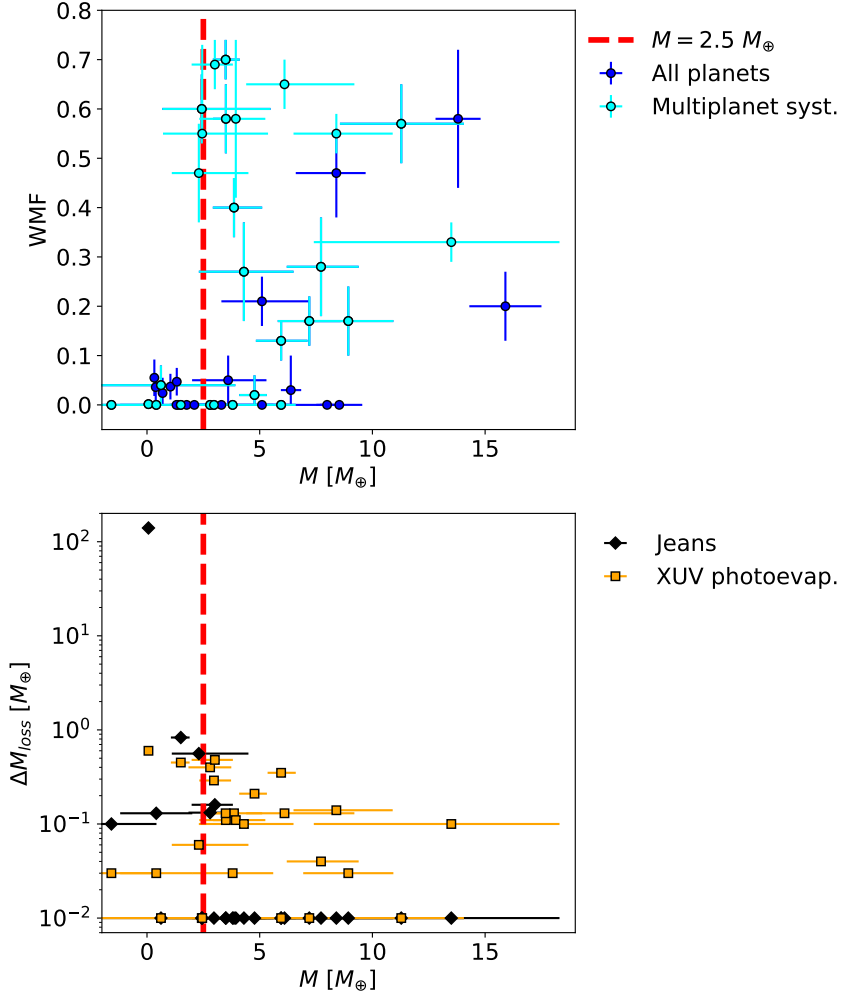


Figure 6.15.: Water mass fraction (upper panel) and lost atmospheric mass (lower panel) as a function of total planetary mass. The mass-WMF diagram includes all analysed planets (dark blue), while the lost mass diagram shows only planets in multiplanetary systems, since these are the only ones for which we have Jeans and XUV photoevaporation mass loss estimates.

Similarly to the radius-WMF diagram, I compare the water content with the total planetary mass for our sample of analysed planets. In Fig. 6.15 (upper panel) we observe that the planets in the lower mass range, $M < 2.5 M_{\oplus}$, are confined within a $WMF \leq 0.10$, while planets with more than $2.5 M_{\oplus}$ have very diverse compositions, from dry ($WMF \approx 0$) up to our model's upper limit of $WMF = 0.70$. Within these two quadrants in the mass-WMF diagram, no correlation is observed between

the two variables. This suggests that some process (or processes) prevent planets in the low range of mass from accreting large volatile reservoirs during planet formation, or conserving them during planet evolution. Planetesimal accretion is able to produce planets with WMF $> 30\%$, being mostly efficient for $M < 5 M_{\oplus}$ (Miguel et al. 2020), while pebble accretion can form planetary low-mass embryos with ice mass fractions up to 30% in turbulent disks (Liu, Beibei et al. 2019; Liu et al. 2020). Therefore, the initial formation of very low-mass planets with WMFs greater than 10% is perfectly feasible. This highlights the loss of the volatile reservoir as the cause of the gap in the mass-WMF diagram. Jeans atmospheric escape is more dependent on the planetary bulk mass than XUV photoevaporation, which is mostly caused by the high-energy emission of the stellar host. To test whether the gap in the mass-WMF diagram is produced by Jeans escape, I display in Fig. 6.15 (lower panel), our Jeans mass loss estimation for multiplanetary systems (see Table 6.4) as a function of total planetary mass. For planets with very low Jeans escape atmospheric loss ($\Delta M_{H_2} < 10^{-2} M_{\oplus}$, their total masses are uniformly distributed. However, planets with a substantial atmospheric loss by Jeans escape (black diamonds) are confined within the $M = 0$ to $\approx 3 M_{\oplus}$ mass range, whereas planets with strong XUV photoevaporation (orange squares) disperse over the whole total low-mass planet range (0 to $15 M_{\oplus}$, as seen in Fig. 6.15, lower panel). Thus, Jeans atmospheric escape is likely to be the mechanism responsible for the lack of volatile-rich planets at $M < 3 M_{\oplus}$ in the mass-WMF diagram.

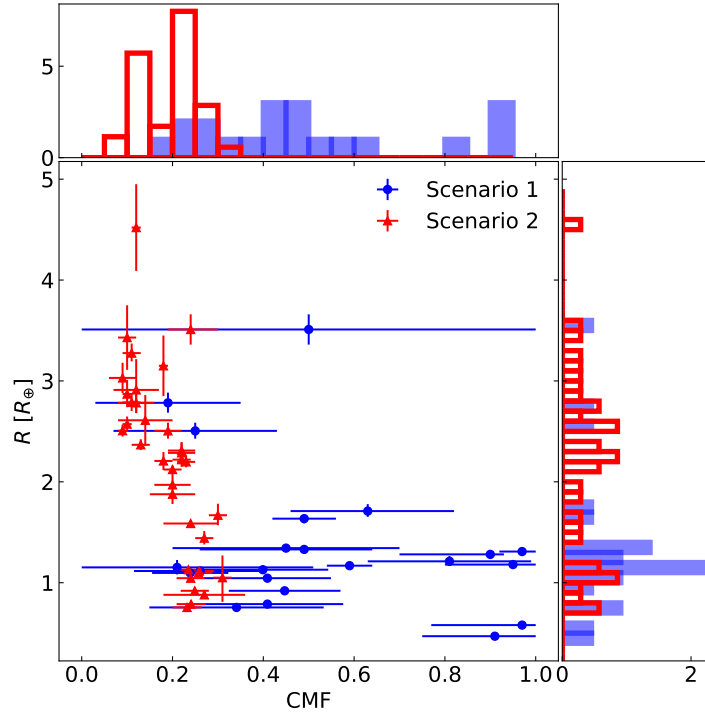


Figure 6.16.: CMF-radius diagram and their respective histograms for all planets in our sample. Scenario 1 corresponds to CMF estimates obtained based only on mass and radius data, while in scenario 2, the Fe/Si mole ratio is considered in addition to the mass and the radius.

In Fig. 6.16, I show the total planetary radius as a function of CMF, separating the data depending on whether we considered the Fe/Si mole ratio in our MCMC analysis (scenario 2) or not (scenario 1). For scenario 1, most of our estimates are for planets whose radius is below $1.9 R_{\oplus}$. For the three sub-Neptunes in scenario 1, the CMF presents large 1σ confidence intervals, since their radius is

mostly determined by the mass and the volatile mass fraction. I select the super-Earths ($R < 1.9 R_{\oplus}$) in scenario 1, and I calculate Pearson's correlation coefficient. I obtain $r_{CMF-R, superEarth} = -0.13 \pm 0.10$, which is compatible with a zero value, indicating that there is no correlation between the two variables. This selection contains planets that, despite being dominated by refractories, present thin envelopes. This might be contributing to the lack of correlation we observe in scenario 1 for super-Earths. Therefore, I choose the dry scenarios within scenario 1 data, where the WMF is constant and equal to zero. In some of the dry scenario 1 cases, the assumption of $WMF = 0$ is not enough to explore the full 1σ confidence interval of the mass and radius input data with the MCMC, especially the total mass. This means that we need to combine dry and wet simulations to explore the complete mass-radius region. I discard these cases, leaving us with five pure dry scenario 1 planets: TOI-178 b, Kepler-80 d and e, K2-233 b, and K2-106 b. This sample is not large enough to draw any conclusions on the radius-CMF relations of dry low-mass planets. Therefore, I conclude that super-Earths in our sample have a wide variety of CMFs, from $CMF = 0.20$ up to more than 0.90. We need a larger sample of rocky planets with precise masses to determine whether there is a correlation between radius and CMF for rocky, dry planets in scenario 1. In scenario 2, I estimate Pearson's correlation coefficient $r_{CMF-R, scenario2} = -0.74 \pm 0.07$, which indicates that the radius is correlated to the CMF. This correlation is even stronger for planets with $R < 3 R_{\oplus}$ (see Fig. 6.16), for which the correlation coefficient is $r = -0.81$. In addition, in scenario 2, the stellar Fe/Si mole ratio produces a CMF distribution whose mean and dispersion are lower than those of the CMF distribution in scenario 1. In the former, the mean $CMF = 0.19$, with a standard deviation of 0.07, while for the latter the CMF distribution is centered at $CMF = 0.53$, with a standard deviation of 0.27. Plotnykov and Valencia (2020) and Schulze et al. (2021) find that the primordial stellar CMFs (equivalent to our scenario 2) also have a narrower distribution than the CMF distribution calculated with the densities of rocky planets only. Plotnykov and Valencia (2020) select planets with less than 25% uncertainties in mass and radius, and whose position in the mass-radius diagram is below the 100% mantle composition line, named the Rocky Threshold Radius (RTR), while I do not carry any selection in our sample. Despite this difference, our general CMF agrees within uncertainties with their estimate ($CMF = 0.24^{+0.33}_{-0.18}$) for scenario 1. In scenario 2, our mean value is the lower limit of their 1σ confidence interval ($CMF = 0.32^{+0.14}_{-0.12}$), being 10% lower than their mean estimate. Since the CMFs have narrow uncertainties when the stellar Fe/Si is considered, scenario 2 might be more sensitive to selection effects than scenario 1. In our sample, I not only include planets below the RTR, but also planets above it, with WMFs ranging from 10% up to $\approx 50\%$. These planets might be lowering the mean of the general distribution of the CMF in scenario 2, in comparison to a sample that would only include strictly dry planets. Additionally, the stellar Fe/Si mole ratio of our sample is $Fe/Si = 0.82 \pm 0.10$, whereas Plotnykov and Valencia (2020) have a sample whose distribution is centered at the solar value ($Fe/Si_{\odot} = 0.97$). Therefore, our sample of planets in scenario 2 might be hosted by stars that are Fe-poor compared to the Sun, producing an overall CMF that is lower than an Earth-like $CMF \approx 0.32$.

Finally, I explore how the WMF and CMF are affected by the uncertainties of the mass, radius and stellar Fe/Si data. In Fig. 6.17 (upper left panel), we can observe that there is a hint of a correlation between the uncertainty of the WMF and the radius error for sub-Neptunes. Pearson's correlation coefficient for Δx_{H_2O} vs ΔR is $r = 0.67$, while the mass presents a coefficient of $r = 0.60$, indicating a slightly weaker correlation. This agrees with Otegi et al. (2020), who find that the uncertainty of the radius of sub-Neptunes determines mostly the error bars of the volatile mass fraction in comparison to the mass. This is because an increase of VMF will change significantly the radius of the volatile layer, and thus the volatile radius ratio, while the atmospheric mass will still remain very small compared to the mass of the core or the mantle. In the case of super-Earths, we do

not observe any trends between the uncertainties of the WMF and the data. We can therefore conclude that only for sub-Neptunes, an improvement of the data uncertainties (especially the radius) entails a refinement of the water content. For super-Earths, we only find a correlation between the uncertainty of the radius and the errors of the CMF in scenario 2 for very rocky planets ($R < 1.5 R_{\oplus}$). This confirms that an improvement of the observational error bars for planets with rock-dominated compositions yields a reduced uncertainty in the CMF only (Otegi et al. 2020). For sub-Neptunes, the uncertainties of the CMF in scenario 2 are completely uncorrelated to the errors of the mass and radius (see Fig. 6.17, lower left), since for $R > 1.9 R_{\oplus}$, the CMF is determined by the Fe/Si mole ratio.

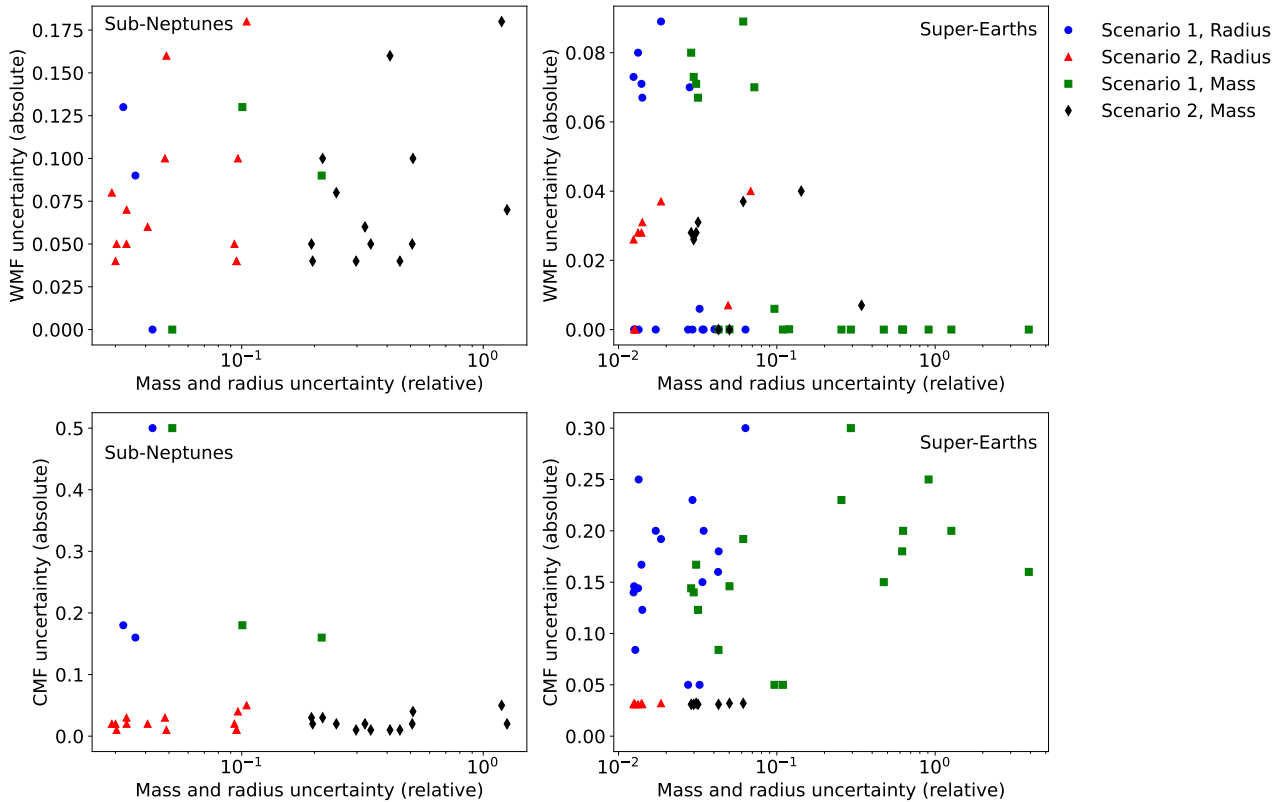


Figure 6.17.: WMF and CMF absolute uncertainties as a function of mass and radius relative errors. The panels on the right corresponds to sub-Neptunes ($R > 1.9 R_{\oplus}$), while the left-side panels display uncertainties for super-Earths ($R < 1.9 R_{\oplus}$).

7. Discussion

In this Chapter, I discuss the caveats and limitations of the interior-atmosphere model I developed in this thesis.

The first of the limitations relates to the fixed composition we consider of the mantle. When a mantle forms, its chemical components are fractionated into different minerals. The EOSs of the chemical components (Mg_2SiO_4 , SiO_2 , FeO) are used to calculate the final combined EOS of the mineral mixture following the additive volume law (Nettelmann et al. 2008), which means that the contribution of each chemical component to the mixture EOS depends on its concentration, which is different for each mineral. This fractionation is usually calculated by assuming chemical equilibrium and Gibbs minimisation (Dorn and Lichtenberg 2021). Therefore, we can conclude that mantles with different mineralogies present different EOS for their mantle mixtures. In our interior model, we have considered that the mineralogy is similar to that of Earth (Brugger et al. 2016, 2017). However, the evolution of a water-rich planet on long time scales can produce changes in the minerals that compose the mantle. Spaargaren et al. (2020) find that the compositional profile of the final solid mantle after magma ocean crystallisation depends mainly on the initial bulk Mg/Si. However, Lichtenberg (2021) develops an interior evolution model that includes a detailed description of redox reactions. These reactions can change the initial primitive composition of the mantle. The efficiency of these reactions, and therefore the extent with which they change the primordial composition of the mantle depends on the accretional heat. Consequently, sub-Neptunes may present more redox reactions that change the thermal evolution and composition of the mantle compared to the primordial one. The EOS of the different chemical components might not change, but the combined EOS of the final mixture (the minerals) is different. In addition, the solubility (ability of rock to mix with water) affects mantle EOSs as well (Dorn and Lichtenberg 2021). Solubility depends on the thermal evolution (Krissansen-Totton and Fortney 2022), so this parameter is also a source of variation in the mantle EOS due to planetary evolution. There is an available implementation of the mantle in our interior model (Brugger 2018) that uses `Perple_X` (Connolly 2005, 2009). Given a mantle mineralogy, `Perple_X` computes the Grüneisen parameter and density. Therefore, coupling this implementation of the interior model with the atmosphere model has the potential to provide an interior-atmosphere model that reflects the changes in the mass-radius diagram due to different mantle mineralogies. An additional interior evolution model would be necessary to provide the final mineral composition as a result of the planet's thermal evolution, such as the models presented in Krissansen-Totton and Fortney (2022) and Lichtenberg (2021).

Other chemical modifications in the mantle and core composition can be taken into account as well. Examples on how this is Dorn et al. (2015), where the effect of light elements in the mantle (Na, Ca, Al) on the mantle's EOS is obtained with the NCFMAS model chemical system. Similarly, Unterborn et al. (2016) use `BurnMan`, which is a self-consistent thermoelastic compression calculator. They find that not including light elements in the core can lead to an overestimation of the core's density of 20%, which can increase the total planetary density by 1 g/cc. An increase of 1 g/cc shifts downwards the mass-radius relationship of an Earth-like planet by $0.05 R_{\oplus}$. The inclusion of an upper mantle (different mantle phase transitions) can increase the planetary radius of an

Earth-like planet by $0.02 R_{\oplus}$. Modifications in mantle composition and mantle mineralogy have more secondary effects on the total radius compared to these two parameters ($< 0.02 R_{\oplus}$).

Likewise, a different chemical composition in the ices and liquid water may affect the thickness of a condensed water layer. The effect of incorporating salts in water decreases the freezing point of water. For a shallow ocean like the one on Earth, increasing the salinity would maintain a higher percentage of the hydrosphere in liquid instead of low pressure ice, compared to a pure water ocean layer. Since the density of liquid water is lower than that of ices, the inclusion of salty water in the interior model may yield a higher radius for the same planetary mass, and water and core mass fractions. The extent of this increase in radius for an Earth-like planet is yet to be explored, but it could be calculated by our interior structure model if the EOS of water was changed to that of seawater, such as that of [Schmidt et al. \(2018\)](#). For water-rich planets, not only the phase transitions between liquid water and high pressure (HP) ices may change, but also the density itself of the high pressure ices. Experimental works report both a change of the volume of the lattice structure in HP ices, which is only explored up to a temperature of 300 K ([Journaux 2022](#)). [Hernandez et al. \(2022\)](#) provide a state-of-the-art EOS for water with NaCl incorporation in HP ices, obtained by combining available experimental data and molecular dynamics simulations. The effects of salty water on habitability ([Olson et al. 2020](#); [Journaux 2022](#)) and climate ([Olson et al. 2022](#)) in exoplanets have been widely explored. However, the extent of its effect on the mass-radius relationships remains unknown. The EOS presented in [Hernandez et al. \(2022\)](#) could be incorporated in our interior structure model to investigate this effect.

The assumption of a H₂O or CO₂-dominated atmosphere is quite limited given the diversity in composition that exoplanet atmospheres can have. The modifications in the interior-atmosphere model to include other chemical compounds would be:

- Implementation of the EOS of these gases in the atmospheric model to calculate the atmospheric thickness. For H/He, the EOS is non-ideal ([Chabrier et al. 2019](#)), while other gases found in secondary atmospheres, such as N₂, could be assumed to be ideal. The thickness of the atmosphere for the same pressure-temperature (PT) profile should be lower than that of a pure water atmosphere, since adding these compounds increases the mean molecular weight. If H/He is included, then the mean molecular weight decreases, yielding a more extended atmosphere.
- Opacities for H/He, N₂ and other gases. These opacities would be line opacities (k-tables), collision-induced absorption (CIA) and Rayleigh scattering opacities. In comparison to H/He and N₂, CO₂ and H₂O are strong greenhouse gases, so including the former gases may lower the surface temperature under the same irradiation conditions, producing a colder PT profile. This lowers the altitude at which clouds form, making the cloud layer deeper, and enhancing its effect in emission and transmission spectra (flatten spectral lines).
- The inclusion of more diverse chemical compounds, such as metal hydrides, metal oxides or ion species (H⁻, Ca⁺, Fe⁺) may modify the atmospheric PT profile by producing an isothermal layer in the middle of the convective layer ([Malik et al. 2019](#)). This PT profile is warmer (by 500 K) in the near-surface region of the atmosphere, which may lead to more extended atmospheres, although isothermal radiative layers are expected to decrease the atmospheric thickness, so the two effects may cancel out. Silicate vapour produces a radiative layer near the surface, decreasing the thickness of the atmosphere compared to a convective atmosphere ([Misener and Schlichting 2022](#)). Temperature inversions in the upper atmospheres of hot sub-Neptunes may occur due to the absorption of short-wavelength stellar irradiation by TiO,

VO (seen in hot Jupiters), Fe and alkalis. For a hot Neptune of of an equilibrium temperature of 2100 K and a solar composition, [Moses et al. \(2022\)](#) estimate that the inversion starts at a pressure of 100 mbar. Therefore, the pressure range between 100 mbar and our transit radius pressure, 20 mbar, might be affected by inversions, producing a slightly shorter atmospheric thickness. Interior-atmosphere models with H/He dominated atmospheres are required to measure the extent of this effect on the total radius, and 2D and 3D models may be needed to simulate their effect on emission spectra. Temperature inversions are not expected in super-Earths ([Miller-Ricci and Fortney 2010](#); [Innes and Pierrehumbert 2022](#)).

8. Conclusion

In this thesis, we present a self-consistent model built to obtain the internal compositions and structures of low-mass planets given their observed mass and radius, irradiation conditions, and their host stellar abundances. I developed this model by starting with the interior model version of [Brugger et al. \(2017\)](#), who had a comprehensive description of a Fe-rich core, a silicate mantle, liquid and high-pressure ice phases, being only adapted for planets with an Earth-like irradiation. Many of the planets we discover are highly irradiated, which means they cannot sustain liquid or ice water on their surfaces. To make the model applicable to highly irradiated exoplanets, it required a thorough study of the equation of state (EOS) and Grüneisen parameter from different formulations. I concluded that the M19 formulation ([Mazevet et al. 2019](#)) is the most appropriate one. This formulation is consistent with experimental data in the supercritical regime of water. For higher temperatures and pressure in the plasma phase of water, it is based on quantum molecular simulations, complementing the experimental data to cover the conditions reached in planetary interiors. Additionally, I revised another version of the interior model for icy planets (irradiations below that of Earth), and improved the accuracy of the ice phase changes.

Consecutively, the complete model for supercritical planets has been assembled by coupling together our interior model and an atmospheric model within a MCMC Bayesian algorithm. The supercritical water layer in the interior model goes down to a pressure of 300 bar. Therefore, for pressures below 300 bar, it is necessary to use an atmospheric model. In addition, the atmospheric model is also needed to calculate the atmospheric thickness, which together with the bulk radius obtained by the interior model, constitute the total planetary radius. The atmospheric model is also essential to carry radiative transfer calculations to compute the surface temperature at which the atmosphere is in radiative-convective equilibrium. Radiative-convective equilibrium is established when the emitted radiation equals the absorbed radiation. Consequently, the atmospheric model calculates the outgoing longwave radiation and the Bond albedo to calculate these, respectively. The bulk radius, which is the radius from the center of the planet up to the top of its outermost layer, is the output of the interior model and a necessary input for the atmospheric model, while the surface temperature is an input to both models. This means that the assembly of the interior-atmosphere model is not straightforward. Therefore, I developed an algorithm to couple the interior and the atmosphere self-consistently, allowing us to obtain the total radius and surface temperature of a highly-irradiated planet given its mass, core mass fraction, and water mass fraction.

The initial interior-atmosphere coupling was done by using grid files of the atmospheric parameters, instead of running the interior and the atmosphere models simultaneously. Multiple input files were required to make the model flexible enough to apply it to surface pressures below 300 bar. These grids also had the disadvantage that no spectra were being generated, which would be useful to predict observations. Therefore, I further continued the development of the atmospheric model to run it at the same time as the interior model. I changed its EOS to an EOS valid at high temperatures ($T > 2500$ K), which is necessary to estimate the atmospheric thickness. In addition, we updated the atomic opacity data (k-tables), that enables the atmospheric model to be more flexible in the spectral and g-point binning. This flexibility in g-point selection permitted me to adapt the atmospheric model to the k-uncorrelated approximation. A flexible spectral binning

allowed me to produce a version of the k-correlated atmospheric model that produces emission spectra at higher resolution ($R \simeq 200$), which is useful to assess observability with atmospheric characterization facilities, such as JWST. The low-resolution, k-uncorrelated version of the atmospheric model ($R \simeq 10$), is used only for fast OLR and Bond albedo calculations within the MCMC framework.

We have a self-consistent interior-atmosphere forward model, that I optimized and included in a user-friendly Python interface. I also implemented it within a MCMC Bayesian algorithm to derive the posterior distribution functions of the non-observable variables, which are the CMF, WMF, and the atmospheric parameters of water-dominated and CO₂-rich atmospheres.

The model was used to perform the analysis of various planetary systems during the course of its development. We first applied it to the well-known planetary system TRAPPIST-1. Our results on this system can be summarized in three points:

- The most dense planets in the system have densities similar to that of Earth, suggesting that they have not undergone post-formation processes that could increase their Fe-to-Si ratio (collisions, mantle evaporation). We retrieve the Fe/Si mole ratio of each individual planet based on masses and radii, and define the system's common Fe/Si as the range where the individual 1σ intervals of the Fe/Si mole ratio overlap. TRAPPIST-1 common Fe/Si range is 0.45 - 0.97, which is also in agreement with the possible Fe/Si mole ratio of the host star. This Fe/Si range is equivalent to a common CMF = 0.23 - 0.40, which includes an Earth-like core mass.
- The water mass fraction presents a trend with semi-major axis (or incident stellar flux). This trend consists of an increase of water content as we move further away from the star for the inner planets (b to e), followed by the constant WMF of the outermost planets (f to h).
- TRAPPIST-1 d seems to be the only exception to the trend within the system, since it presents a slightly higher WMF than planet e in our initial analysis, where we considered that its hydrosphere is a pure condensed water layer. However, we find that TRAPPIST-1 c's density and incident flux are also compatible with a CO₂-dominated atmosphere in equilibrium, with a surface pressure of $\simeq 300$ bar. This suggests that water could be present in gaseous form in a secondary atmosphere, yielding a lower volatile mass fraction than TRAPPIST-1 e's WMF.

We carried a similar analysis on a sample of other multiplanetary systems to derive in a homogeneous manner their composition and further search for any compositional trends. The main findings of our homogeneous interior analysis are:

- Multiplanetary systems show a clear separation between the inner, dry planets; and the outer, volatile-rich planets.
- The system K2-138 presents the gradient with a plateau similar to the one we highlighted in the TRAPPIST-1 system. Kepler-102 could potentially present a constant WMF in the outer part of the system as well, given the uncertainties of the composition of its outermost planets.
- Other systems, such as Kepler-11, Kepler-80 and the inner planets in TOI-178, present only a gradient in WMF. We are able to explain exceptions to this trend within each system with Jeans atmospheric escape and/or XUV photoevaporation.

- Despite not including H/He envelopes in the current version of our interior-atmosphere model, we are able to identify planets that present a H/He-rich atmosphere, based on the difference between the observed and the retrieved planetary radius.
- Dry, high-CMF planets are always located in the innermost part of the system.

These results give insights into the formation and evolution of low-mass planets in multiplanetary systems. The gradient-plus-plateau trend could be shaped by XUV photoevaporation (or core-powered mass loss) for the inner planets gradient, while the constant WMF could be produced by the formation in the vicinity of the water ice line of the outermost planets, with later inward migration to their current position. In addition, planets with a CMF similar to that of Mercury could be formed in the vicinity of rocklines, or undergo collisions or mantle evaporation that strip away their silicate layer.

Our interior-atmosphere model can also be used to assess the observability of the atmospheres of rocky Earth-sized planets and super-Earths. We derive the expected emission spectra for TRAPPIST-1 c and 55 Cancri e to simulate observations with JWST. This required a prior modelling of the interior of 55 Cancri e to constrain the most likely planetary surface conditions. We find that:

- TRAPPIST-1 c could have a H₂O-dominated atmosphere of up to 25 bar of surface pressure, or no atmosphere at all. A CO₂-rich envelope would have higher surface pressures, and therefore it would be more massive than the water case. For $\lambda > 12.8 \mu\text{m}$, we might be able to distinguish between these two compositions by observing with JWST's photometric filters F1500, F1800 and F2100.
- For 55 Cancri e, a pure CO₂ atmosphere is not extended enough to explain its low planetary density. A massive envelope with more than 300 bar of surface pressure that contains water is necessary to fit its low density, since its molecular weight is lower than that of CO₂.
- A combination of observations with NIRCam and MIRI LRS of 55 Cancri e will allow us to detect water spectral lines for $\lambda > 3.5 \mu\text{m}$, where the noise level is low.

The derivation of compositional parameters, and the constraints they provide on planet formation mechanisms, as well as the assessment of the observability of atmospheres, are the main applications of the model. To achieve a proper coupling of the atmosphere and the interior, I have also explored the requirements that coupled interior and atmosphere models need to follow to obtain accurate and precise estimates on the composition of low-mass planets. We highlight that:

- The choice of the EOS and Grüneisen parameter formulations for supercritical water in interior models with a water envelope can produce differences of up to 20% in radius. Other formulations ([Duan and Zhang 2006](#)) tend to overestimate the radius compared to the others we considered in this thesis, namely M19 ([Mazevet et al. 2019](#)), and IAPWS95 ([Wagner and Pruß 2002](#)). This is especially the case for WMFs greater than 20%.
- Differences in opacity data in atmospheric models can produce differences of approximately $\approx 200 \text{ K}$ in the surface temperature at which the atmosphere is in equilibrium. This is due to the outgoing longwave radiation, which is more dependent on the opacity data than the Bond albedo. This difference in surface temperature causes a difference in total radius of up to 7% for a planet of $1.5 M_{\oplus}$ and WMF = 70%. The discrepancy in total radius decreases with increasing mass and decreasing volatile content.

- Using a constant step size when sampling the prior distribution in a MCMC scheme is not efficient in exploring the parameter space. This causes an underestimation of the uncertainties of the compositional parameters. Therefore, it is necessary to use an adaptive MCMC when performing retrieval with interior models, especially for planets whose compositional parameters can reach the maximum or minimum possible values. This is the case of rocky Earth-sized planets and super-Earths, whose WMFs are close to zero, but are nonetheless important to determine their surface pressure.
- The mass and radius retrieved by MCMC methods in interior modelling need to be compared with the observed values to check that they are similar. If this is not the case, it means that it is not possible to reproduce the observed density under the assumptions the model considers. I defined the metric $d_{obs-ret}$, which quantifies the difference between the observed and retrieved masses and radii to determine the likelihood of the presence of a water-rich envelope, as opposed to a H/He atmosphere or not atmosphere at all. For planets whose density can only be explained with a H/He envelope, I combined our interior model core and mantle with the H/He relations from [Zeng et al. \(2019\)](#) to estimate their volatile mass fractions.

We take advantage of the complete sample of planets that we have analysed with the interior-atmosphere model during its development. This sample is constituted by 46 exoplanets, including both super-Earths and sub-Neptunes. We summarize the statistical properties of the composition of these planets in the context of planet observation programs we are associated to:

- The radii of the planets in our sample span from 0.5 to 3.5 R_{\oplus} . This range covers the the Fulton gap, the super-Earth and the sub-Neptune peaks, being representative of the population of low-mass planets.
- We find two regimes in the WMF distribution. The super-Earth regime corresponds to planets with WMF < 10%, whereas the sub-Neptune regime is constituted by planets with WMFs > 20%, up to 70%. Furthermore, within the sub-Neptune sample, we distinguish two maxima in the water mass fraction distribution: one at WMF \approx 0.20 and another one at WMF \approx 0.60. The gap at WMF = 0.30 - 0.40 might indicate a transition between water-rich envelopes and H/He atmospheres.
- We found that planets with $M < 2.5 M_{\oplus}$ do not present WMFs greater than 10%. I show that this could be due to Jeans atmospheric escape preventing them from retaining massive envelopes.
- The CMF distribution of our sample shows two populations: Earth-like planets with CMFs = 0.2 - 0.4, and super-Mercuries with CMF > 0.80. When the core mass fraction is obtained taking into account the Fe/Si mole ratio from the host stellar abundances, its distribution is Gaussian, with a 1 σ confidence interval in agreement with previous studies ([Plotnykov and Valencia 2020](#)).

The sample also presents a wide range of relative uncertainties in mass and radius, allowing us to observe how the refinement of the precision in the observables can improve the determination of the composition. We find that the increase in the precision of the radius for sub-Neptunes improves the volatile mass fraction estimate, but has no effects on the core mass fraction, in agreement with previous work by ([Otegi et al. 2020](#)). Moreover, considering the Fe/Si mole ratio as input data helps

constrain better the core mass fraction for sub-Neptunes, but provides very similar estimates for the water mass fraction compared to when we only take into account mass and radius data. Therefore, the Fe/Si mole ratio does not break the degeneracy in composition for sub-Neptunes. In the case of super-Earths, the uncertainty of the CMF is reduced when the radius precision is improved and the Fe/Si mole ratio is included in the modelling.

Perspectives

The field of interior and atmosphere modelling faces the challenge of degeneracies in composition. The first degeneracy consists of the fact that different combinations of core mass fraction and water mass fraction can yield the same radius. It is widely adopted to use the host stellar abundances to constrain the Fe/Si mole ratio to break this degeneracy (Dorn et al. 2015; Brugger et al. 2017; Wang et al. 2022). Nonetheless, the Fe/Si derived from stellar abundances and that obtained from rocky planet densities are not exactly a 1:1 relationship (Schulze et al. 2021; Adibekyan et al. 2021), which could be due to different formation processes. For this reason, in this thesis we presented two scenarios in most of our planet analyses: scenario 1, where the mass and radius are the data for the retrieval framework; while in scenario 2, we consider the mass, radius and the Fe/Si mole ratio derived from the stellar abundances. The presentation of both scenarios gives us a complete picture of the possible core and water mass fractions.

This degeneracy could be broken by measuring the Love number. The rigidity of the solid material, as well as tidal forces, influence the shape of the planet, which might differ significantly from a perfect sphere. The deviation from a perfect spherical shape depends on the Love number, which is a proxy of the interior's density profile. The Love number has been measured only for hot Jupiters with the radial velocity method (Csizmadia et al. 2019) and transit photometry (Barros et al. 2022). The on-going JWST, and the future PLATO mission will be able to provide more precise data to infer the Love number of more exoplanets (Hellard et al. 2020; Nettelmann and Valencia 2021). The planet's interior density profile not only depends on the Love number, but also on whether the refractories and volatiles in the interior are mixed or stratified in two separate layers. Although modelling work has done progress in determining what evolution and planetary conditions facilitate mixing of ice and rocks (Vazan et al. 2022), measuring the Love number will contribute to break this second degeneracy in planetary interiors. In addition, the generation of magnetic fields by liquid iron cores could also be used to constrain the size of the planetary core, which would add more observational data to break the degeneracy between the core and the water layer. Current radio telescopes are able to detect magnetic activity emission from hot Jupiters (Lynch et al. 2018; Cendes et al. 2022), although irregular magnetic and cyclotron emission from smaller planets could be detected with future ground facilities (Driscoll and Olson 2011).

The interior composition of the core and mantle has a connection with the composition of the atmosphere via outgassing (Dorn et al. 2018; Spaargaren et al. 2020; Wang et al. 2022). Atmospheric characterization of low-mass planets seems to be the most accessible way to constrain the iron-to-silicate ratio and redox state, given the upcoming amount of data from JWST. Furthermore, these data will also help break a third degeneracy we find in planetary interiors, between the mass and the composition of the atmosphere. In other words, a secondary atmosphere (water, CO₂) and a H/He atmosphere can yield a similar atmospheric thickness with different atmospheric masses. If the composition of the atmosphere is known, an interior-atmosphere model, such as the one we developed in this thesis, could be used to constrain the surface pressure and temperature.

The atmospheric compositions of sub-Neptunes are proving to be a mixture of H/He, water and other compounds via observations and models (Madhusudhan et al. 2020; Bézard et al. 2020;

Guzmán-Mesa et al. 2022), while super-Earths can have more exotic atmospheric compositions, such as mineral atmospheres (Keles et al. 2022). Therefore, the scope of future work will be to include more gases in the atmospheric model, as well as the calculation of transmission spectra in addition to the existing implementation of emission and reflection spectra. This will constitute a forward model that could be used, together with Bayesian algorithms, to perform a simultaneous retrieval of the spectrum and mass and radius data. This type of retrieval, together with spectroscopic characterization data, would break the degeneracy between the composition of the envelope and the surface pressure. Photochemistry models have shown that the surface conditions also determine the abundances of specific atmospheric gases in sub-Neptunes that can be detected with JWST (Yu et al. 2021; Tsai et al. 2021; Hu et al. 2021). These equilibrium chemistry networks can be included in interior-atmospheric models to provide further constraints on the surface conditions of low-mass planets.

JWST has several targets of interest to which this model could be applied once data is acquired. These targets include the sub-Neptunes K2-18 b, for which more sensitive transmission spectra will be obtained (Hu and Damiano 2021; Madhusudhan et al. 2021); and GJ1412 b, which will be characterised by phase curves (Bean et al. 2021), in addition to its already existing HST transmission spectrum (Kreidberg et al. 2014). Other possible cases to which it could be applied are super-Earths 55 Cancri e (Hu et al. 2021), and Gl 486 b (Mansfield et al. 2021), and the Earth-sized TRAPPIST-1 c (Kreidberg et al. 2021). These will be the first cases of a sample of low-mass exoplanets with available atmospheric characterization data. A number of future missions will provide more of this type of data. Whenever 3D effects are negligible, or at least poorly constrained due to the lack of in-depth characterization of the atmosphere, one-dimensional interior-atmosphere models constitute fast and accurate tools to identify planets that are the most promising for future atmospheric observations. They can also help the interpretation of atmospheric observations at low resolution. In addition, they provide unique information on their internal composition, which is particularly valuable in helping us to better understand how these diverse, low-mass planets form. The model presented in this thesis could be used as a starting point to develop more detailed models necessary for upcoming missions, which include:

1. Ariel (Tinetti et al. 2018) will increase the number of planets with atmospheric characterization data by carrying out a survey. It will also obtain the phase curves of a large sample of exoplanets across a wide range of masses and radii, which will enable us to obtain atmospheric metallicity and chemistry constraints, as well as albedo measurements, which are necessary to determine the atmospheric energy balance (Charnay et al. 2022). The metallicity and C/O ratio have been obtained for Jupiter-sized planets (Giacobbe et al. 2021), which are parameters that contain information about the formation location and the dominant type of accretion (dust, pebble or planetesimal) (Khorshid et al. 2021; Mollière et al. 2022). An atmosphere characterization survey such as Ariel will provide data necessary to connect atmospheric compositions and planet formation for low-mass planets, similarly to what it is being done with gaseous giants.
2. The detection and characterization of temperate Earth-sized planets are a priority for PLATO (Rauer et al. 2014). PLATO will be able to detect temperate Earth-sized planets around solar-like stars and provide precise radius measurements to characterise their density, together with radial velocity follow-up with ESPRESSO, HARPS and HiRES (Udry et al. 2014). PLATO will be the first survey of habitable planets, which will supply a sample of targets for atmospheric characterization to search for biosignatures.
3. The ground-based Extremely Large Telescopes (ELTs) will acquire atmospheric characteriza-

tion data to search for biosignatures. Biosignatures are combinations of gaseous species at given abundances that are not expected in atmospheres in chemical equilibrium, produced by the action of life. The processes that could drive an atmosphere in a terrestrial planet out of chemical equilibrium could also be geological. It is therefore essential to model the disequilibrium chemistry produced by abiotic geological processes (i.e. volcanism, weathering), to avoid biosignature false positives (Truong and Lunine 2021). This search for biosignatures will be done with transmission spectroscopy, high-resolution spectroscopy and reflected-light imaging in the next decade (Snellen et al. 2013; Lopez-Morales et al. 2019).

4. In the next 20 to 30 years, the number of Earth-sized planets characterized by atmospheric data could be increased by future space missions Luvor (The LUVOIR Team 2019), HabEx (Gaudi et al. 2018) and LIFE (Konrad et al. 2021).

Simultaneously, an in-situ probe exploration of the Solar System ice giants is a priority in the next 30 to 40 years. The interiors and atmospheres of Uranus and Neptune are still poorly understood, as we only have data acquired remotely (Mousis et al. 2018; Fletcher et al. 2020). Such a mission is essential to break the degeneracies we have discussed above for sub-Neptunes, because our interior and atmosphere models for this class of planets are based on our knowledge of the Solar System ice giants, similar to Earth and Venus being the references for super-Earth models.

A. First-author contributions

A.1. Characterisation of the hydrospheres of TRAPPIST-1 planets ([Acuña et al. 2021](#))

L. Acuña, M. Deleuil, O. Mousis, E. Marcq, M. Levesque, and A. Aguichine

The following is the first first-author publication as a result of the work presented in this thesis. It describes the implementation of different water phases in two separate versions of the interior model, one which includes the supercritical water layer (see Sect. 2.2), and another one that includes ice phases for cold planets. Furthermore, the interior-atmosphere algorithm is presented (see Chapter 3) to couple our interior model to the atmospheric model developed by [Marcq et al. \(2017\)](#) and [Pluriel et al. \(2019\)](#). Additionally, the non-adaptive MCMC (see Sect. 5.1) is also explained and used to infer the compositional parameters of the planets in the TRAPPIST-1 system. I present the final results, and discuss on the composition scenarios and possible formation pathways of the TRAPPIST-1 system in Sect. 6.1.

This work was also presented in the PLATO mission conference 2020 as a contributed talk, in December 2020 (virtual).

Characterisation of the hydrospheres of TRAPPIST-1 planets

Lorena Acuña¹, Magali Deleuil¹, Olivier Mousis¹, Emmanuel Marcq², Maëva Levesque¹, and Artyom Agüichine¹

¹ Aix-Marseille Univ., CNRS, CNES, LAM, Marseille, France
e-mail: lorena.acuna@lam.fr

² LATMOS/IPSL, UVSQ Université Paris-Saclay, Sorbonne Université, CNRS, Guyancourt, France

Received 10 November 2020 / Accepted 19 January 2021

ABSTRACT

Context. Planetary mass and radius data suggest that low-mass exoplanets show a wide variety of densities. This includes sub-Neptunes, whose low densities can be explained with the presence of a volatile-rich layer. Water is one of the most abundant volatiles, which can be in the form of different phases depending on the planetary surface conditions. To constrain their composition and interior structure, models must be developed that accurately calculate the properties of water at its different phases.

Aims. We present an interior structure model that includes a multiphase water layer with steam, supercritical, and condensed phases. We derive the constraints for planetary compositional parameters and their uncertainties, focusing on the multi-planetary system TRAPPIST-1, which presents both warm and temperate planets.

Methods. We use a 1D steam atmosphere in radiative-convective equilibrium with an interior whose water layer is in supercritical phase self-consistently. For temperate surface conditions, we implement liquid and ice Ih to ice VII phases in the hydrosphere. We adopt a Markov chain Monte Carlo inversion scheme to derive the probability distributions of core and water compositional parameters.

Results. We refine the composition of all planets and derive atmospheric parameters for planets ‘b’ and ‘c’. The latter would be in a post-runaway greenhouse state and could be extended enough to be probed by space missions such as JWST. Planets ‘d’ to ‘h’ present condensed ice phases, with maximum water mass fractions below 20%.

Conclusions. The derived amounts of water for TRAPPIST-1 planets show a general increase with semi-major axis, with the exception of planet d. This deviation from the trend could be due to formation mechanisms, such as migration and an enrichment of water in the region where planet d formed, or an extended CO₂-rich atmosphere.

Key words. planets and satellites: interiors – planets and satellites: composition – planets and satellites: atmospheres – planets and satellites: individual: TRAPPIST-1 – methods: statistical – methods: numerical

1. Introduction

Ongoing space missions such as CHEOPS (Benz 2017) and TESS (Ricker et al. 2015), and their follow-up with ground-based radial velocity telescopes, are confirming the existence of low-mass exoplanets with a wide range of densities. These densities range from the values typically inferred for the Earth or Mercury to those measured in Uranus and Neptune. The exoplanets in the former class are mainly composed of a Fe-rich core and a silicate mantle, while the latter class has a layer that is rich in volatiles. Water is the most abundant and least dense volatile after H and He (Forget & Leconte 2014), which makes it a likely species to constitute the volatile reservoir in these planets. Several studies have investigated the interior structure and composition of water-rich planets (Sotin et al. 2007; Seager et al. 2007; Dorn et al. 2015; Zeng et al. 2019), but focused mainly on its condensed phases. Nonetheless, many sub-Neptunes are close to their host star and receive enough irradiation to trigger a runaway greenhouse state in which water is present as steam. In some cases, the high pressure and temperature conditions can render the hydrosphere supercritical and plasma, or even lead to superionic phases (Mazevet et al. 2019; French et al. 2016). Therefore, it is crucial to include the modelling of all possible phases of water in order to provide an accurate description of its presence on the planetary surface. Moreover, the surface conditions are determined by the greenhouse effect caused by atmospheric gases, making the modelling of radiative-convective equilibrium

in atmospheres a key parameter to determine the phase in which water could be present on the surface. Most interior structure models represent the planetary atmosphere as a gas layer with a simplified isothermal temperature profile (Dorn et al. 2018, 2017b), which is very different from the temperature profile in the convective deep layers of thick atmospheres (Marcq 2012).

Multi-planetary systems are unique environments that present both planets that can hold condensed phases and highly irradiated planets with steam atmospheres. In this study, we develop a planet interior model suitable for the different conditions at which water can be found in low-mass planets. Our implementation includes a supercritical water layer, introduced in Mousis et al. (2020), coupled with a 1D radiative-convective atmosphere model (Marcq 2012; Marcq et al. 2017; Pluriel et al. 2019) to calculate the total radius of the highly irradiated planets with water self-consistently. Furthermore, for temperate planets, we have updated the interior model presented in Brugger et al. (2016, 2017) to include ice phases Ih, II, III, V, and VI. We introduce these models in a Markov chain Monte Carlo (MCMC) Bayesian analysis scheme adapted from Dorn et al. (2015). This allows us to derive the water mass fraction (WMF) and core mass fraction (CMF) that reproduce the observed radius, mass, and stellar composition measurements.

We use this model to explore the possible water content of the TRAPPIST-1 system, an ultra-cool M dwarf that hosts seven low-mass planets in close-in orbits. Three of these planets are located in the habitable zone (Grimm et al. 2018), meaning that

they can hold liquid water or ice Ih on their surfaces. Although all planets in TRAPPIST-1 system have masses and radii that are characteristic of rocky planets, their differences in density indicate that each planet has a different volatile content. This makes this planetary system ideal for testing planet interior, atmospheric structure, and formation scenarios.

In Sect. 2 we describe the complete interior structure model, including the new updates for the supercritical and ice phases, the coupling between the interior and the atmosphere for steam and supercritical planets, and the MCMC Bayesian algorithm. The parameters for the TRAPPIST-1 planets used in this study are summarised in Sect. 3, including mass, radius, and Fe/Si molar ratio. The results of our analysis of the hydrospheres of TRAPPIST-1 planets are described in Sect. 4. We compare our results with those of previous works and discuss the implications of our water estimates for planet formation in Sect. 5. We finally provide conclusions in Sect. 6.

2. Planetary structure model

For consistency, we recall the main principles of the interior structure model. The basis of our model is explained in Brugger et al. (2016, 2017). The 1D interior structure model takes as input the mass and the composition of the planet, which are parameterised by the CMF and WMF. The structure of the planet is stratified in three layers: a core, a mantle, and a hydrosphere. The pressure, temperature, gravity acceleration, and density are computed at each point of the 1D spatial grid along the radius of the planet. The pressure, $P(r)$, is obtained by integrating the hydrostatic equilibrium (Eq. (1)); the gravitational acceleration, $g(r)$, by solving Gauss's theorem (Eq. (2)); the temperature, $T(r)$, with the adiabatic gradient (Eq. (3)); and the density, $\rho(r)$, with the equation of state (EOS). In Eq. (2), m is the mass at radius r , G is the gravitational constant, and γ and ϕ are the Gruneisen and the seismic parameters, respectively. Their formal macroscopic definitions are shown in Eq. (4), where E is the internal energy and V is the volume. The Gruneisen parameter is a thermodynamic parameter that describes the dependence of the vibrational properties of a crystal with the size of its lattice. It relates the temperature in a crystalline structure to the density, which is calculated by the EOS. The seismic parameter defines how seismic waves propagate inside a material. It is related to the slope of the EOS at constant temperature (Brugger et al. 2017; Sotin et al. 2007).

$$\frac{dP}{dr} = -\rho g, \quad (1)$$

$$\frac{dg}{dr} = 4\pi G \rho - \frac{2Gm}{r^3}, \quad (2)$$

$$\frac{dT}{dr} = -g \frac{\gamma T}{\phi}, \quad (3)$$

$$\begin{cases} \phi = \frac{dP}{d\rho} \\ \gamma = V \left(\frac{dP}{dE} \right)_V \end{cases}. \quad (4)$$

The boundary conditions are the temperature and pressure at the surface, and the gravitational acceleration at the centre of the planet. The value of the latter is zero. The total mass of the planet is calculated with Eq. (5), which is derived from the conservation of mass (Brugger et al. 2017; Sotin et al. 2007). Once the total

input mass of the planet is reached and the boundary conditions are fulfilled, the model has converged.

$$\frac{dm}{dr} = 4\pi r^2 \rho. \quad (5)$$

Depending on the surface conditions, the hydrosphere can be present in supercritical, liquid, or ice states. For each of these phases of water, we use a different EOS and Gruneisen parameter to compute their P-T profiles and density accurately. In Sect. 2.1 we describe the updates to the supercritical water layer with respect to the model depicted in Mousis et al. (2020), while in Sect. 2.2 we present the implementation of the hydrosphere in ice phases. Finally, the coupling between the atmosphere and the interior model with planets whose hydrosphere is in steam or supercritical phases is explained in Sect. 2.3, followed by the description of the MCMC algorithm in Sect. 2.4.

2.1. Supercritical water

If the planet is close enough to its host star, the upper layer of the hydrosphere corresponds to a hot steam atmosphere, whose temperature at the base is determined by the radiative-convective balance calculated by the atmosphere model (Marcq 2012; Marcq et al. 2017). When the pressure and temperature at the surface, which is defined as the base of the hydrosphere layer, are above the critical point of water, we include a supercritical water layer extending from the base of the hydrosphere to a height corresponding to the phase change to steam (Mousis et al. 2020). We updated the EOS for this layer to the EOS introduced by Mazevet et al. (2019), which is a fit to the experimental data provided by the International Association for the Properties of Water and Steam (IAPWS; Wagner & Pruß 2002) for the supercritical regime, and quantum molecular dynamics (QMD) simulations data for plasma and superionic water (French et al. 2009). The IAPWS experimental data span a temperature range of 251.2 to 1273 K and of 611.7 to 10^9 Pa in pressure, while their EOS can be extrapolated up to 5000 K in temperature and 10^{11} Pa in pressure (Wagner & Pruß 2002). The validity range of the EOS presented in Mazevet et al. (2019) includes that of the IAPWS plus the region in which the QMD simulations are applicable, which corresponds to a temperature from 1000 K to 10^5 K and densities in the $1-10^2$ g cm $^{-3}$ range. These densities are reached at high pressures, that is, in the 10^9-10^{12} Pa range. Following Eq. (3), the adiabatic gradient of the temperature is specified by the Gruneisen and the seismic parameters. These are dependent on the derivatives of the pressure with respect to the density and the internal energy (Eq. (4)). We make use of the specific internal energy and density provided by Mazevet et al. (2019) to calculate them.

2.2. Ice phases

We extended the hydrosphere in Brugger et al. (2016, 2017) with liquid and high-pressure ice VII by adding five more condensed phases: ice Ih, II, III, V, and VI. An EOS for ice Ih was developed by Feistel & Wagner (2006) with minimisation of the Gibbs potential function from the fit of experimental data. It covers the whole pressure and temperature range in which water forms ice Ih.

Fei et al. (1993) proposed a formalism to derive the EOSs of ices II, III, and V. These EOSs take the form $V = V(P, T)$, which can be found by integrating the following differential equation (Tchijov et al. 2004):

$$\frac{dV}{V} = \alpha dT - \beta dP, \quad (6)$$

Table 1. EOS and reference thermal parameters for ices Ih, II, III, V, and VI.

Phase	ρ_0 [kg m ⁻³]	T_0 [K]	K_{T_0} [GPa]	K'_{T_0}	$C_p(T_0)$ [J kg ⁻¹ K ⁻¹]	α_0 [10 ⁻⁶ K ⁻¹]	References
Ih	921.0	248.15	9.50	5.3	1913.00	147	1, 8
II	1169.8	237.65	14.39	6.0	2200.00	350	1, 2, 7
III	1139.0	237.65	8.50	5.7	2485.55	405	3, 4, 5, 7
V	1235.0	237.65	13.30	5.2	2496.63	233	1, 4, 5, 7
VI	1270.0	300.00	14.05	4.0	2590.00	146	4, 6, 7

Notes. This includes the reference values for the density ρ_0 , the temperature T_0 , the bulk modulus K_{T_0} and its derivative K'_{T_0} , the heat capacity $C_p(T_0)$, and the thermal expansion coefficient α_0 .

References. (1) Gagnon et al. (1990); (2) Bález & Clancy (1995); (3) Tulk et al. (1997); (4) Tchijov et al. (2004); (5) Shaw (1986); (6) Bezacier et al. (2014); (7) Choukroun & Grasset (2010); (8) Feistel & Wagner (2006).

where α is the thermal expansion coefficient and β the isothermal compressibility coefficient. If the relationship between the specific volume, V , and the pressure, P , at a constant temperature $T = T_0$ is determined, Eq. (6) can be integrated as:

$$V(P, T) = V(P, T_0) \exp\left(\int_{T_0}^T \alpha(P, T') dT'\right). \quad (7)$$

Fei et al. (1993) proposed the following expression for the thermal expansion coefficient α :

$$\alpha(P, T) = \alpha(P_0, T) \frac{\left(1 + \frac{K'_{T_0}}{K_{T_0}} P\right)^{-\eta}}{\left(1 + \frac{K'_{T_0}}{K_{T_0}} P_0\right)^{-\eta}} = -\frac{1}{\rho} \frac{d\rho(T)}{dT} \frac{\left(1 + \frac{K'_{T_0}}{K_{T_0}} P\right)^{-\eta}}{\left(1 + \frac{K'_{T_0}}{K_{T_0}} P_0\right)^{-\eta}}, \quad (8)$$

where η is an adjustable parameter estimated from the fitting of experimental data. Its value is 1.0 for ice II and ice III (Leon et al. 2002) and 7.86 for ice V (Shaw 1986). ρ is the density, $\alpha(P_0, T)$ is the coefficient of thermal expansion at a reference pressure P_0 , K_{T_0} is the isothermal bulk modulus at the reference temperature T_0 , and K'_{T_0} is the first derivative of the isothermal bulk modulus at the reference temperature. Hence, by substituting Eq. (8) in Eq. (7) and integrating, we obtain the following EOS for high-pressure ice:

$$V(P, T) = V(P, T_0) \exp\left[\ln\left(\frac{\rho(T_0)}{\rho(T)}\right) \frac{\left(1 + \frac{K'_{T_0}}{K_{T_0}} P\right)^{-\eta}}{\left(1 + \frac{K'_{T_0}}{K_{T_0}} P_0\right)^{-\eta}}\right]. \quad (9)$$

The final expression (Eq. (9)) requires knowledge of the variation of the specific volume, $V(P, T_0)$, with pressure at the reference temperature T_0 . Moreover, the variation of the density with temperature, $\rho(T)$, and the bulk modulus with its derivative at the reference temperature, K_{T_0} and K'_{T_0} , must also be provided. In Table 1 we specify the data and references to obtain these parameters for each ice phase.

In the case of ice VI, we adopt the second-order Birch-Murnaghan (BM2) formulation, which is

$$P = \frac{3}{2} K_{T_0} \left[\left(\frac{\rho}{\rho_0}\right)^{\frac{7}{3}} - \left(\frac{\rho}{\rho_0}\right)^{\frac{5}{3}} \right], \quad (10)$$

where ρ_0 is the reference density for ice VI. We also introduce a thermal correction to the density because the pressure also depends on the temperature:

$$\rho(T) = \rho_0 \exp(\alpha_0 (T - T_0)), \quad (11)$$

where α_0 is the reference coefficient of thermal expansion. Interfaces between liquid and ice layers are established by phase transition functions from Dunaeva et al. (2010).

2.3. Interior-atmosphere coupling

We use a 1D atmosphere model designed to compute radiative transfer and pressure–temperature (P, T) profiles for water and CO₂ atmospheres (Marcq 2012; Marcq et al. 2017). The formation of water clouds is considered in the computation of the albedo. The atmosphere is in radiative equilibrium, and presents a composition of 99% water and 1% CO₂. The density of steam is obtained using a non-ideal EOS (Haar et al. 1984).

If the surface pressure is below 300 bar, the atmosphere and the interior are coupled at the atmosphere–mantle boundary and water does not reach the supercritical regime. However, if the surface pressure is greater than 300 bar, the atmosphere and the interior are coupled at this pressure level and a layer of water in supercritical phase forms between the atmosphere and the mantle. The pressure level at 300 bar is close enough to the critical point of water at 220 bar to avoid the atmosphere model to take over pressures and temperatures where the temperature profile is adiabatic.

The pressure at the top of the atmosphere is set to 20 mbar, which corresponds to the observable transiting radius (Mousis et al. 2020; Grimm et al. 2018). We denote the radius and mass from the centre of the planet to this pressure level the total radius and mass, R_{total} and M_{total} , respectively. We also define the radius and the mass that comprise the core, mantle, and supercritical layers as the bulk radius and mass, R_{bulk} and M_{bulk} , respectively. The atmosphere model provides the outgoing longwave radiation (OLR), albedo, thickness, and mass of the atmosphere as a function of the bulk mass and radius, and the surface temperature. If the atmosphere of the planet is in radiative equilibrium, the OLR is equal to the radiation the planet absorbs from its host star, F_{abs} . The OLR depends on the effective temperature because $\text{OLR} = \sigma T_{\text{eff}}^4$, where σ corresponds to the Stefan-Boltzmann constant. To calculate the absorbed radiation F_{abs} , we first compute the equilibrium temperature, which is

$$T_{\text{eq}} = (1 - A_B)^{0.25} \left(\frac{R_\star}{2a_d}\right)^{0.5} T_\star, \quad (12)$$

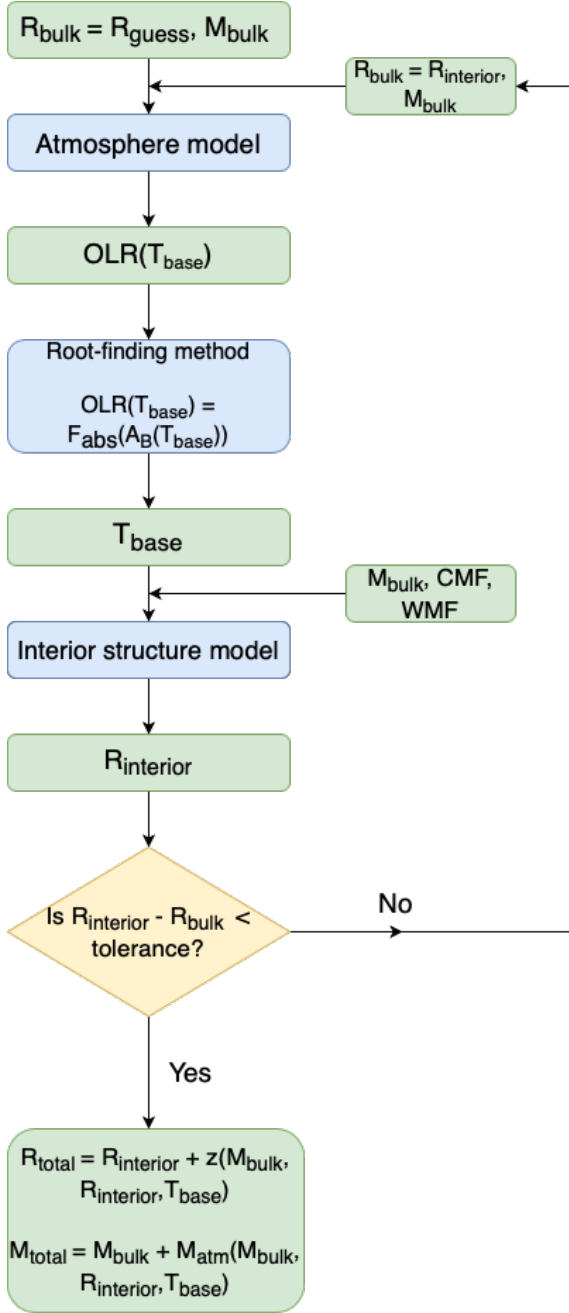


Fig. 1. Structural diagram of the coupling between the interior structure model and the atmosphere model. T_{base} is the temperature at the bottom of the steam atmosphere in radiative-convective equilibrium. z and M_{atm} denote the atmospheric thickness and mass, respectively. R_{bulk} and M_{bulk} correspond to the planet bulk radius and mass, respectively. R_{guess} refers to the initial guess of the bulk radius, while R_{interior} is the output bulk radius of the interior structure model in each iteration.

where A_B is the planetary albedo, R_* and T_* are the radius and effective temperature of the host star, respectively, and a_d is the semi-major axis of the planet. The absorbed radiation is then calculated as

$$F_{\text{abs}} = \sigma T_{\text{eq}}^4. \quad (13)$$

Figure 1 shows the algorithm we implemented to couple the planetary interior and the atmosphere. The interior structure model calculates the radius from the centre of the planet to the base of the steam atmosphere. For a fixed set of bulk mass and radius, the OLR depends on the surface temperature. Consequently, the surface temperature at which the OLR is equal to the absorbed radiation corresponds to the surface temperature that yields radiative equilibrium in the atmosphere. This is estimated with a root-finding method. As the bulk radius is an output of the interior model (R_{interior}) and an input of the atmosphere model, we first need to calculate the surface temperature for a certain mass and composition with an initial guess bulk radius. This surface temperature is then the input for the interior model, which provides the bulk radius. With this bulk radius, we can generate a new value of the surface temperature. This scheme is repeated until the bulk radius converges to a constant value, to which we add the thickness of the atmosphere, z , to get the total radius of the planet R_{total} . The total mass M_{total} is obtained as the sum of the bulk mass M_{bulk} plus the atmospheric mass M_{atm} . The tolerance used to determine if the bulk radius has achieved convergence is 2% of the bulk radius in the previous iteration. This is approximately $0.02 R_{\oplus}$ for an Earth-sized planet.

2.4. MCMC Bayesian analysis

We adapted the MCMC Bayesian analysis algorithm described in Dorn et al. (2015) to our coupled interior and atmosphere model. The input model parameters are the bulk planetary mass M_{bulk} , the CMF, and the WMF. Therefore, $\mathbf{m} = \{M_{\text{bulk}}, \text{CMF}, \text{WMF}\}$ following the notation in Dorn et al. (2015). Depending on the planetary system and their available data, we can have observational measurements of the total planetary mass and radius and the stellar composition, or only the total planetary mass and radius. The available data in the former case are denoted as $\mathbf{d} = \{M_{\text{obs}}, R_{\text{obs}}, \text{Fe/Si}_{\text{obs}}\}$, while the data in the latter case are represented as $\mathbf{d} = \{M_{\text{obs}}, R_{\text{obs}}\}$. The uncertainties on the measurements are $\sigma(M_{\text{obs}})$, $\sigma(R_{\text{obs}})$, and $\sigma(\text{Fe/Si}_{\text{obs}})$.

The CMF and WMF prior distributions are uniform distributions between 0 and a maximum limit. This maximum limit is 75% for the CMF, which is derived from the maximum estimated Fe/Si ratio of the proto-Sun (Lodders et al. 2009). With this limit on the CMF, we are assuming that the exoplanets have not been exposed to events during or after their formation that could have stripped away all of their mantle, such as mantle evaporation or giant impacts. In addition, the maximum WMF is set to 80%, which is the average water proportion found in comets in the Solar System (McKay et al. 2019). The prior distribution for the mass is a Gaussian distribution whose mean and standard deviations correspond to the central value and uncertainties of the observations.

The MCMC scheme first starts by randomly drawing a value for each of the model parameters from its respective prior distributions. This set of values is denoted $\mathbf{m}_1 = \{M_{\text{bulk},1}, \text{CMF}_1, \text{WMF}_1\}$. The index $i = 1$ corresponds to the first proposed set of input values within the first chain, $n = 1$. The model calculates the total mass and radius and the theoretical Fe/Si mole ratio, which are the set of output parameters $g(\mathbf{m}_1) = \{R_1, M_1, \text{Fe/Si}_1\}$. The likelihood of a set of model parameters is then calculated via the following relationship (Dorn et al. 2015):

$$L(\mathbf{m}_i | \mathbf{d}) = C \exp \left(-\frac{1}{2} \left[\left(\frac{R_i - R_{\text{obs}}}{\sigma(R_{\text{obs}})} \right)^2 + \left(\frac{M_i - M_{\text{obs}}}{\sigma(M_{\text{obs}})} \right)^2 + \left(\frac{\text{Fe/Si}_i - \text{Fe/Si}_{\text{obs}}}{\sigma(\text{Fe/Si}_{\text{obs}})} \right)^2 \right] \right). \quad (14)$$

where the normalisation constant of the likelihood function C is defined as:

$$C = \frac{1}{(2\pi)^{3/2} [\sigma^2(M_{\text{obs}}) \cdot \sigma^2(R_{\text{obs}}) \cdot \sigma^2(\text{Fe/Si}_{\text{obs}})]^{1/2}}. \quad (15)$$

When the Fe/Si mole ratio is not available as data, the square residual term of the Fe/Si mole ratio is removed from Eq. (14), as is its squared uncertainty in Eq. (15).

Subsequently we draw a new set of input parameters, $\mathbf{m}_2 = \{M_{\text{bulk},2}, \text{CMF}_2, \text{WMF}_2\}$, from the prior distributions within the same chain, n . We assure that the absolute difference between the values for $i = 1$ and $i = 2$ is lower than a fixed step, which is the maximum size of the perturbation. This guarantees that the new state \mathbf{m}_2 is uniformly bounded and centred around the old state, \mathbf{m}_1 . The maximum perturbation size is selected so that the acceptance rate of the MCMC, which is defined as the ratio between the number of models that are accepted and the number of proposed models, is above 20%. After \mathbf{m}_2 is chosen, the forward model calculates its corresponding output parameters and obtains their likelihood $L(\mathbf{m}_2 | \mathbf{d})$, as shown in Eq. (14). The acceptance probability is estimated with the log-likelihoods $l(\mathbf{m} | \mathbf{d}) = \log(L(\mathbf{m} | \mathbf{d}))$ as:

$$P_{\text{accept}} = \min \left\{ 1, e^{l(\mathbf{m}_2 | \mathbf{d}) - l(\mathbf{m}_1 | \mathbf{d})} \right\}. \quad (16)$$

If P_{accept} is greater than a number drawn from a uniform distribution between 0 and 1, \mathbf{m}_2 is accepted and the chain moves to the state characterised by \mathbf{m}_2 , starting the next chain $n + 1$. Otherwise, the chain remains in the state of \mathbf{m}_1 and a different set of model parameters is proposed as \mathbf{m}_3 . To ensure that the posterior distributions converge and that the whole parameter space is explored, we run 10^4 chains. In other words, with acceptance rates between 0.2 and 0.6, the MCMC proposes between 1.6 and 5×10^4 sets of model inputs.

3. System parameters of TRAPPIST-1

Agol et al. (2021) performed an analysis of transit timing variations (TTVs) that includes all transit data from *Spitzer* since the discovery of the system. We adopt these data for the mass, radius, and semi-major axis in our interior structure analysis (Table 2).

Data are not available regarding the chemical composition of TRAPPIST-1. However, the Fe/Si abundance ratio can be estimated assuming that TRAPPIST-1 presents a similar chemical composition to that of other stars of the same metallicity, age, and stellar population. As proposed by Unterborn et al. (2018), we select a sample of stars from the Hypatia Catalogue (Hinkel et al. 2014, 2016, 2017). We choose the set of stars by constraining the C/O mole ratio to be less than 0.8, and the stellar metallicity between -0.04 and 0.12 , as this is the metallicity range calculated for TRAPPIST-1 by Gillon et al. (2017). We discard thick disc stars because TRAPPIST-1 is likely a thin disc star. Our best-fit Gaussian to the distribution of the Fe/Si mole ratio shows a mean of 0.76 and a standard deviation of 0.12 . As this Fe/Si value is an estimate based on the chemical composition of a sample of stars that belong to the same stellar population as TRAPPIST-1, we present two scenarios for each planet. In scenario 1, the only available data are the planetary mass and radius, while scenario 2 includes the estimated stellar Fe/Si mole ratio to constrain the bulk composition.

For temperate planets that cannot have a steam atmosphere, we set the surface temperature in our interior model to their equilibrium temperatures assuming an albedo zero (Table 2).

Table 2. Masses, radii, and semi-major axis for all planets in TRAPPIST-1 (Agol et al. 2021).

Planet	$M [M_{\oplus}]$	$R [R_{\oplus}]$	$a_d [10^{-2} \text{ AU}]$	$T_{\text{eq}} [\text{K}]$
b	1.374 ± 0.069	$1.116^{+0.014}_{-0.012}$	1.154	398
c	1.308 ± 0.056	$1.097^{+0.014}_{-0.012}$	1.580	340
d	0.388 ± 0.012	$0.788^{+0.011}_{-0.010}$	2.227	286
e	0.692 ± 0.022	$0.920^{+0.013}_{-0.012}$	2.925	250
f	1.039 ± 0.031	$1.045^{+0.013}_{-0.012}$	3.849	218
g	1.321 ± 0.038	$1.129^{+0.015}_{-0.013}$	4.683	197
h	0.326 ± 0.020	0.775 ± 0.014	6.189	172

Notes. Equilibrium temperatures are calculated assuming a null albedo, with the stellar effective temperature, stellar radius, and semi-major axis provided by Agol et al. (2021).

Although surface temperatures for thin atmospheres are lower than that obtained with this assumption, the dependence of the bulk radius on surface temperature for planets with condensed water is low. For example, if we assume a pure-water planet of $1 M_{\oplus}$ with a surface pressure of 1 bar, the increase in radius due to a change of surface temperature from 100 to 360 K is $0.002 R_{\oplus}$, which is less than 0.2% of the total radius, which is ten times less than our convergence criterion. Additionally, the atmospheres of TRAPPIST-1 planets in the habitable zone and farther are significantly thinner than those of the highly irradiated planets. Lincowski et al. (2018) estimated thicknesses of approximately 80 km for temperate planets in TRAPPIST-1, which is negligible compared to their total radius. Therefore, we only calculate the atmospheric parameters (OLR, surface temperature, albedo, and thickness of the atmosphere) for planets that present their hydrospheres in steam phase.

4. Characterisation of hydrospheres

4.1. CMF and WMF posterior distributions

Tables 3 and 4 show the retrieved parameters, including the total planetary mass and radius, and the Fe/Si mole ratio. In both scenarios, we retrieve the mass and radius within the 1σ -confidence interval of the measurements for all planets. In scenario 1, where only the mass and radius data are considered, we retrieve Fe/Si mole ratios without any assumptions on the chemical composition of the host star. Although the uncertainties on these estimates are more than 50% in some cases, we can estimate a common Fe/Si mole ratio for the planetary system. This common Fe/Si range is determined by the overlap of the 1σ confidence intervals of all planets, which corresponds to Fe/Si = 0.45 – 0.97 . This interval is compatible with the Fe/Si mole ratio of 0.76 ± 0.12 proposed by Unterborn et al. (2018). This overlap can also be seen in Fig. 2, which presents the 1σ -confidence regions derived from the 2D marginalised posterior distributions of the CMF and WMF. The minimum value of the common CMF is determined by the lower limit of the confidence region of planet g, which is approximately 0.23, whereas the common maximum CMF value corresponds to the upper limit of planets b and c, which is 0.4. This is partially in agreement with the CMF obtained in scenario 2, where we assume the Fe/Si mole ratio proposed by Unterborn et al. (2018), which is found to be

Table 3. Output parameters retrieved by the MCMC method for all TRAPPIST-1 planets.

Planet	$M_{\text{ret}} [M_{\oplus}]$	$R_{\text{ret}} [R_{\oplus}]$	CMF	WMF	Fe/Si _{ret}
b	1.375 ± 0.041	1.116 ± 0.013	0.261 ± 0.146	$(3.1^{+5.0}_{-3.1}) \times 10^{-5}$	1.00 ± 0.56
c	1.300 ± 0.036	1.103 ± 0.015	0.239 ± 0.084	$(0.0^{+4.4}_{-0.0}) \times 10^{-6}$	0.71 ± 0.26
d	0.388 ± 0.007	0.790 ± 0.010	0.409 ± 0.167	0.084 ± 0.071	$1.22^{+1.30}_{-1.22}$
e	0.699 ± 0.013	0.922 ± 0.015	0.447 ± 0.123	0.094 ± 0.067	1.75 ± 1.17
f	1.043 ± 0.019	1.047 ± 0.015	0.409 ± 0.140	0.105 ± 0.073	1.44 ± 1.14
g	1.327 ± 0.024	1.130 ± 0.016	0.399 ± 0.144	0.119 ± 0.080	1.33 ± 1.29
h	0.327 ± 0.012	0.758 ± 0.013	0.341 ± 0.192	$0.081^{+0.089}_{-0.081}$	$0.13^{+1.80}_{-0.13}$

Notes. Columns are: total mass (M_{ret}) and radius (R_{ret}), CMF, WMF, and Fe/Si molar ratio. In this case the mass and radius are considered as input data (scenario 1).

Table 4. Output parameters retrieved by the MCMC method for all TRAPPIST-1 planets.

Planet	$M_{\text{ret}} [M_{\oplus}]$	$R_{\text{ret}} [R_{\oplus}]$	CMF	WMF	Fe/Si _{ret}
b	1.359 ± 0.043	1.124 ± 0.016	0.259 ± 0.032	$(0.0^{+3.4}_{-0.0}) \times 10^{-6}$	0.79 ± 0.10
c	1.299 ± 0.034	1.103 ± 0.014	0.257 ± 0.031	$(0.0^{+2.7}_{-0.0}) \times 10^{-6}$	0.79 ± 0.11
d	0.387 ± 0.007	0.792 ± 0.010	0.241 ± 0.032	0.036 ± 0.028	0.76 ± 0.12
e	0.695 ± 0.012	0.926 ± 0.012	0.249 ± 0.031	$0.024^{+0.027}_{-0.024}$	0.78 ± 0.12
f	1.041 ± 0.019	1.048 ± 0.013	0.240 ± 0.031	0.037 ± 0.026	0.76 ± 0.12
g	1.331 ± 0.023	1.131 ± 0.015	0.235 ± 0.031	0.047 ± 0.028	0.75 ± 0.12
h	0.326 ± 0.011	0.758 ± 0.013	0.232 ± 0.032	0.055 ± 0.037	0.75 ± 0.12

Notes. Columns are: total mass (M_{ret}) and radius (R_{ret}), CMF, WMF, and Fe/Si molar ratio. In this case, the Fe/Si mole ratio estimated by following Unterborn et al. (2018) is also included as data (scenario 2).

between 0.2 and 0.3 (Table 4). Thus, the CMF of the TRAPPIST-1 planets could be compatible with an Earth-like CMF ($\text{CMF}_{\oplus} = 0.32$).

In scenario 1, the retrieved WMFs for all planets in the system are below 20% within their uncertainties. This maximum WMF limit reduces to 10% for scenario 2. This indicates that the TRAPPIST-1 system is poor in water and other volatiles, especially the inner planets b and c. Both planets are compatible with a dry composition in both scenarios, although the presence of an atmosphere cannot be ruled out given the possible CMF range estimated in scenario 1.

4.2. Water phases

Figure 3 shows the OLR calculated by the atmosphere model and the absorbed radiation (Eqs. (12) and (13)) for planets b, c, and d. For temperatures lower than $\sim T_{\text{surf}} = 2000$ K, the OLR has little dependency on the surface temperature. This is caused by the nearly constant temperature (between 250 and 300 K) of the radiating layers in the thermal IR range (Goldblatt et al. 2013) and it is related to the runaway greenhouse effect (Ingersoll 1969). We obtain a constant OLR or an OLR limit (Nakajima et al. 1992) of 274.3, 273.7, and 254.0 W m^{-2} for planets b, c, and d, respectively. These are close to the OLR limit obtained by Katyal et al. (2019) of 279.6 W m^{-2} for an Earth-like planet. The small difference is due to their different surface gravities. As explained in Sect. 2.3, if the atmosphere model can find a surface temperature at which the OLR and the absorbed radiation are equal, their atmospheres are in global radiative balance. This is the case for planets b and c, whose surface temperatures are approximately 2450 and 2250 K, respectively. These are above the temperatures

where the blanketing effect is effective, named T_{ε} in Marcq et al. (2017), implying that the atmospheres of planets b and c are in a post-runaway state. However, planet d is not in global radiative balance as its absorbed radiation never exceeds its OLR. This means that planet d would be cooling down, and an internal flux of approximately 33 W m^{-2} would be required to supply the extra heat to balance its radiative budget. TRAPPIST-1 inner planets are likely to present an internal heat source due to tidal heating (Barr et al. 2018; Dobos et al. 2019; Turbet et al. 2018). The tidal heat flux estimated for planet d is $F_{\text{tidal}} = 0.16 \text{ W m}^{-2}$ (Barr et al. 2018), which is one order of magnitude lower than needed for radiative-convective balance of a steam atmosphere. Due to the blanketing effect of radiation over the surface of planet d, the OLR limit is larger than the absorbed radiation and hence the planet can cool enough to present its hydrosphere in condensed phases.

Figure 4 shows the (P, T) profiles and the different phases of water we can find in the hydrospheres of the TRAPPIST-1 planets. The maximum WMF of planets b and c are 8.1×10^{-5} and 4.4×10^{-6} , which correspond to a surface pressure of 128.9 and 4.85 bar, respectively.

The thermal structure of their steam atmospheres are dominated by a lower, unsaturated troposphere where water condensation does not occur. Then the atmosphere consists of a middle, saturated troposphere where cloud formation would be possible, extending up to 10 mbar, and finally an isothermal mesosphere above. As we consider a clear transit radius of 20 mbar (Grimm et al. 2018; Mousis et al. 2020), the presence of clouds above this pressure level would flatten the water features in the planetary spectrum (Turbet et al. 2019; Katyal et al. 2020). On the other hand, planets d and e could present water in liquid phase,

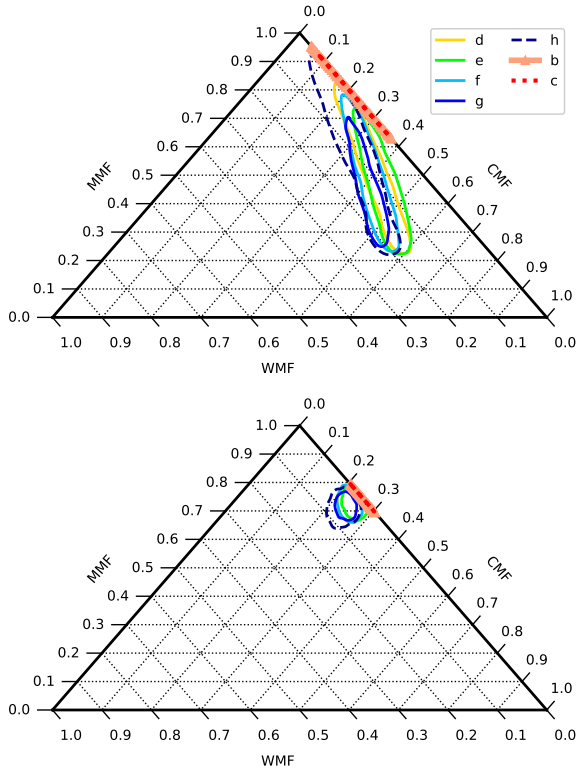


Fig. 2. *Top panel:* 1σ -confidence regions derived from the 2D posterior distributions for the first scenario, where only the masses and radii are available as data. *Bottom panel:* 1σ -confidence regions derived from the 2D posterior distributions for the second scenario, where the Fe/Si abundance ratio from Unterborn et al. (2018) is considered together with the mass and radius for each planet. The axis of the ternary diagram indicates the CMF, the WMF, and the mantle mass fraction $\text{MMF} = 1 - \text{CMF} - \text{WMF}$.

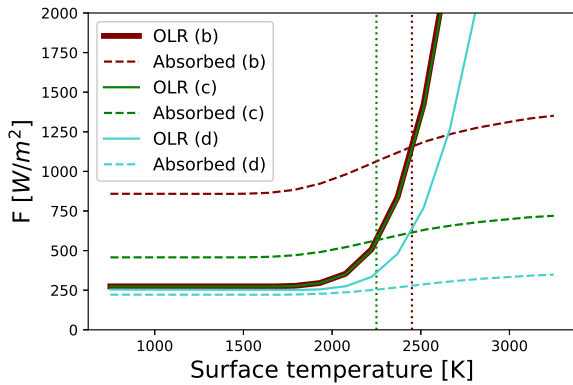


Fig. 3. Outgoing longwave radiation (OLR) and absorbed radiation as a function of surface temperature for the steam atmospheres of TRAPPIST-1 b, c, and d. Vertical dotted lines indicate the surface temperature at which the absorbed flux is equal to the OLR for planets b and c.

which could be partially or completely covered in ice Ih. While the hydrosphere of planet h is not massive enough to attain the high pressures required for ice VII at its base, planets d to g can reach pressures up to a 100 GPa.

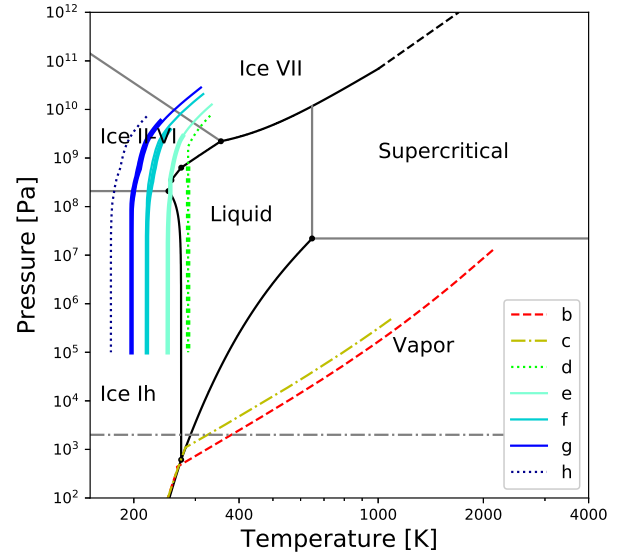


Fig. 4. (P,T) profiles of the hydrospheres of TRAPPIST-1 planets. The dashed-dotted grey horizontal line indicates the 20 mbar pressure level (see text). Thicker lines indicate the profile for the minimum WMF estimated for each planet in scenario 1, while thinner lines mark the profile for the maximum WMF under the same scenario. The minimum WMF of planets b, c, and h is zero.

4.3. Retrieval of atmospheric parameters

Figure 5 shows the output atmospheric parameters (surface temperature, atmospheric thickness, albedo, and atmospheric mass) of TRAPPIST-1 b and c for a water-dominated atmosphere in scenario 1. The total thickness of an atmosphere is related to its scale height, which is defined as $H = RT/\mu g$, where $R = 8.31 \text{ J K}^{-1} \text{ mol}$ is the gas constant, T is the mean atmospheric temperature, μ the mean molecular mass, and g the surface gravity acceleration. For planets b and c, the mean atmospheric temperatures are 940.4 and 499.4 K, and their surface gravities are 10.8 and 10.7 m s^{-2} , respectively. The mean molecular mass for a 99% water and 1% CO_2 atmosphere is 18.3 g mol^{-1} . The mean temperature increases with surface temperature, while the mean molecular mass is determined by the composition of the atmosphere.

For the same composition and surface gravity, the scale height and therefore the thickness of the atmosphere are directly correlated to the surface temperature. As shown in Fig. 5, the atmospheric thickness, z_{atm} , increases with the surface temperature T_{surf} . This is known as the runaway greenhouse radius inflation effect (Goldblatt 2015; Turbet et al. 2019), where a highly irradiated atmosphere is more extended than a colder one despite having similar composition. For planet b, its atmosphere can extend up to 450 km, while planet c presents a maximum extension of 300 km. The minimum limit for the thicknesses is zero, which corresponds to the case of a dry composition. Ortenzi et al. (2020) estimated that for a planet of $1\text{--}1.5 M_{\oplus}$ the maximum atmospheric thickness due to the outgassing of an oxidised mantle is 200 km, which is compatible with the ranges we obtain for the atmospheric thicknesses. Scenario 2 shows the same trends for the atmospheric parameters but with lower atmospheric mass and surface pressure. With their WMF posterior distributions centred in zero and low standard deviation, the surface pressure is below 1 bar and atmospheric thicknesses are

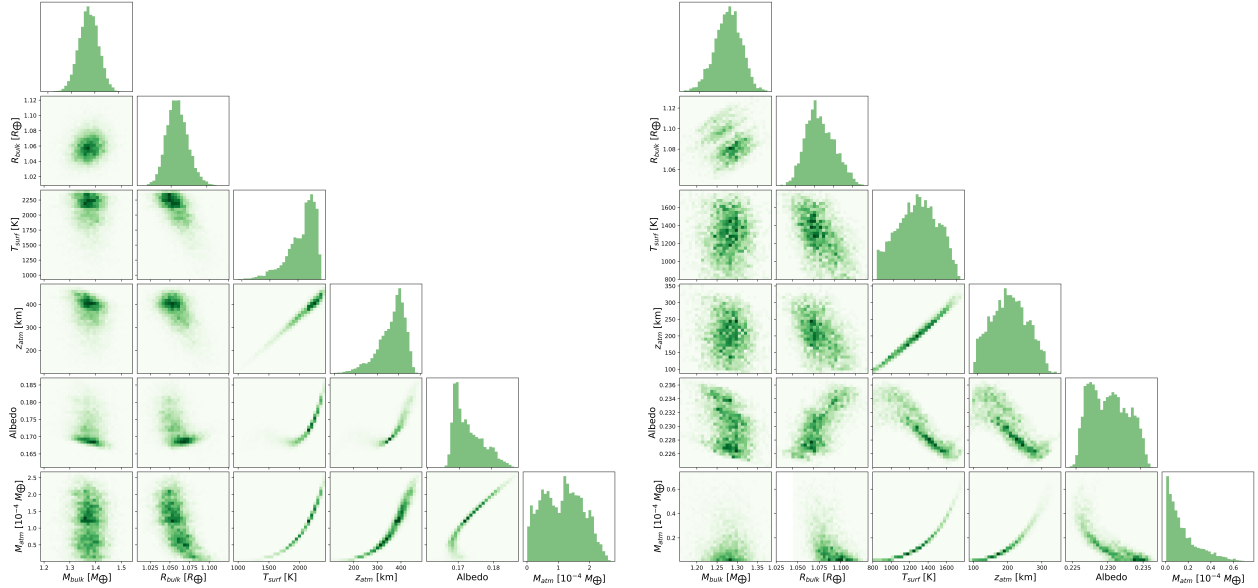


Fig. 5. One- and two-dimensional marginal posterior distributions for the atmospheric parameters (surface temperature T_{surf} , atmospheric thickness z_{atm} , albedo, and atmospheric mass M_{atm}), and bulk mass and radius, M_{bulk} and R_{bulk} , of TRAPPIST-1 b (left panel) and c (right panel). These were derived under scenario 1, where we do not consider Fe/Si data.

below 100 km in most of the accepted models, which means that in scenario 2 planets b and c are most likely dry rocky planets.

5. Discussion

5.1. WMF comparison with previous works

Agol et al. (2021) use the interior and atmosphere models presented in Dorn et al. (2018) and Turbet et al. (2020b) to obtain the WMF estimated for the TRAPPIST-1 planets with updated and more precise radii and masses data from *Spitzer* TTVs (Agol et al. 2021). We therefore limited the comparison to the sole results of Agol et al. (2021) with the same input values. By doing so, we can be certain that the variations in WMF estimates are due to our different modelling approach. Figure 6 shows that planets b and c are most likely dry in scenario 2, where the resulting CMFs are between 0.2 and 0.3 for the whole system. We obtain maximum estimates of 3.4×10^{-6} and 2.7×10^{-6} for b and c, respectively. For the same density, the estimated value of the WMF depends on the CMF that is considered. Therefore, we compare WMF estimates for similar CMFs in this work and Agol et al. (2021). We show our WMF in scenario 2, because the CMF of all planets spans a narrow range between 0.2 and 0.3, which are the most similar values to one of the CMFs assumed by Agol et al. (2021), $\text{CMF} = 0.25$. Our WMF for the steam planets of the TRAPPIST-1 system are in agreement with those of Agol et al. (2021), who calculated a maximum WMF of 10^{-5} for a constant CMF of 0.25. We are able to reduce the maximum limit of the water content of the highly irradiated planets compared to previous studies and establish the most likely WMF with our coupled atmosphere–interior model. The calculation of the total radius requires precise determination of the atmospheric thickness. This depends strongly on the surface temperature and the surface gravity, which are obtained with radiative transfer in the atmosphere, and the calculation of the gravity profile for a bulk mass and composition in the interior self-consistently.

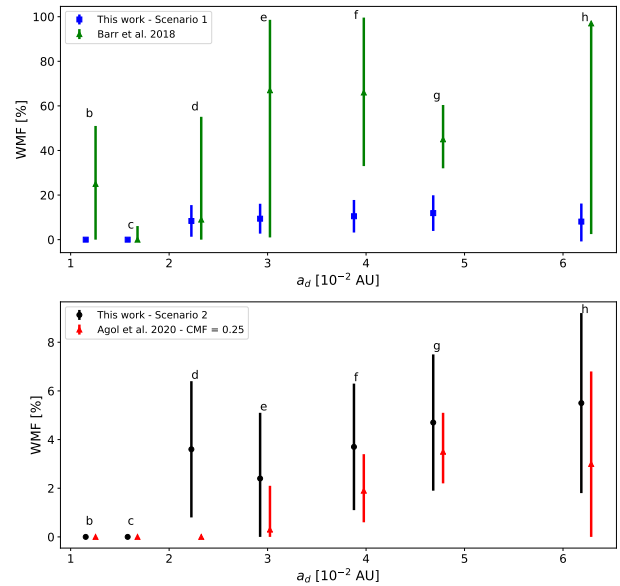


Fig. 6. Water mass fraction as a function of the distance to the star for the TRAPPIST-1 system. *Upper panel:* our estimates for scenario 1 and those of Barr et al. (2018), where only mass and radius data were taken into account. *Lower panel:* corresponds to scenario 2, whose CMF is constrained in a narrow range between 0.2 and 0.3, while for Agol et al. (2021) we show the WMF for a CMF of 0.25.

In the case of planet d, we estimate a WMF of 0.036 ± 0.028 , while Agol et al. (2021) obtain an upper limit of 10^{-5} . The latter estimate considers that water is in vapour form, which is less dense than condensed phases, while our model shows that the surface conditions allow liquid or ice phases, resulting in

Table 5. Comparison between our 1D 1σ confidence regions for the CMF and those of Barr et al. (2018).

Planet	CMF	
	Barr et al. (2018)	This study (2020)
b	0.00–0.43	0.12–0.41
c	0.00–0.98	0.16–0.32
d	0.00–0.39	0.24–0.58
e	0.00–0.40	0.32–0.57
f	0.00	0.24–0.58
g	0.00	0.26–0.54
h	0.00	0.15–0.53

Notes. We show only estimates for scenario 1, because Barr et al. (2018) did not consider any constraints on the Fe/Si ratio based on stellar composition.

a higher WMF. This discrepancy in the possible water phases on the surface of planet d is due to different atmospheric compositions. We consider a water-dominated atmosphere with 1% CO₂, while Agol et al. (2021) and Turbet et al. (2020b) assume a N₂ and H₂O mixture. This difference in composition changes the radiative balance because CO₂ is a strong absorber in the IR compared to N₂, which is a neutral gas. Nonetheless, N₂ is subject to stellar-wind-driven escape and is unlikely to be stable for the inner planets of TRAPPIST-1, while CO₂ is more likely to survive thermal and ion escape processes (Turbet et al. 2020a).

Our WMF for planets e to h are in agreement within uncertainties with those of Agol et al. (2021), although their central values are significantly lower. The EOS employed to compute the density of the water layers in Agol et al. (2021) is also used in Dorn et al. (2018) and Vazan et al. (2013), which agrees well with the widely used SESAME and ANEOS EOSs (Baraffe et al. 2008). These EOS are not consistent with experimental and theoretical data because they overestimate the density at pressures higher than 70 GPa (Mazevet et al. 2019). This yields an underestimation of the WMF for the same total planetary density and CMF.

For the specific case of scenario 1, with no assumptions on the stellar composition and the Fe/Si mole ratio, we compared our CMF and WMF with those obtained in Barr et al. (2018) (Fig. 6 and Table 5). These latter authors use masses and radii data given by Wang et al. (2017); they obtain lower masses compared to Agol et al. (2021) while their radii are approximately the same, which would explain why Barr et al. (2018) tend to overestimate the water content of the TRAPPIST-1 planets. Moreover, most of the mass uncertainties in Wang et al. (2017) are 30–50%, while the mass uncertainties obtained by Agol et al. (2021) are 3–5%. This causes Barr et al. (2018) to calculate wider CMF and WMF 1σ confidence intervals. In addition, there are differences between our interior modelling approach and that of Barr et al. (2018). For example, according to the results of these latter authors, planet b can have up to 50% of its mass as water. This high WMF value is due to the assumption that the hydrosphere is in liquid and ice I phases, and high-pressure ice polymorphs (HPPs), which are more dense than the steam atmosphere we consider. In contrast, the CMF seems to be closer to our estimates, especially for planets b, d, and e, where their maximum CMF is approximately 0.40, in agreement with our CMF 1σ intervals.

We can also discuss the possible habitability of the hydrospheres of the TRAPPIST-1 planets by comparing our WMF

estimates with the layer structure as a function of planetary mass and water content obtained by Noack et al. (2016). According to Noack et al. (2016), a habitable hydrosphere must be structured in a single liquid water ocean or in several ice layers that enable the formation of a lower ocean layer. This lower ocean would be formed by the heat supplied by the mantle that melts the high-pressure ice in the ice-mantle boundary (Noack et al. 2016). For planet d, a surface liquid ocean would form for all its possible WMF if the atmosphere allows for the presence of condensed phases. For planets e, f, and g, the hydrosphere could be stratified in a surface layer of ice Ih and a liquid or ice II-VI layer. In the case of low-pressure ices II-VI, their base could be melted by the heat provided by the mantle, and form a lower ocean layer as suggested by Noack et al. (2016). At WMF ≥ 0.10 , less than 50% of the possible configurations enable a habitable sub-surface ocean layer, and at a WMF ≥ 0.14 , the hydrosphere is uninhabitable. In scenario 1, planets e to g reach these values within uncertainties, although their minimum values extend down to 0–0.03 in WMF, which would be the habitable regime.

5.2. System formation and architecture

In the case of scenario 1, where no Fe/Si data are assumed, the WMF increases with the distance to the star with the exception of planet h, whose WMF is similar to that of planet d. In the case of scenario 2, where a common Fe/Si of 0.76 ± 0.12 is assumed for the whole system, the WMF increases with the distance to the star (Fig. 6) with the exception of planet d whose WMF is similar to that of planet f, which is more water-rich than planet e. This slight deviation from the observed trend could be explained by migration, where planet d could have formed beyond the snow line before migrating inwards (Raymond et al. 2018). In addition, pebble ablation and water recycling back into the disc could have been less efficient for planet d than for planet e (Coleman et al. 2019). On the other hand, the gas at the distance at which planet d formed could have been more enriched in volatiles than the outer planets, accreting more water ice than planet e in a ‘cold finger’ (Stevenson & Lunine 1988; Cyr et al. 1998). Pebble formation in the vicinity of the water ice line can induce important enhancements of the water ice fraction in those pebbles due to the backward diffusion of vapour through the snowline and the inward drift of ice particles. Therefore, if a planet forms from this material, it should be more water-rich than those formed further out (Mousis et al. 2019). These formation scenarios could explain the high WMF of planet d when we assume that its water layer is in condensed phases. Post-formation processes could also have shaped the trend of the WMF with axis, such as atmospheric escape due to XUV and X-ray emission from their host star. Bolmont et al. (2017) estimated a maximum water loss of 15 Earth Oceans (EO) for TRAPPIST-1 b and c and 1 EO for planet d. If we were to assume that the current WMFs are the central values of the posterior distributions we derived in scenario 1, planets b, c, and d would have had an initial WMF of 2.37×10^{-3} , 2.50×10^{-3} and 0.085, respectively. Therefore, atmospheric escape would have decreased the individual WMF of each planet, but the increase of WMF with distance from the star would have been preserved.

In addition to the WMF-axis trend, we can differentiate the very water-poor, close-in planets, b and c, from the outer, water-rich planets, d to h. This has been reported as a consequence of pebble accretion in the formation of other systems, such as the Galilean moons. While Io is dry, Callisto and Ganymede are water-rich, with Europa showing an intermediate WMF of 8% (Ronnet et al. 2017). Pebble-driven formation can produce

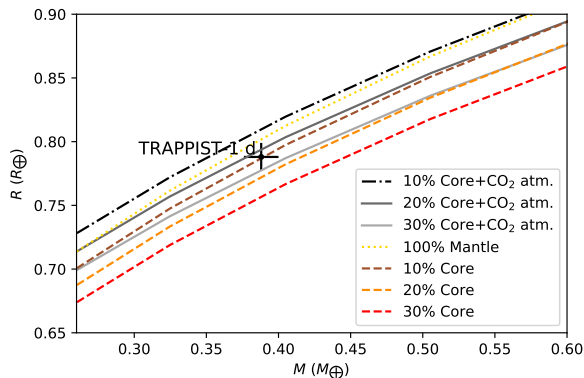


Fig. 7. Mass–radius relationships for planets with CO₂-dominated atmospheres assuming different CMFs. The surface pressure is 300 bar. The black dot and its error bars indicate the location and uncertainties of planet d in the mass–radius diagram.

planets with WMF $\geq 15\%$ if these are formed at the water ice line (Coleman et al. 2019; Schoonenberg et al. 2019). In contrast, planets formed within the ice line would present WMFs of less than 5% (Liu et al. 2020; Coleman et al. 2019), which is close to the mean value we calculate for planet h, 5.5%. The maximum WMF limit in the first scenario is approximately 20%. This maximum limit is significantly lower than the typical WMF generated by the planetesimal accretion scenario, which is 50–40% (Miguel et al. 2020). Therefore, our results are consistent with the pebble-driven formation scenario.

5.3. Alternative atmospheric compositions

However, the atmosphere of planet d could be dominated by other atmospheric gases different from H₂O-based mixtures, which could produce an extended atmosphere and increase the total planetary radius. Hydrogen-dominated atmospheres have been deemed unlikely as one of the possible atmospheric compositions for all planets in the TRAPPIST-1 system, both cloud-free (de Wit et al. 2016, 2018) and with high-altitude clouds and hazes (Grimm et al. 2018; Ducrot et al. 2020). Similarly, CH₄-dominated atmospheres are not probable according to the photometry data of the *Spitzer* Space Telescope (Ducrot et al. 2020). Therefore, our best candidate to explain the low density of planet d in a water-poor scenario is CO₂. We find that a CO₂-dominated atmosphere with 1% water vapour in planet d would be in radiative-convective equilibrium by computing the OLR and absorbed radiation, as we have for water-dominated atmospheres. The resulting surface temperature is approximately 950 K, which is slightly higher than the surface temperature of Venus (700 K) with a higher water vapour mixing ratio. Figure 7 introduces the mass–radius relationships for different CMFs, assuming a CO₂-dominated atmosphere with a surface pressure of 300 bar. Planet d appears to be compatible with a planet with a CO₂-dominated atmosphere and CMF between 0.2 and 0.3, which is a very likely CMF range for TRAPPIST-1 planets based on our analysis. Surface pressures lower than 300 bar would yield lower atmospheric thicknesses, and so it would be necessary to consider a lower CMF to explain the observed density of planet d. CO₂ in the case of planet d can be provided by volcanic outgassing (Ortenzi et al. 2020), as its internal heat flux produced by tidal heating is in the range 0.04–2 W m⁻², which favours plate tectonics (Papaloizou et al. 2018). Secondary CO₂-dominated atmospheres could have traces of O₂, N₂, and water vapour.

6. Conclusions

We present an interior structure model for low-mass planets at different irradiances that is valid for a wide range of water phases and was derived from the approaches of Brugger et al. (2017) and Mousis et al. (2020). For highly irradiated planets, we couple a 1D water steam atmosphere in radiative–convective equilibrium with a high-pressure convective layer in supercritical phase. The density in this layer is computed using an accurate EOS for high-pressure and high-temperature water phases. For temperate planets whose surface conditions allow the formation of condensed phases, we implemented a hydrosphere with liquid water and ice phases Ih, II, III, V, VI and VII. In addition, we adapted the MCMC Bayesian algorithm described in Dorn et al. (2015) to our interior model to derive the posterior distributions of the compositional parameters, WMF and CMF, given mass, radius, and stellar composition data. We then applied our interior model to the particular case of TRAPPIST-1 planets using the latest mass and radius data from *Spitzer* (Agol et al. 2021).

We characterised the hydrospheres of TRAPPIST-1 planets by calculating their *P-T* profiles and thermodynamic phases. Planets b and c are warm enough to present steam atmospheres. They could hold post-runaway greenhouse atmospheres with thicknesses of up to 450 km and surface temperatures of up to 2500 K, which means that they are extended enough to be suitable targets for atmospheric characterisation by future space-based facilities such as *James Webb* Space Telescope (JWST). Moreover, planets d to g present hydrospheres in condensed phases. These hydrospheres can contain high-pressure ices that start to form at 10⁹–10¹⁰ Pa.

We obtained CMF and WMF probability distributions for all planets in the system. We find that the Fe/Si mole ratio of the system is in the 0.45–0.97 range without considering any assumption on the chemical composition of the stellar host. This Fe/Si range corresponds to a CMF value in the 0.23–0.40 range, making the CMF of TRAPPIST-1 planets compatible with an Earth-like value (0.32). In addition, our WMF estimates agree within uncertainties with those derived by Agol et al. (2021), although their most likely values are considerably lower for planets with condensed phases. In the case of planets with steam hydrospheres, their densities are compatible with dry rocky planets with no atmospheres. Nevertheless, we cannot rule out the presence of an atmosphere with the Fe/Si range we derive without any assumption on the chemical composition of the host star. When considering a possible estimate of the Fe/Si ratio of the host star (scenario 2), we obtain lower maximum limits of the WMF for planets b and c compared to previously calculated limits by Agol et al. (2021) for a similar CMF of 0.25. Our estimated WMFs in steam and condensed phases are consistent with an increase in the WMF with progressing distance from the host star. This trend, as well as the maximum WMF we calculate, favour pebble-driven accretion as a plausible formation mechanism for the TRAPPIST-1 system. However, planet d presents a slightly higher WMF than planet e. This could be due to processes that took place during planet formation, such as migration, a low-efficient ablation of pebbles, and gas recycling, or an enhancement of the water ice fraction in pebbles at the distance of the disc where planet d formed. An extended atmosphere dominated by greenhouse gases different from a water-dominated atmosphere, such as CO₂, could also explain the low-density of planet d compared to planet e.

Future work should include more atmospheric processes and species that determine the mass–radius relations of planets with secondary atmospheres in the super-Earth and sub-Neptune regime. These can vary the atmospheric thickness and increase

the total planetary radius with varying atmospheric mass while other compositional parameters change the bulk radius. These should be integrated into one single interior-atmosphere model combined in a MCMC Bayesian framework such as the one we used in this study. This statistical approach has been employed with interior models for planets with H/He-dominated atmospheres (Dorn et al. 2017b,a, 2018), or dry planets (Plotnykov & Valencia 2020), but not for planets with secondary, CO₂ and steam-dominated atmospheres. The integrated model should also include a description of escape processes, such as hydrodynamic or Jeans escapes, which is particularly interesting for exploration of the lifetime of secondary atmospheres. Close-in, low-mass planets are likely to outgas atmospheric species such as CO₂, and form O₂ via photodissociation of outgassed H₂O during their magma ocean stage or due to plate tectonics (Chao et al. 2020). A mixture of these gases should therefore be considered to study the thermal structure of planets with secondary atmospheres. Planets b and c in the TRAPPIST-1 system could present magma oceans due to their high surface temperatures ($T \geq 1300$ K) (Barr et al. 2018; Chao et al. 2020), and the maximum surface pressure we have obtained here can be used to assess the current outgassing rate in magma ocean studies (Noack et al. 2017; Baumeister et al., in prep.) and better constrain the WMF for the interior magma ocean models (e.g. Katyal et al. 2020) in the future.

Acknowledgements. M.D. and O.M. acknowledge support from CNES. We acknowledge the anonymous referee whose comments helped improve and clarify this manuscript.

References

- Agol, E., Dorn, C., Grimm, S. L., et al. 2021, *Planet. Sci. J.*, **2**, 1
- Báez, L. A., & Clancy, P. 1995, *J. Chem. Phys.*, **103**, 9744
- Baraffe, I., Chabrier, G., & Barman, T. 2008, *A&A*, **482**, 315
- Barr, A. C., Dobos, V., & Kiss, L. L. 2018, *A&A*, **613**, A37
- Benz, W. 2017, *CHEOPS Fifth Science Workshop*, 1
- Bezacier, L., Journaux, B., Perrillat, J.-P., et al. 2014, *J. Chem. Phys.*, **141**, 104505
- Bolmont, E., Selsis, F., Owen, J. E., et al. 2017, *MNRAS*, **464**, 3728
- Brugger, B., Mousis, O., Deleuil, M., & Lunine, J. I. 2016, *ApJ*, **831**, L16
- Brugger, B., Mousis, O., Deleuil, M., & Deschamps, F. 2017, *ApJ*, **850**, 93
- Chao, K.-H., deGraffenried, R., Lach, M., et al. 2020, arXiv e-prints, [arXiv:2012.07337]
- Choukroun, M., & Grasset, O. 2010, *J. Chem. Phys.*, **133**, 144502
- Coleman, G. A. L., Leleu, A., Alibert, Y., & Benz, W. 2019, *A&A*, **631**, A7
- Cyr, K. E., Sears, W. D., & Lunine, J. I. 1998, *Icarus*, **135**, 537
- de Wit, J., Wakeford, H. R., Gillon, M., et al. 2016, *Nature*, **537**, 69
- de Wit, J., Wakeford, H. R., Lewis, N. K., et al. 2018, *Nat. Astron.*, **2**, 214
- Dobos, V., Barr, A. C., & Kiss, L. L. 2019, *A&A*, **624**, A2
- Dorn, C., Khan, A., Heng, K., et al. 2015, *A&A*, **577**, A83
- Dorn, C., Hinkel, N. R., & Venturini, J. 2017a, *A&A*, **597**, A38
- Dorn, C., Venturini, J., Khan, A., et al. 2017b, *A&A*, **597**, A37
- Dorn, C., Mosegaard, K., Grimm, S. L., & Alibert, Y. 2018, *ApJ*, **865**, 20
- Ducrot, E., Gillon, M., Delrez, L., et al. 2020, *A&A*, **640**, A112
- Dunaeva, A. N., Antsyshkin, D. V., & Kuskov, O. L. 2010, *Sol. Syst. Res.*, **44**, 202
- Fei, Y., Mao, H.-K., & Hemley, R. J. 1993, *J. Chem. Phys.*, **99**, 5369
- Feistel, R., & Wagner, W. 2006, *J. Phys. Chem. Ref. Data*, **35**, 1021
- Forget, F., & Leconte, J. 2014, *Phil. Trans. R. Soc. London Ser. A*, **372**, 20130084
- French, M., Mattsson, T. R., Nettelmann, N., & Redmer, R. 2009, *Phys. Rev. B*, **79**, 054107
- French, M., Desjarlais, M. P., & Redmer, R. 2016, *Phys. Rev. E*, **93**, 022140
- Gagnon, R. E., Kieffe, H., Clouter, M. J., & Whalley, E. 1990, *J. Chem. Phys.*, **92**, 1909
- Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, *Nature*, **542**, 456
- Goldblatt, C. 2015, *Astrobiology*, **15**, 362
- Goldblatt, C., Robinson, T. D., Zahnle, K. J., & Crisp, D. 2013, *Nat. Geosci.*, **6**, 661
- Grimm, S. L., Demory, B.-O., Gillon, M., et al. 2018, *A&A*, **613**, A68
- Haar, L., Gallagher, J. S., & Kell, G. S. 1984, NBS/NRC steam tables thermodynamic and transport properties and computer programs for vapor and liquid states of water in SI units (Taylors & Francis Inc.) <https://www.osti.gov/biblio/5614915>
- Hinkel, N. R., Timmes, F. X., Young, P. A., Pagano, M. D., & Turnbull, M. C. 2014, *AJ*, **148**, 54
- Hinkel, N. R., Young, P. A., Pagano, M. D., et al. 2016, *ApJS*, **226**, 4
- Hinkel, N. R., Mamajek, E. E., Turnbull, M. C., et al. 2017, *ApJ*, **848**, 34
- Ingersoll, A. P. 1969, *J. Atm. Sci.*, **26**, 1191
- Katyal, N., Nikolauou, A., Godolt, M., et al. 2019, *ApJ*, **875**, 31
- Katyal, N., Ortenzi, G., Lee Grenfell, J., et al. 2020, *A&A*, **643**, A81
- Leon, G. C., Rodriguez Romo, S., & Tchijov, V. 2002, *J. Phys. Chem. Solids*, **63**, 843
- Lincowski, A. P., Meadows, V. S., Crisp, D., et al. 2018, *ApJ*, **867**, 76
- Liu, B., Lambrechts, M., Johansen, A., Pascucci, I., & Henning, T. 2020, *A&A*, **638**, A88
- Lodders, K., Palme, H., & Gail, H. P. 2009, *Landolt Börnstein*, **4B**, 712
- Marcq, E. 2012, *J. Geophys. Res. Planets*, **117**, E01001
- Marcq, E., Salvador, A., Massol, H., & Davaille, A. 2017, *J. Geophys. Res. Planets*, **122**, 1539
- Mazevet, S., Licari, A., Chabrier, G., & Potekhin, A. Y. 2019, *A&A*, **621**, A128
- McKay, A. J., DiSanti, M. A., Kelley, M. S. P., et al. 2019, *AJ*, **158**, 128
- Miguel, Y., Cridland, A., Ormel, C. W., Fortney, J. J., & Ida, S. 2020, *MNRAS*, **491**, 1998
- Mousis, O., Ronnet, T., & Lunine, J. I. 2019, *ApJ*, **875**, 9
- Mousis, O., Deleuil, M., Aguichine, A., et al. 2020, *ApJ*, **896**, L22
- Nakajima, S., Hayashi, Y.-Y., & Abe, Y. 1992, *J. Atm. Sci.*, **49**, 2256
- Noack, L., Höning, D., Rivoldini, A., et al. 2016, *Icarus*, **277**, 215
- Noack, L., Rivoldini, A., & Van Hoolst, T. 2017, *Phys. Earth Planet. Inter.*, **269**, 40
- Ortenzi, G., Noack, L., Sohl, F., et al. 2020, *Sci. Rep.*, **10**, 10907
- Papaloizou, J. C. B., Szuszkiewicz, E., & Terquem, C. 2018, *MNRAS*, **476**, 5032
- Plotnykov, M., & Valencia, D. 2020, *MNRAS*, **499**, 932
- Pluriel, W., Marcq, E., & Turbet, M. 2019, *Icarus*, **317**, 583
- Raymond, S. N., Boulet, T., Izidoro, A., Esteves, L., & Bitsch, B. 2018, *MNRAS*, **479**, L81
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, *J. Astron. Telesc. Instrum. Syst.*, **1**, 014003
- Ronnet, T., Mousis, O., & Vernazza, P. 2017, *ApJ*, **845**, 92
- Schoonenberg, D., Liu, B., Ormel, C. W., & Dorn, C. 2019, *A&A*, **627**, A149
- Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007, *ApJ*, **669**, 1279
- Shaw, G. H. 1986, *J. Chem. Phys.*, **84**, 5862
- Sotin, C., Grasset, O., & Mocquet, A. 2007, *Icarus*, **191**, 337
- Stevenson, D. J., & Lunine, J. I. 1988, *Icarus*, **75**, 146
- Tchijov, V., Ayala, R. B., Leon, G. C., & Nagornov, O. 2004, *J. Phys. Chem. Solids*, **65**, 1277
- Tulk, C. A., Gagnon, R. E., Kieffe, H., & Clouter, M. J. 1997, *J. Chem. Phys.*, **107**, 10684
- Turbet, M., Bolmont, E., Leconte, J., et al. 2018, *A&A*, **612**, A86
- Turbet, M., Ehrenreich, D., Lovis, C., Bolmont, E., & Fauchez, T. 2019, *A&A*, **628**, A12
- Turbet, M., Bolmont, E., Bourrier, V., et al. 2020a, *Space Sci. Rev.*, **216**, 100
- Turbet, M., Bolmont, E., Ehrenreich, D., et al. 2020b, *A&A*, **638**, A41
- Unterborn, C. T., Desch, S. J., Hinkel, N. R., & Lorenzo, A. 2018, *Nat. Astron.*, **2**, 297
- Vazan, A., Kovetz, A., Podolak, M., & Helled, R. 2013, *MNRAS*, **434**, 3283
- Wagner, W., & Prub, A. 2002, *J. Phys. Chem. Ref. Data*, **31**, 387
- Wang, S., Wu, D.-H., Barclay, T., & Laughlin, G. P. 2017, arXiv e-prints, [arXiv:1704.04290]
- Zeng, L., Jacobsen, S. B., Sasselov, D. D., et al. 2019, *Proc. Natl. Acad. Sci.*, **116**, 9723

A.2. Water content trends in K2-138 and other low-mass multi-planetary systems (Acuña et al. 2022)

L. Acuña, T. A. Lopez, T. Morel, M. Deleuil, O. Mousis, A. Aguichine, E. Marq, and A. Santerne

This is the second first-author publication resulting from the research carried out in this thesis. The model implementation is similar to the one used in [Acuña et al. \(2021\)](#). See Sect. 2.2 to read more about the version of the interior model that includes supercritical water phases, Chapter 3 for the interior-atmosphere coupling algorithm, and Sect. 5.1 for the non-adaptive MCMC. The results of applying this version of the interior-atmosphere model to a sample of multiplanetary systems with low-mass planets can be found in Sect. 6.2, together with a discussion on planet formation scenarios.

The work shown in this publication was also presented in the following conferences:

- PLATO mission conference 2021, as a contributed talk, in October 2021 (virtual)
- Europlanet Science Conference (EPSC) 2021, as a contributed talk, in September 2021 (virtual)
- Semaine de la Astrophysique Française (SF2A) 2021, as a contributed talk, in June 2021 (virtual)

Water content trends in K2-138 and other low-mass multi-planetary systems^{*}

L. Acuña¹, T. A. Lopez¹, T. Morel², M. Deleuil¹, O. Mousis¹, A. Aguichine¹, E. Marcq³, and A. Santerne¹

¹ Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
e-mail: lorena.acuna@lam.fr

² Space sciences, Technologies and Astrophysics Research (STAR) Institute, Université de Liège, Quartier Agora,
Allée du 6 Août 19c, Bât. B5C, 4000 Liège, Belgium

³ LATMOS/CNRS/Sorbonne Université/UVSQ, 11 boulevard d'Alembert, Guyancourt, 78280, France

Received 5 October 2021 / Accepted 26 January 2022

ABSTRACT

Context. Both rocky super-Earths and volatile-rich sub-Neptunes have been found simultaneously in multi-planetary systems, suggesting that these systems are appropriate to study different composition and formation pathways within the same environment.

Aims. We perform a homogeneous interior structure analysis of five multi-planetary systems to explore compositional trends and their relation with planet formation. For one of these systems, K2-138, we present revised masses and stellar host chemical abundances to improve the constraints on the interior composition of its planets.

Methods. We conducted a line-by-line differential spectroscopic analysis on the stellar spectra of K2-138 to obtain its chemical abundances and the planetary parameters. We selected multi-planetary systems with five or more low-mass planets ($M < 20 M_{\oplus}$) that have both mass and radius data available. We carried out a homogeneous interior structure analysis on the planetary systems K2-138, TOI-178, Kepler-11, Kepler-102, and Kepler-80. We estimated the volatile mass fraction of the planets in these systems assuming a volatile layer constituted of water in steam and supercritical phases. Our interior-atmosphere model took the effects of irradiation on the surface conditions into account.

Results. K2-138 inner planets present an increasing volatile mass fraction with distance from their host star, while the outer planets present an approximately constant water content. This is similar to the trend observed in TRAPPIST-1 in a previous analysis with the same interior-atmosphere model. The Kepler-102 system could potentially present this trend. In all multi-planetary systems, the low volatile mass fraction of the inner planets could be due to atmospheric escape, while the higher volatile mass fraction of the outer planets can be the result of accretion of ice-rich material in the vicinity of the ice line with later inward migration. Kepler-102 and Kepler-80 present inner planets with high core mass fractions which could be due to mantle evaporation, impacts, or formation in the vicinity of rocklines.

Key words. stars: abundances – stars: individual: K2-138 – planets and satellites: interiors – planets and satellites: composition – planets and satellites: individual: K2-138 – methods: numerical

1. Introduction

Multi-planetary systems appear to be suitable distant laboratories to explore the diversity of small planets, as well as their formation and evolution pathways. This is the case for Kepler-36 (Carter et al. 2012), where its two planets, b and c, present periods of 14 and 16 days with densities of 7.5 and 0.9 g cm⁻³, respectively. This suggests that these planets may have formed in different environments within the same protoplanetary disc before migrating inwards. Furthermore, a decreasing density gradient with distance from the host star in multi-planetary systems with six to seven planets, such as TRAPPIST-1 (Acuña et al. 2021; Agol et al. 2021) and TOI-178 (Leleu et al. 2021), suggest that there might be a transition between the rocky, inner super-Earths and the outer, volatile-rich sub-Neptunes. This transition is most probably due to the presence of the snowline in the protoplanetary disc (Ruden 1999).

Nevertheless, there are currently several limitations to determining the variation of the volatile mass fraction of planets

within their systems, including the precision reached on the fundamental parameters of both the planets and the star as well as the different assumptions considered between different interior structure models. These assumptions include whether the volatile layer of the planet is fully constituted of H/He (Lopez & Fortney 2014), an ice layer (Zeng et al. 2019), an ice layer with a H/He atmosphere on top (Dorn et al. 2015), or a steam and/or supercritical water layer (Mousis et al. 2020; Turbet et al. 2020). To overcome the differences in volatile mass fraction estimates of multi-planetary systems due to the different compositions of the volatile layer between interior structure models, we performed a homogeneous analysis of the interior structure and composition of several multi-planetary systems. For our interior structure model, we assumed that the volatile layer is water-dominated, following the approach of Mousis et al. (2020) and Acuña et al. (2021). This analysis allowed us to uncover volatile and core mass fraction trends, and their connection with planet formation and evolution. We used previously published masses, radii, and stellar composition data for four systems, and we performed our own spectroscopic analysis to improve the parameters of one system, K2-138, whose detection was reported in Christiansen et al. (2018). K2-138 harbours six small planets

^{*} Based on observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 198.C-0.168.

in a chain of near 3:2 mean-motion resonances and benefitted from a radial velocity ground-based follow-up with HARPS on the 3.6 m telescope at La Silla Observatory, leading to the confirmation and mass measurements of the four inner planets (Lopez et al. 2019), with relatively good precisions given the standard today. In order to bring stronger constraints on the stellar parameters and abundances and further reduce the degeneracies in the planetary structure modelling, we carried out an in-depth analysis of K2-138.

Section 2 presents the new detailed analysis of the stellar host in the K2-138 system, which allowed us to derive stellar fundamental parameters and the elemental abundances using the Sun and α Cen B as benchmarks. Section 3 describes a new Bayesian analysis of the HARPS radial velocities and K2 photometry, using the new stellar parameters.

We describe our interior-atmosphere modelling in Sect. 4, including our calculation of atmospheric mass-loss rates to infer the current presence or absence of volatiles. We present the volatile and core mass fraction trends for each mutiplanetary system as a result of our homogeneous analysis in Sect. 5. Finally, we discuss the planet formation and evolution mechanisms that could have shaped these compositional trends in Sect. 6. We present our concluding remarks in Sect. 7.

2. Spectroscopic analysis

K2-138 stellar parameters and abundances were derived based on a differential, line-by-line analysis relative to the Sun. The solar abundances are determined as part of such an analysis (e.g. Meléndez et al. 2012) and a set of reference values is not assumed. We used the HARPS spectra retrieved under programme ID 198.C-0.168. These were corrected from systemic velocity and planetary reflex motion, removing the spectra with a signal-to-noise ratio (S/N) lower than 10 in order 47 (550 nm) and the ones contaminated by the moonlight (S/N above 1.0 in fibre B). We then co-added the spectra in a single 1D spectrum and normalised it to the continuum. For the Sun, we used the HARPS spectra extracted from the ESO instrument archives¹, acquired under programme ID 088.C-0323. The reduction of the solar spectrum, obtained just as the spectrum of the light reflected by Vesta, is detailed in Haywood et al. (2016) and the co-addition was performed as for K2-138.

The stellar parameters and abundances of 24 metal species were self-consistently determined from the spectra, plane-parallel MARCS model atmospheres (Gustafsson et al. 2008), and the 2017 version of the line-analysis software MOOG originally developed by Sneden (1973). The equivalent widths (EWs) were measured manually using IRAF² tasks assuming Gaussian profiles. Strong lines with $RW = \log(EW/\lambda) > -4.80$ were discarded. This constraint on the line strength was relaxed for Mg because it would result in no Mg I lines left.

2.1. Stellar parameters

The stellar parameters of K2-138 and α Cen B appear to be similar (see below). Therefore, we also analysed the latter for benchmarking because it has accurate and nearly model-independent T_{eff} and $\log g$ estimates from long-baseline

interferometry and asteroseismology, respectively. K2-138 and α Cen B were observed with exactly the same instrumental setup, which ensures the highest consistency (Bedell et al. 2014). The α Cen B spectra were selected from the ESO archive, keeping those corrected from the blaze and with an S/N higher than 350 in order 47. For α Cen B, we adopt in the following $T_{\text{eff}} = 5231 \pm 21$ K derived by Kervella et al. (2017) from their Very Large Telescope Interferometer (VLTI)/Precision Integrated-Optics Near-infrared Imaging Experiment (PIONIER) measurements and the bolometric flux of Boyajian et al. (2013). We also assumed $\log g = 4.53 \pm 0.02$ dex (Heiter et al. 2015) based on scaling relations making use of the frequency of maximum oscillation power, ν_{max} , determined from radial-velocity time series by Kjeldsen et al. (2008).

The model parameters (T_{eff} , $\log g$, ξ , and [Fe/H]) were iteratively modified until the excitation and ionisation balance of iron was fulfilled and the Fe I abundances exhibited no trend with RW. The abundances of iron and the α elements were also required to be consistent with the values adopted for the model atmosphere. For the solar analysis, T_{eff} and $\log g$ were held fixed to 5777 K and 4.44 dex, respectively, whereas the microturbulence, ξ , was left as a free parameter. The uncertainties in the stellar parameters were computed as in Morel (2018).

We first carried out the analysis of α Cen B and K2-138 using various iron line lists (Biazzo et al. 2012; Doyle et al. 2017; Feltzing & Gonzalez 2001; Jofré et al. 2014; Meléndez et al. 2014; Morel et al. 2014; Reddy et al. 2003; Tsantaki et al. 2019). For Jofré et al. (2014), we adopted their FG Dwarf “FGDa” line list. The goal was to identify the line list that provides the most accurate parameters based on a comparison with the interferometric and asteroseismic constraints at hand for α Cen B. To ensure the highest consistency, the spectral features on which the analysis is based for a given line list were exactly the same for the three stars.

The parameters obtained are given in Table 1 and shown in Fig. 1. The surface gravity of α Cen B appears to be underestimated in most cases. We also experimented with the LW13 Ti line list of Tsantaki et al. (2019) to constrain this quantity through Ti ionisation balance. As discussed by these authors, this leads to a larger value, amounting to ~ 0.11 dex here. However, it still falls short of matching the seismic value. As can be seen in Fig. 1, the only notable difference between the parameters of α Cen B and K2-138 is that the latter is slightly poorer in metals. Indeed, a differential analysis of K2-138 with respect to α Cen B adopting the line list of Biazzo et al. (2012) gives the following results: $\Delta T_{\text{eff}} = -10 \pm 45$ K, $\Delta \log g = +0.02 \pm 0.09$ dex, $\Delta \xi = +0.03 \pm 0.09$ km s⁻¹, and $\Delta[\text{Fe}/\text{H}] = -0.11 \pm 0.04$. For the abundance analysis of K2-138, in the following we adopt the parameters provided by the line list of Biazzo et al. (2012): $T_{\text{eff}} = 5275 \pm 50$ K, $\log g = 4.50 \pm 0.11$, $\xi = 0.95 \pm 0.10$ km s⁻¹, and $[\text{Fe}/\text{H}] = +0.08 \pm 0.05$. This choice was motivated by the fact that it leads to parameters that reproduce those of the reference for α Cen B within the errors. In addition, the metallicity is within the range of accepted values for the binary system (Morel 2018, and references therein).

However, from the comparison to the interferometric-based T_{eff} in Fig. 1, we cannot rule out that the effective temperature of K2-138 is slightly overestimated at the ~ 50 K level. The analysis was also repeated using Kurucz atmosphere models (Castelli & Kurucz 2003). The following modest deviations with respect to the default values (Kurucz – MARCS) were found: $\Delta T_{\text{eff}} \sim +10$ K, $\Delta \log g \sim +0.02$ dex, and $\Delta[\text{Fe}/\text{H}] \sim +0.02$ dex. We examine the robustness of our abundance results against such putative systematic errors in Sect. 2.2. In any case, we find that K2-138 is

¹ <http://archive.eso.org>

² IRAF is distributed by the National Optical Astronomy Observatories, operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.

Table 1. Stellar parameters of α Cen B and K2-138, as obtained from the various iron line lists.

Iron line list	α Cen B				K2-138			
	T_{eff} (K)	$\log g$	ξ (km s^{-1})	[Fe/H]	T_{eff} (K)	$\log g$	ξ (km s^{-1})	[Fe/H]
Biazzo et al. (2012)	5285 \pm 60	4.49 \pm 0.14	0.909 \pm 0.121	0.200 \pm 0.051	5275 \pm 50	4.50 \pm 0.11	0.945 \pm 0.099	0.084 \pm 0.043
Doyle et al. (2017)	5245 \pm 32	4.35 \pm 0.08	0.490 \pm 0.146	0.185 \pm 0.043	5235 \pm 30	4.43 \pm 0.07	0.450 \pm 0.146	0.083 \pm 0.034
Feltzing & Gonzalez (2001)	5330 \pm 41	4.48 \pm 0.11	0.890 \pm 0.100	0.220 \pm 0.040	5280 \pm 38	4.46 \pm 0.10	0.915 \pm 0.084	0.100 \pm 0.035
Jofré et al. (2014)	5210 \pm 77	4.31 \pm 0.11	0.500 \pm 0.221	0.181 \pm 0.063	5210 \pm 66	4.37 \pm 0.11	0.555 \pm 0.190	0.069 \pm 0.054
Meléndez et al. (2014)	5270 \pm 35	4.37 \pm 0.08	0.755 \pm 0.133	0.174 \pm 0.044	5255 \pm 24	4.44 \pm 0.06	0.767 \pm 0.105	0.070 \pm 0.031
Morel et al. (2014)	5265 \pm 31	4.35 \pm 0.09	0.795 \pm 0.102	0.197 \pm 0.031	5275 \pm 31	4.45 \pm 0.08	0.870 \pm 0.089	0.089 \pm 0.032
Reddy et al. (2003)	5320 \pm 38	4.51 \pm 0.11	0.900 \pm 0.062	0.218 \pm 0.036	5295 \pm 29	4.52 \pm 0.09	0.958 \pm 0.046	0.092 \pm 0.027
Tsantaki et al. (2019)	5190 \pm 64	4.26 \pm 0.09	0.590 \pm 0.149	0.163 \pm 0.048	5140 \pm 81	4.35 \pm 0.08	0.485 \pm 0.198	0.050 \pm 0.049

Notes. For iron, 42 Fe I and 4 Fe II lines were used.

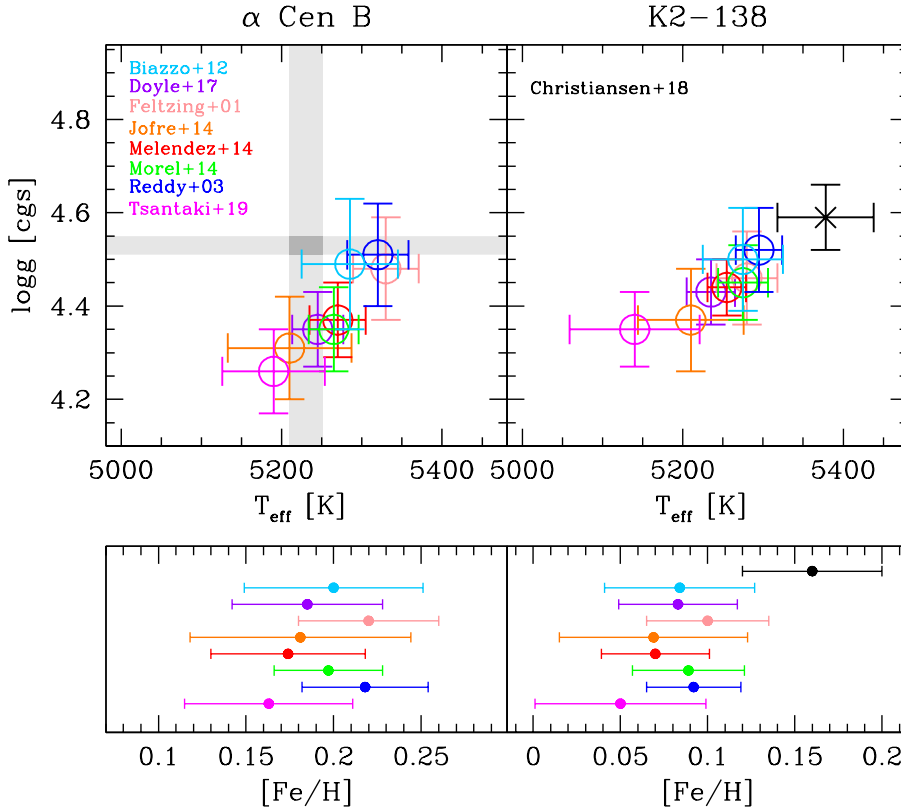


Fig. 1. Results of the analysis of α Cen B (left panels) and K2-138 (right panels) using the various iron line lists. The colour coding for each line list is indicated in the upper left panel. The parameters of K2-138 determined by Christiansen et al. (2018) are shown in the right panels. The grey-shaded areas for α Cen B delimit the interferometric T_{eff} and seismic $\log g$ values ($\pm 1 \sigma$; see Sect. 2.1 for details).

cooler and less metal rich than concluded by Christiansen et al. (2018).

2.2. Stellar abundances

We proceed for the abundance analysis with the extensive line list of Meléndez et al. (2014) because the lines of some important elements (e.g. Mg) in Biazzo et al. (2012) are not covered by our observations. A hyperfine structure was taken into account for Sc, V, Mn, Co, and Cu using atomic data from the Kurucz database³, while the Eu data were taken from Ivans et al. (2006).

³ Available at <http://kurucz.harvard.edu/linelists.html>

A classical curve-of-growth analysis making use of the EWs was performed for most species. However, the determination of some abundances relied on spectral synthesis. The oxygen abundance was based on [O I] λ 630.0, while the C abundance was also estimated from the C₂ lines at 508.6 and 513.5 nm. Readers can refer to Morel et al. (2014) for further details on the modelling of the [O I] and C₂ features. Finally, the Eu abundance was based on a synthesis of a number of Eu II lines (for details, see Wang et al. 2020). For K2-138, $v \sin i = 2.5$ and a macroturbulence of 1.9 km s^{-1} were assumed based on the analysis reported in Lopez et al. (2019). An attempt was made to model Li I λ 670.8. The line is not detected in K2-138, but the Li abundance appears to be much lower than solar.

Table 2. Abundance results for K2-138.

Abundance ratio	Default T_{eff} scale	Cooler T_{eff} scale
[Fe/H]	+0.08 ± 0.05 (42+4)	+0.01
[C I/Fe]	−0.04 ± 0.08 (3)	+0.03
[C ₂ /Fe]	−0.07 ± 0.09 (2)	−0.01
[O I/Fe]	+0.03 ± 0.10 (1)	−0.01
[Na I/Fe]	+0.02 ± 0.06 (3)	−0.04
[Mg I/Fe]	−0.06 ± 0.08 (3)	−0.05
[Al I/Fe]	+0.01 ± 0.05 (2)	−0.04
[Si I/Fe]	+0.01 ± 0.04 (10)	+0.00
[Ca I/Fe]	+0.04 ± 0.06 (3)	−0.05
[Sc I/Fe]	−0.03 ± 0.10 (4)	−0.06
[Sc II/Fe]	−0.01 ± 0.05 (5)	−0.01
[Ti I/Fe]	+0.01 ± 0.08 (14)	−0.07
[Ti II/Fe]	+0.01 ± 0.06 (10)	+0.00
[V I/Fe]	+0.03 ± 0.08 (5)	−0.07
[Cr I/Fe]	+0.03 ± 0.05 (7)	−0.04
[Cr II/Fe]	+0.08 ± 0.04 (4)	+0.01
[Mn I/Fe]	+0.04 ± 0.07 (5)	−0.05
[Co I/Fe]	+0.00 ± 0.06 (7)	−0.03
[Ni I/Fe]	+0.00 ± 0.04 (14)	−0.02
[Cu I/Fe]	−0.02 ± 0.03 (2)	−0.02
[Zn I/Fe]	−0.01 ± 0.03 (3)	+0.00
[Sr I/Fe]	+0.01 ± 0.09 (1)	−0.07
[Y II/Fe]	+0.02 ± 0.07 (4)	−0.01
[Zr II/Fe]	+0.06 ± 0.06 (2)	−0.02
[Ba II/Fe]	+0.02 ± 0.07 (1)	−0.02
[Ce II/Fe]	+0.01 ± 0.08 (5)	−0.02
[Nd II/Fe]	+0.07 ± 0.05 (3)	−0.02
[Eu II/Fe]	+0.04 ± 0.08 (3)	−0.02
[C I/O I]	−0.07 ± 0.13	+0.04
[C ₂ /O I]	−0.10 ± 0.12	+0.00
[Mg I/Si I]	−0.07 ± 0.08	−0.05

Notes. The last column shows the impact of lowering T_{eff} by 50 K (see Sect. 2.1), while keeping $\log g$ and ξ unchanged. The number in brackets gives the number of lines the abundance is based on. For iron, the number of Fe I and Fe II lines is given.

The abundances are provided in Table 2. The random uncertainties were estimated following Morel (2018). For the spectral synthesis, additional sources of errors (e.g. continuum placement) were taken into account (see Morel et al. 2014). The O abundance is based on a single line that is weak ($EW < 10 \text{ m}\text{\AA}$) and blended with a Ni line. It is therefore uncertain. The same is true for the Mg abundance that is based on three strong lines exhibiting quite a large line-to-line scatter ($\sim 0.05 \text{ dex}$).

The impact of lowering T_{eff} by 50 K (see Sect. 2.1) is also given in Table 2. The Sc, Ti, and Cr abundances were derived from both neutral and singly ionised species. Ionisation balance is fulfilled within the uncertainties in all cases assuming the default parameters. However, it can be noted that the agreement systematically degrades for the cooler T_{eff} scale.

3. PASTIS analysis

The joint analysis of the HARPS radial velocities, K2 light curve and spectral energy distribution (SED) was carried out using the Bayesian software PASTIS (Díaz et al. 2014). Improvements with respect to our previous analysis in Lopez et al. (2019) include the following: (1) the radial velocities were binned nightly to average out the correlated high-frequency noise resulting from granulation and instrumental calibrations, and (2) the new stellar parameters, as derived in Sect. 2.1, were used as priors. We ran two sets of analysis with the adopted T_{eff} and lowered by 50 K, as the latter cannot be ruled out, as reported in Sect. 2.1.

The magnitudes used to construct the SED were taken from the American Association of Variable Star Observers Photometric All-Sky Survey (Henden et al. 2015) archive in the optical, from the Two-Micron All-Sky Survey (Munari et al. 2014) and the Wide-field Infrared Survey Explorer (Cutri & et al. 2014) archives in the near-infrared. The SED was modelled with the BTSettl stellar atmospheric models (Allard et al. 2012). The radial velocities were modelled with Keplerian orbit models for the planetary contribution and with a Gaussian process regression for the correlated noise induced by the activity. For the latter, the following quasi-periodic kernel was used:

$$k(t_i, t_j) = A^2 \exp \left[-\frac{1}{2} \left(\frac{t_i - t_j}{\lambda_1} \right)^2 - \frac{2}{\lambda_2^2} \sin^2 \left(\frac{\pi |t_i - t_j|}{P_{\text{rot}}} \right) \right] + \delta_{ij} \sqrt{\sigma_i^2 + \sigma_j^2} \quad (1)$$

where A corresponds to the radial velocity modulation amplitude, P_{rot} to the stellar rotation period, λ_1 to the correlation decay timescale of the active regions, λ_2 to the relative contribution between the periodic and the decaying components, and σ_j to the radial velocity jitter. To model the photometry, we used the JKT Eclipsing Binary Orbit Program (Southworth 2008) with an oversampling factor of 30 to account for the long integration time of *Kepler* (Kipping 2010). The star was modelled with the PARSEC evolution tracks (Bressan et al. 2012), taking the asterodensity profiling into account (Kipping 2014), and with the limb darkening coefficients taken from Claret & Bloemen (2011).

We ran 80 Markov chain Monte Carlo (MCMC) chains with 10^6 iterations for the two different effective temperatures to explore the posterior distributions of the parameters. The convergence was assessed with a Kolmogorov–Smirnov test (Brooks et al. 2003). The burn-in phase was then removed (Díaz et al. 2014) and the remaining iterations of the different chains having converged were merged. Both analyses, with T_{eff} and T_{eff} lowered by 50 K, converged towards the same distributions, and in particular the same median effective temperature was found for both. Therefore we only report the posteriors for the analysis based on $T_{\text{eff}} = 5275 \text{ K}$, along with the priors used. These are shown in Table A.1.

The parameters obtained are fully compatible with that of Lopez et al. (2019). In particular, we found masses of $2.80^{+0.94}_{-0.96} M_{\oplus}$, $5.95^{+1.17}_{-1.12} M_{\oplus}$, $7.20 \pm 1.40 M_{\oplus}$, and $11.28^{+2.78}_{-2.72} M_{\oplus}$ for planets b, c, d, and e, respectively, giving a precision of 34%, 20%, 19%, and 25%. For planets f and g, the median values on the masses are $2.43^{+3.05}_{-1.75} M_{\oplus}$ and $2.45^{+2.92}_{-1.74} M_{\oplus}$, respectively, giving a significance of 1.4σ for both planets. For planet g, the non detection is not surprising given the relatively long orbital period for a planet with a radius compatible with a low-density planet. Conversely, for planet f, we cannot exclude absorption of the signal by the Gaussian process given that its orbital period

is half the stellar rotation period. Further discussion on the constraints and upper limits of the planetary masses can be found in Lopez et al. (2019). The parameters of the planets were then used as input for the planets' modelling described in the following section (see Table A.1).

4. Composition analysis

4.1. Interior-atmosphere model

We used the internal structure model initially developed by Brugger et al. (2017) and Mousis et al. (2020), and recently updated by Acuña et al. (2021) for their internal composition's study. The model can accommodate a surface water layer. To consider the effect of the stellar irradiation on this layer, we included a water-rich atmosphere on top of the high-pressure water layer or the mantle by coupling the interior to an atmosphere model. The atmospheric model computes the temperature at the bottom of the atmosphere, which is the boundary condition for the interior model. As a result, our current atmosphere-interior model allows us to assess in detail how well a close-in planet, similar to the ones we analyse in Sect. 5, can support a water-rich layer either in a liquid, vapour, or supercritical state depending on the surface temperature.

Our atmosphere-interior model takes the irradiation received by the planet into account and calculates the surface temperature assuming a water-rich atmosphere on top of a high-pressure water layer or a mantle. Therefore, in Sect. 5, we use the terms volatile mass fraction and water mass fraction interchangeably. The planets in the multi-planetary systems we analyse are highly irradiated, with irradiation temperatures ranging from approximately 1300 K to 500 K (see Table 5). Depending on the corresponding surface conditions, if water is present, it can be in a vapour or supercritical state.

The input variables of the interior structure model are the total planetary mass, the core mass fraction (CMF), and the water mass fraction (WMF), while the model outputs the total planetary radius and the Fe/Si mole ratio. In order to explore the parameter space, we performed a complete Bayesian analysis to obtain the probability density distributions of the parameters. This Bayesian analysis was carried out via the implementation of a MCMC algorithm, by adapting the method proposed by Dorn et al. (2015) to our interior and atmosphere model as described in Acuña et al. (2021).

Initial values of the three input parameters were randomly drawn from their prior distributions, which correspond to a Gaussian distribution for the mass, and uniform distributions for the CMF and the WMF. We established a maximum WMF in the uniform prior of 80% based on the maximum water content found in Solar System bodies (McKay et al. 2019). For the atmosphere, we have considered a composition of 99% water and 1% carbon dioxide. The atmosphere and the interior are coupled at a pressure of 300 bar. We considered the stellar spectral distribution of a Sun-like star for the calculation of the Bond albedo. The atmospheric mass, thickness, Bond albedo, and temperature at the bottom of the atmosphere are provided by a grid generated with the atmospheric model described in Marcq et al. (2017) and Pluriel et al. (2019).

4.2. Atmospheric escape

Atmospheric mass loss in super-Earths and sub-Neptunes can be produced by thermal or non-thermal escape, with Jeans escape (Jeans 1925), XUV photoevaporation (Owen & Jackson 2012), or

core-powered mass loss (Ginzburg et al. 2016). These processes might shape the trend of the volatile mass fraction (water, H/He, or a combination of both) in the inner region of multi-planetary systems. An estimate of the mass loss rates of different species can help to discriminate between two possible interior compositions. In our Solar System, Jeans' escape efficiently removed lighter gases as H₂ and He on telluric planets, leaving heavier molecules. For the planets in the K2-138 system, we estimated Jeans mass loss rates (Aguichine et al. 2021) by using the masses, radii, and equilibrium temperatures we obtained as a result of our spectroscopic analysis as input (Sect. 2). For the rest of the multi-planetary systems we analysed, we used the parameters provided by the references we mention in Sect. 4.3.

The hydrodynamic escape of H-He is driven by the incident XUV flux from the host star. A star's XUV luminosity L_{XUV} is usually constant at early stages, called a saturation regime (a few tens of megayears), and then it evolves as a power-law function of time $L_{\text{XUV}} \propto t^\alpha$, with $\alpha \approx -1.5$ (Sanz-Forcada et al. 2011). Computing the mass loss rate from Owen & Jackson (2012), we get the following:

$$\dot{m} = \eta \frac{L_{\text{XUV}} R_b^3}{GM_b (2a_b)^2}, \quad (2)$$

where G is the gravitational constant and $\eta=0.1$ is an efficiency factor (Owen & Jackson 2012). Following the approach in Aguichine et al. (2021), we integrated Eq. (2) over time assuming that only L_{XUV} can vary, implying the mass and radius do not change significantly, to calculate the total lost mass.

4.3. Multi-planetary systems parameters

In addition to K2-138, we selected a sample of multi-planetary systems that host only low-mass planets ($M < 20 M_\oplus$), with five or more planets that have masses and radii available. These systems are TOI-178, Kepler-11, Kepler-102, and Kepler-80. For K2-138, we took the planetary mass and radius derived in Sect. 3, and the corrected Fe/Si molar ratio. The latter was estimated as Fe/Si = 0.77 ± 0.07 , using the metallicity and the Mg, Al, Si, Ca, and Ni abundances presented in Sect. 2.2, following Sotin et al. (2007) and Brugger et al. (2017).

For the other systems, we performed the same modelling, taking masses, radii, and stellar abundances from Leleu et al. (2021) for TOI-178; Lissauer et al. (2011) and Brewer et al. (2016) for Kepler-11; Marcy et al. (2014) and Brewer & Fischer (2018) for Kepler-102; and MacDonald et al. (2016, 2021) for Kepler-80. We show a summary of the parameters we used in Table 3. The Fe/Si mole ratios of these systems were computed similarly to the Fe/Si mole ratio of K2-138 from their respective host stellar abundances.

5. Compositional trends in multi-planetary systems

Table 4 shows the retrieved CMF and WMF and their one-dimensional 1σ uncertainties as a result of our Bayesian analysis, as well as their atmospheric mass loss estimates. To assess how compatible a water-rich composition is with the data, we also show the difference between the observational mean and the retrieved mean, which is calculated as $d_{\text{obs-ret}} = \max\{|R_{\text{data}} - R|, |M_{\text{data}} - M|\}$. If $d_{\text{obs-ret}}$ is below 1σ , the retrieved mass and radius agree within the 1σ confidence intervals with the observed mass and radius, meaning that the density of a planet is compatible with a volatile layer dominated by water. A high $d_{\text{obs-ret}}$ ($>1 \sigma$) and a high WMF in our model

Table 3. Masses, radii, semi-major axis, and irradiation temperature for the multi-planetary systems TOI-178, Kepler-11, Kepler-102, and Kepler-80.

System	Planet	$M (M_{\oplus})$	$R (R_{\oplus})$	a_d (AU)	T_{irr} (K)
TOI-178	b	$1.5^{+0.39}_{-0.44}$	$1.152^{+0.073}_{-0.070}$	0.026	1040
	c	$4.77^{+0.55}_{-0.68}$	$1.669^{+0.114}_{-0.099}$	0.037	873
	d	$3.01^{+0.80}_{-1.03}$	$2.572^{+0.075}_{-0.078}$	0.059	691
	e	$3.86^{+1.25}_{-0.94}$	$2.207^{+0.088}_{-0.090}$	0.078	600
	f	$7.72^{+1.67}_{-1.52}$	$2.287^{+0.108}_{-0.110}$	0.104	521
	g	$3.94^{+1.31}_{-1.62}$	$2.87^{+0.14}_{-0.13}$	0.128	471
Kepler-11	b	$4.3^{+2.2}_{-2.0}$	1.97 ± 0.19	0.091	953
	c	$13.5^{+4.8}_{-6.1}$	3.15 ± 0.30	0.106	883
	d	$6.1^{+3.1}_{-1.7}$	3.43 ± 0.32	0.159	721
	e	$8.4^{+2.5}_{-1.9}$	4.52 ± 0.43	0.194	653
	f	$2.3^{+2.2}_{-1.2}$	2.61 ± 0.25	0.250	575
Kepler-102	b	0.41 ± 1.6	0.47 ± 0.02	0.055	868
	c	-1.58 ± 2.0	0.58 ± 0.02	0.067	786
	d	3.80 ± 1.8	1.18 ± 0.04	0.086	597
	e	8.93 ± 2.0	2.22 ± 0.07	0.117	694
	f	0.62 ± 3.3	0.88 ± 0.03	0.165	501
Kepler-80	d	$5.95^{+0.65}_{-0.60}$	$1.309^{+0.036}_{-0.032}$	0.033	990
	e	$2.97^{+0.76}_{-0.65}$	$1.330^{+0.039}_{-0.038}$	0.044	863
	b	$3.50^{+0.63}_{-0.57}$	$2.367^{+0.055}_{-0.052}$	0.058	750
	c	$3.49^{+0.63}_{-0.57}$	$2.507^{+0.061}_{-0.058}$	0.071	679
	g	$0.065^{+0.044}_{-0.038}$	$1.05^{+0.22}_{-0.24}$	0.094	588

Notes. References can be found in Sect. 4.3.

Table 4. Retrieved core mass fraction (CMF) and water mass fraction (WMF) of planets in the multi-planetary systems K2-138, TOI-178, Kepler-11, Kepler-102, and Kepler-80, with our interior-atmosphere model.

System	Planet	CMF	WMF	$d_{\text{obs-ret}}$	$\Delta M_{\text{H}_2} (M_{\oplus})$	$\Delta M_{\text{H}_2\text{O}} (M_{\oplus})$	$\Delta M_{\text{XUV}} (M_{\oplus})$
K2-138	b	0.27 ± 0.02	$0.000^{+0.007}_{-0.000}$	1.5σ	0.132	<0.01	0.40
	c	0.23 ± 0.02	0.13 ± 0.04	<1 σ	<0.01	<0.01	<0.01
	d	0.22 ± 0.03	0.17 ± 0.05	<1 σ	<0.01	<0.01	<0.01
	e	0.11 ± 0.02	0.57 ± 0.08	<1 σ	<0.01	<0.01	<0.01
	f	0.11 ± 0.02	0.60 ± 0.07	<1 σ	<0.01	<0.01	<0.01
	g	0.12 ± 0.05	0.55 ± 0.18	1.3σ	<0.01	<0.01	<0.01
TOI-178	b	0.21 ± 0.30	0	<1 σ	0.83	<0.01	0.45
	c	0.30 ± 0.02	$0.02^{+0.04}_{-0.02}$	<1 σ	<0.01	<0.01	0.21
	d	0.10 ± 0.01	0.69 ± 0.05	1.3σ	0.16	<0.01	0.48
	e	0.18 ± 0.02	0.40 ± 0.06	<1 σ	<0.01	<0.01	0.13
	f	0.22 ± 0.03	0.28 ± 0.10	<1 σ	<0.01	<0.01	0.04
	g	0.10 ± 0.01	0.58 ± 0.16	3.0σ	<0.01	<0.01	0.11
Kepler-11	b	0.20 ± 0.04	0.27 ± 0.10	<1 σ	<0.01	<0.01	0.10
	c	0.18 ± 0.01	0.33 ± 0.04	1.7σ	<0.01	<0.01	0.10
	d	0.10 ± 0.02	0.65 ± 0.05	2.4σ	<0.01	<0.01	0.13
	e	0.12 ± 0.01	0.55 ± 0.04	4.4σ	<0.01	<0.01	0.14
	f	0.14 ± 0.06	0.47 ± 0.10	1.9σ	0.56	<0.01	0.06
Kepler-102	b	$0.91^{+0.09}_{-0.16}$	0	<1 σ	0.13	<0.01	0.03
	c	$0.95^{+0.05}_{-0.30}$	0	<1 σ	0.10	<0.01	0.03
	d	0.80 ± 0.14	0	<1 σ	<0.01	<0.01	0.03
	e	0.22 ± 0.02	0.17 ± 0.07	<1 σ	0.01	<0.01	0.03
	f	0.27 ± 0.09	0.04 ± 0.04	<1 σ	0.02	<0.01	0.01
Kepler-80	d	$0.97^{+0.03}_{-0.05}$	0	<1 σ	<0.01	<0.01	0.35
	e	0.43 ± 0.18	0	<1 σ	<0.01	<0.01	0.29
	b	0.13 ± 0.02	0.58 ± 0.07	<1 σ	<0.01	<0.01	0.11
	c	0.09 ± 0.01	0.70 ± 0.04	<1 σ	<0.01	<0.01	0.13
	g	0.31 ± 0.02	$<1.5 \times 10^{-3}$	<1 σ	140	3.23	0.60

Notes. A low $d_{\text{obs-ret}}$ indicates that the assumption of a water-dominated atmosphere is adequate for a particular planet (see text). We note that ΔM_{H_2} , $\Delta M_{\text{H}_2\text{O}}$, and ΔM_{XUV} correspond to the maximum estimate of atmospheric escape mass loss due to H_2 , water Jeans escape, and XUV photoevaporation, respectively.

Table 5. Atmospheric parameters retrieved for the planets whose composition can accommodate a water-dominated atmosphere (see text).

Planet	T_{irr} (K)	T_{300} (K)	z_{atm} (km)	A_B
K2-138 b	1291	4110 ± 44	932 ± 151	0.213 ± 0.001
K2-138 c	1125	3900 ± 23	711 ± 103	0.214 ± 0.002
K2-138 d	978	3614 ± 56	635 ± 84	0.218 ± 0.002
K2-138 e	850	3383 ± 39	673 ± 90	0.231 ± 0.001
K2-138 f	735	3396 ± 116	1483 ± 546	0.260 ± 0.004
TOI-178 c	873	3344 ± 33	500 ± 60	0.226 ± 0.001
TOI-178 d	691	3254 ± 45	1181 ± 224	0.264 ± 0.004
TOI-178 e	600	2930 ± 31	690.7 ± 133	0.225 ± 0.018
TOI-178 f	521	2610 ± 23	368 ± 60	0.298 ± 0.007
Kepler-11 b	953	3697 ± 133	840 ± 313	0.221 ± 0.005
Kepler-102 e	694	2947 ± 29	360 ± 55	0.243 ± 0.004
Kepler-102 f	501	2784 ± 102	837 ± 290	0.347 ± 0.013
Kepler-80 b	750	3344 ± 33	1133 ± 148	0.253 ± 0.002
Kepler-80 c	679	3219 ± 29	1128 ± 114	0.266 ± 0.003

Notes. These parameters are the equilibrium temperature assuming a null albedo (T_{irr}), the atmospheric temperature at 300 bar (T_{300}), the thickness of the atmosphere from the 300 bar to 20 mbar (z_{atm}), and the planetary Bond albedo (A_B).

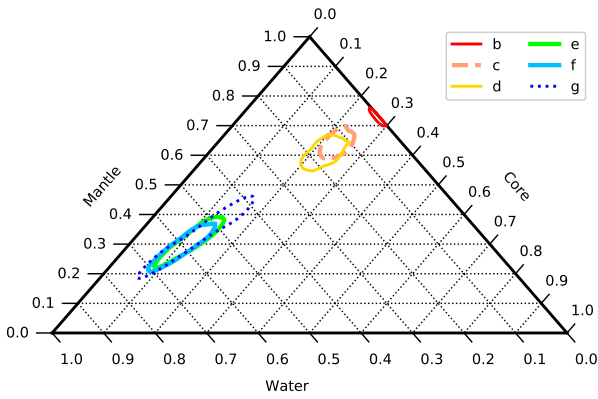


Fig. 2. $1\text{-}\sigma$ confidence regions derived from the 2D posterior distributions of the CMF and WMF obtained with the planetary interior Bayesian analysis. Axes indicate the core mass fraction (CMF), water mass fraction (WMF), and the mantle mass fraction (MMF). The latter is defined as $\text{MMF} = 1 - (\text{CMF} + \text{WMF})$.

simultaneously indicate that a water-dominated atmosphere is not inflated enough to account for the low density of the planet, pointing to an atmosphere with more volatile gases, which are probably H and He. Table 5 shows the irradiation temperatures and the retrieved atmospheric parameters of the planets whose density is compatible with the presence of a volatile layer dominated by water.

5.1. K2-138

Figure 2 displays the 1σ confidence intervals derived from the 2D distributions of the WMF and CMF of the K2-138 in a ternary diagram. We can see that the confidence regions are aligned along a line almost parallel to the lines where the CMF is constant. This alignment is due to the constraint on the Fe/Si mole ratio we have considered within the whole planetary system: the confidence regions are spread over the Fe/Si-isolines whose constant values range from Fe/Si=0.70 to 0.84 (see Brugger et al. 2017, their Fig. 4).

For K2-138 b, the results set an upper limit of 0.7% in the WMF, which means that this planet is unlikely to have a significant amount of volatiles, including water. The retrieved planetary radius is $1.538 R_{\oplus}$, which is 1.5σ larger than the measured radius from the analysis in Sect. 3. This is due to the extended atmosphere necessary to produce temperature and pressure conditions to hold supercritical water on the surface ($P_{\text{surf}} > 300$ bar). If we assume a mass of $2.80 M_{\oplus}$ and a CMF of 0.27, a vapour atmosphere with a maximum surface pressure of 300 bar would yield a WMF of 0.01% (WMF of Earth is 0.05%) and a radius of $1.461 R_{\oplus}$, which is well within the 1σ confidence interval of the observed value. Therefore, we can conclude that K2-138 b is a volatile-poor planet, which might present a secondary atmosphere with a low surface pressure ($P_{\text{surf}} \leq 300$ bar) or no atmosphere (WMF=0). In addition, it is the planet with the highest CMF in the system, showing that planets in this system are likely to have less massive cores than Earth (CMF = 0.325) and the other terrestrial planets in the Solar System.

The atmospheric model also establishes a minimum surface gravity of 2 m s^{-1} to retain an atmosphere. Unlike planets b, c, d, and e, in which the $1\text{-}\sigma$ intervals on the masses exclude such low surface gravity, this is not the case for planets f and g. For planet f, a lower limit on the surface gravity of the planet can be translated to a lower limit on the mass. If it is below this limit, the gravity at the surface would not be enough to retain an atmosphere. For planet f, with a total radius of $2.762 R_{\oplus}$ and a CMF of 0.11, this limit would be approximately $2 M_{\oplus}$. This minimum mass value to retain its atmosphere is above the lower limit of the total mass set by its 1σ uncertainties, as can be seen in the upper panel of Fig. 3. Furthermore, planet f is the most water-rich in the K2-138 planetary system, with an upper limit of 66% in the WMF, which is close to the 77% maximum limit on the water content derived from measurements on cometary compositions. Similarly, planet g also presents a lower limit on the mass of the bulk of the planet of $\sim 2 M_{\oplus}$ (see Fig. 3, lower panel). Its retrieved planetary radius is significantly lower than the observational value, with a difference of 1.3σ . Therefore, the atmosphere of K2-138 g is significantly more extended than an atmosphere dominated by water vapour under the same irradiance conditions. This increase in atmospheric thickness is probably due to an atmosphere rich in H and He. K2-138 g could

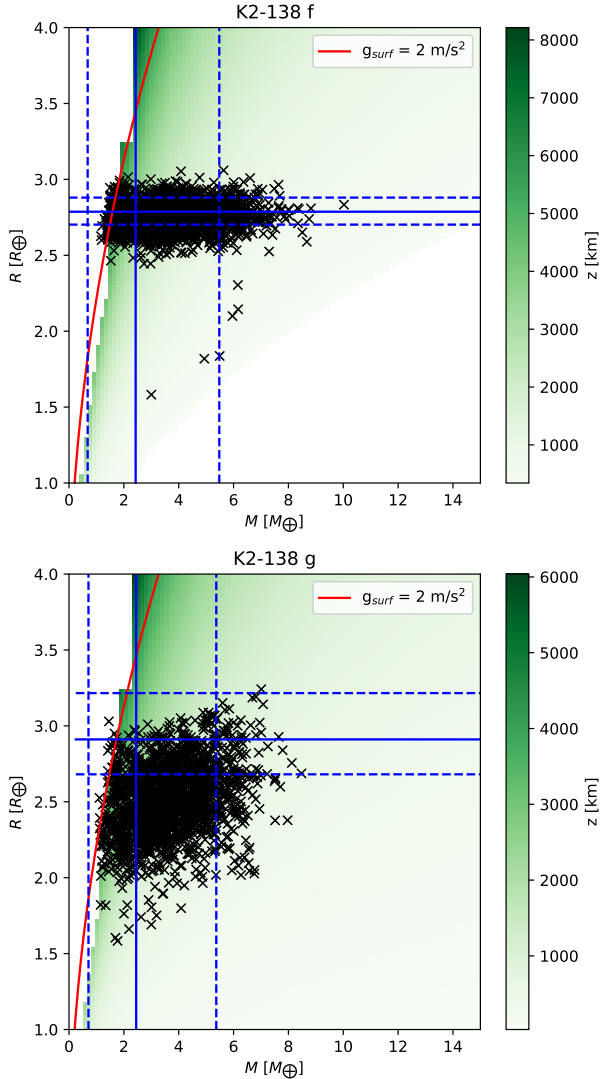


Fig. 3. Total mass and radius of K2-138 f (*upper panel*) and K2-138 g (*lower panel*) from the different realisations of the MCMC (black crosses). The solid blue lines show the mass and radius measurements from PASTIS, and the dashed lines give the related uncertainties. The red line indicates the limit below which the planet cannot maintain an atmosphere.

have up to 5% of the volatile mass fraction assuming a H/He atmosphere (see Fig. 1 in Lopez & Fortney 2014).

A rough estimate of Jeans mass loss rates for K2-138 b yields $6 \times 10^{-7} M_{\oplus} \text{Gyr}^{-1}$ for Jeans escape of H_2 , and $5 \times 10^{-84} M_{\oplus} \text{Gyr}^{-1}$ for Jeans escape of H_2O . For comparison, in the case of Earth, the absence of H_2 is due to an exobase (altitude at which particles escape) temperature much higher than the equilibrium temperature (Hedin 1983). An exobase temperature 2 times higher than the equilibrium temperature gives a mass-loss rate of $4 \times 10^{-2} M_{\oplus} \text{Gyr}^{-1}$. In that case, an envelope of 1–10% of the H-He mixture could be efficiently removed, leaving only heavier species such as H_2O . In the case of hydrodynamic escape, we obtained a mass loss rate of $2 M_{\oplus} \text{Gyr}^{-1}$ during the saturation regime and $1 \times 10^{-2} M_{\oplus} \text{Gyr}^{-1}$ at $t = 3 \text{ Gyr}$. This yields an integrated mass loss of $0.4 M_{\oplus}$, or 14% of planet’s

b total mass. Comparing this value to the WMF derived for planets c and d from the MCMC in Table 4, we conclude that K2-138 b could have formed with a thick envelope of H_2O that has been blown away by XUV photoevaporation.

5.2. TOI-178

In the TOI-178 system, planets b and c have an increasing WMF with a progressing distance from the star, while planets d to g have a WMF equal or greater than 30%. For planets d and g, the volatile layer is likely to present H/He, which would explain why in our analysis their WMFs are in the 60–70% range in addition to $d_{\text{obs-ret}}$ being greater than 1σ . TOI-178 b could have lost up to $0.83 M_{\oplus}$ of its current mass in H_2 due to Jeans escape, and up to $0.45 M_{\oplus}$ due to photoevaporation, while TOI-178 c could have lost $0.21 M_{\oplus}$. In such a scenario, the TOI-178 b and c original volatile mass fraction would be up to 0.36 and 0.10, respectively, compared to their current value.

5.3. Kepler-11

For Kepler-11, the WMF of the innermost planet is 0.27 ± 0.10 , which is compatible with a water-dominated envelope. For Kepler-11 c to e, their radius data are 1.7σ , 2.4σ , and 4.4σ higher than the radius we retrieved with our model, discarding the water-rich envelope hypothesis. The increasing significance level indicates that these planets have an increasing content of H/He with distance from the star. In the case of the outermost planet, Kepler-11 f, the retrieved radius is 1.9σ lower than the data, suggesting that this planet presents less H/He than planets c to e. Nonetheless, this could be because Kepler-11 f not being able to retain a primordial atmosphere due to its low mass ($2.3^{+2.2}_{-1.2} M_{\oplus}$), compared to the higher masses of the rest of the planets in the system ($>6 M_{\oplus}$). Furthermore, Kepler-11 f could have lost up to $0.56 M_{\oplus}$ in H_2 , according to our atmospheric Jeans escape calculation, whereas the other four planets in the system have atmospheric mass losses below $2 \times 10^{-3} M_{\oplus}$.

5.4. Kepler-102

The densities of the three innermost planets of Kepler-102 suggest that these are dry planets with high CMFs. Their core-to-mantle ratios could be even higher than the CMF we would expect from the Fe and Si stellar abundances of their host star. Therefore, we set the WMF equal to zero in our MCMC Bayesian analysis and let the CMF be the only free parameter. We only took the mass and radius into account as observables. Our modelling shows that Kepler-102 b, c, and d are dry Mercury-like planets, with $\text{CMF} = 0.91^{+0.09}_{-0.16}$, $0.95^{+0.05}_{-0.30}$, and 0.80 ± 0.14 , respectively. Their high CMF could be due to mantle evaporation (Cameron 1985), impacts (Benz et al. 1988, 2007; Asphaug & Reufer 2014), or planet formation in the vicinity of the rock-lines (Aguichine et al. 2020; Scora et al. 2020). Kepler-102 e presents a WMF of 0.17 ± 0.07 , suggesting that this planet has a more volatile-rich composition than the planets that precede it. The large uncertainties in the mass of Kepler-102 f prevent us from determining whether this is a bare rocky planet with no atmosphere, or if it presents a thin atmosphere with a maximum $\text{WMF} = 0.08$. In addition, Jeans H_2 atmospheric escape could have removed up to $0.02 M_{\oplus}$ from Kepler-102 f, yielding an original volatile mass fraction between 0.07 and 0.10.

5.5. Kepler-80

Kepler-80 d presents a high CMF, corresponding to a Fe-rich planet, similarly to Kepler-102 b and c. Kepler-80 e is consistent

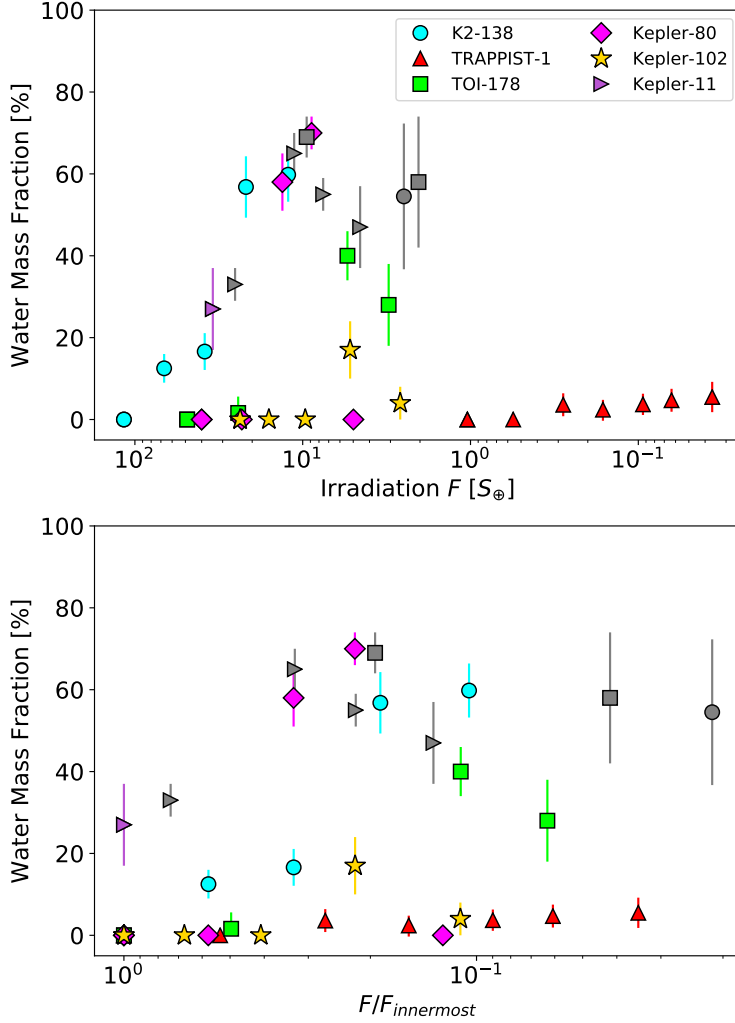


Fig. 4. Volatile mass fraction trends of the six multi-planetary systems analysed with our interior-atmosphere model. We show the water mass fraction estimates (see text) as a function of the stellar incident flux or irradiation, F , in Earth irradiation units ($S_{\oplus} = 1361 \text{ W/m}^2$) in the upper panel. In the lower panel, the incident flux is normalised with respect to the inner, most irradiated planet in each system, $F_{\text{innermost}}$. The planets whose atmospheric composition is likely to be H/He-dominated instead of water-dominated ($d_{\text{obs-ret}} > 1 \sigma$) are indicated in grey.

with a dry planet with an Earth-like CMF, whereas Kepler-80 b and c are volatile-dominated planets. Kepler-80 g shows a WMF of up to 0.15%. Given its low mass $M = 0.065^{+0.044}_{-0.038} M_{\oplus}$ (MacDonald et al. 2021), planet g could not have retained a H/He atmosphere, making a secondary atmosphere with water and/or CO₂ the most likely atmospheric composition for this planet. Based on our MCMC interior-atmosphere analysis, this atmosphere could be of less than 300 bar of surface pressure. This scenario is also supported by our estimated Jeans water escape, which is between $3.26 \times 10^{-3} M_{\oplus}$ and $3.24 M_{\oplus}$. Both Jeans escape and XUV photoevaporation could have removed a H/He envelope efficiently. The total atmospheric mass loss and the current mass add up to a planetary mass that is similar to that of Kepler-80 e, b, and c. Finally, the radius of Kepler-80 g is 2.7σ higher than the radius of a rocky planet with no atmosphere, which suggests that Kepler-80 g probably has retained a gaseous envelope.

6. Discussion

Figure 4 shows the volatile content of the five multi-planetary systems we analysed in this work as a function of the incident

flux normalised with the incident flux received by the innermost planet. In addition, we include in Fig. 4 the WMF of TRAPPIST-1 derived with our interior-atmosphere model by Acuña et al. (2021) for a homogeneous comparison. Of all systems, K2-138 presents a very clear volatile mass fraction trend: an increasing gradient in water content with distance from the host star for planets b to d, followed by a constant volatile mass fraction for the outer planets (planets e to g). A similar trend is observed in the TRAPPIST-1 system if one neglects TRAPPIST-1 d, presenting a higher volatile mass fraction than its two surrounding inner and outer planets in Fig. 4. In Acuña et al. (2021), the WMF was obtained by assuming a condensed water layer. However, water could be in vapour phase and mixed with CO₂ in a CO₂-dominated atmosphere, lowering the overall volatile mass fraction of TRAPPIST-1 d. In that case, the TRAPPIST-1 system could potentially show the increase-plus-plateau volatile trend observed in K2-138. Transmission spectroscopy of TRAPPIST-1 d is needed to probe the composition of its atmosphere. The multi-planetary systems TOI-178 and Kepler-11 do not show a smooth increases in the water mass fraction with orbital distances, although the inner planets present significantly less volatiles than the outer planets. Finally, Kepler-80 and Kepler-102 could form this trend if it was not because of their outermost

planet, which presents a lower volatile mass fraction than the planet that immediately precedes it. In addition, the estimated original volatile mass fraction of Kepler-102 f is well within the uncertainties of the WMF of Kepler-102 e, meaning that planets e and f could potentially form a plateau in the outer part of the Kepler-102 system with a water mass fraction of 10%, similarly to TRAPPIST-1.

In the case of TOI-178 and Kepler-11, it would be necessary to adopt a self-consistent modelling approach that includes the possibility of a H/He-dominated volatile layer to determine whether their volatile mass fraction trend is as clear as that of K2-138 and TRAPPIST-1. For the other multi-planetary systems, which do not present high $d_{\text{obs-ret}}$ combined with high water mass fractions in our analysis, the volatile mass fraction would decrease for each individual planet under the assumption of a H/He envelope. Including H/He as part of the envelope would change the value of the volatile mass fraction of each individual planet, but it would not change our conclusion about the global volatile mass fraction trends in each system (i.e. the gradient and plateau trend in TRAPPIST-1 and K2-138). Furthermore, the water-H/He degeneracy to which volatile-rich planets are subject can only be broken with atmospheric characterisation data, such as transmission spectroscopy and phase curves. In many cases, the volatile envelope of sub-Neptunes might not be dominated by either water or H/He, but it could be a mixture of both. This is supported by transmission spectroscopy of the sub-Neptune K2-18 b (Tsiaras et al. 2019; Benneke et al. 2019; Madhusudhan et al. 2020), where water is detected; although its current trace species could be compatible with a H₂-rich atmosphere (Yu et al. 2021). Additionally, meteorite outgassing experiments show that a significant fraction of H/He could be sustained in a water-dominated secondary atmosphere (Thompson et al. 2021).

The significant difference in the volatile mass fraction between the inner planets and the outer planets of these multi-planetary systems indicates that these planets might have undergone similar formation and evolution histories. The gradient-plus-plateau trend could potentially result from the combination of planetary formation in ice-rich regions of the protoplanetary disc, atmospheric loss, and inward migration. The outer volatile-rich planets could have formed beyond the ice line prior to migration, where ice-rich solids are expected to form (Mousis et al. 2021), producing planets with high volatile contents. In the systems whose planets present water mass fractions lower than 10%, volatiles could have been simply delivered by building blocks made of chondritic minerals bearing this amount of water (Melwani Daswani et al. 2021). In those conditions, the radial drift of icy planetesimals from beyond the snowline is not required. In the case of K2-138, the three-body Laplace resonances are a sign of inner planetary migration (Terquem & Papaloizou 2007; Izidoro et al. 2017; Ramos et al. 2017). For three systems, we found that their outermost planets (Kepler-11 f, Kepler-102 f, and Kepler-80 g) have lower volatile mass fractions than the planets before them in the system. This could be due to their lower masses compared to the other planets in their systems, since they are not massive enough to have a surface gravity that would help them retain their atmospheres. In addition, these three low-mass, low-WMF planets could have formed further away from the water ice line than the water-rich planets in their systems, having less water-rich material available during accretion than those planets that formed in the vicinity of the water ice line.

In contrast to K2-138, the water mass fractions of the outer planets found in the TRAPPIST-1 and Kepler-102 systems are

compatible with 10% (Agol et al. 2021; Acuña et al. 2021), a value found in agreement with the water content of many asteroids of the Main Belt (Vernazza et al. 2015). This similarity suggests that the building blocks of the outer planets of these systems could have agglomerated from a mixture of ice grains coming from the snowline and anhydrous silicates that formed at closer distances from the host star, following the classical formation scenarios invoked for the Main Belt (Rivkin et al. 2002). In that case, this implies that the migration distances of the planets in TRAPPIST-1 and Kepler-102 would have been more restricted than those of the water-rich planets in the K2-138, TOI-178, and Kepler-11 systems.

We have considered the Fe/Si mole ratio as an observable of our MCMC Bayesian analysis in addition to the planetary masses and radii. Even though the Fe/Si derived from stellar abundances and that obtained from rocky planet densities could depart from a 1:1 relationship (Plotnykov & Valencia 2020; Adibekyan et al. 2021), considering the Fe/Si mole ratio contributes to reducing the degeneracy between the rock+mantle layers and the volatile layer (Dorn et al. 2015, 2017; Brugger et al. 2017). Particularly, assuming that the planetary Fe/Si mole ratio is similar to the Fe/Si ratio of the host star improves the determination of the CMF, but it does not necessarily contribute to the determination of the volatile mass fraction in volatile-rich planets (Otegi et al. 2020). This is the case of the TRAPPIST-1 system, where the inclusion of the Fe/Si mole ratio as an observable in the MCMC Bayesian analysis refines the determination of the surface pressure for the inner planets of the system, but slightly reduces the uncertainties of the WMF estimates for the outer planets (see Tables 3 and 4 in Acuña et al. 2021). Therefore, considering the Fe/Si mole ratio does not affect the volatile general trend of the planets within a multi-planetary system.

7. Conclusions

We carried out a homogeneous interior modelling and composition analysis of five multiplanetary systems that have five or more low-mass planets ($M < 20 M_{\oplus}$), rather than compiling the volatile content estimates of previous works to eliminate the differences between interior models as a possible bias when comparing the compositional trends between planetary systems. In the case of the TOI-178, Kepler-11, Kepler-102, and Kepler-80 systems, we used previously published mass, radius, and stellar abundance data. In the case of the K2-138 system, we completed the previous analysis with an in-depth stellar spectroscopic analysis. We performed a line-by-line differential analysis of K2-138 spectra with respect to α Cen B and the Sun to derive the most accurate stellar parameters and abundances given the available data. These were used for a new complete Bayesian analysis of the radial velocities and photometry acquired for the system. We explored the robustness of the planetary parameters and stellar chemical abundances in our spectroscopic analysis. We concluded that the parameters we derived are fully consistent with the ones obtained by Lopez et al. (2019).

With our interior-atmosphere model in a MCMC framework, we obtained the posterior distribution of the compositional parameters (CMF and WMF) and the atmospheric parameters assuming a water-dominated volatile layer of each of the planets in these multi-planetary systems. We found that K2-138 and TRAPPIST-1 present a very clear volatile trend with distance from the host star. Kepler-102 could potentially present

this trend. For the TOI-178 and Kepler-11 systems, our modelling ruled out the presence of a large hydrosphere being responsible for their low density. For such systems, it would be necessary to include H/He as part of the volatile layer in a self-consistent interior-atmosphere model. Nonetheless, all multi-planetary systems showed that the volatile mass fraction is significantly lower for the inner planets than for the outer planets. This is consistent with a formation history that involves formation of the outer planets in the vicinity of the ice line, inward migration, and atmospheric loss of the inner planets. We discussed the possible formation and evolution pathways that might yield these volatile content trends case by case. Similarly, we also commented on the possible causes of the high core mass fractions of the inner planets of Kepler-102 and Kepler-80, which might involve formation in the vicinity of the rocklines.

In addition, the atmospheric thickness that we obtained as a result of our Bayesian analysis (see Table 5) can be used to estimate the scale height of the extended atmospheres of the planets analysed in this work, which is necessary to assess the observing time and number of transits to characterise the composition of these atmospheres with transmission spectroscopy. This would confirm the exact composition of their atmospheres. To better assess possible evolutionary effects on the current composition of the planet, future work should involve the inclusion of atmospheric mass loss processes in the coupled atmosphere-interior model. In this work, we assumed that the planets do not evolve with time. The variation in the water mass fraction could also have been shaped by post-formation processes such as hydrodynamic escape (Bonfanti et al. 2021). Each of the discussed processes has been studied individually with interior models to constrain whether the atmospheres of low-mass planets are primordial or secondary (Dorn & Heng 2018; Gupta & Schlichting 2021), but none has modelled the effects of all these combined processes on the volatile reservoir of low-mass planets.

Acknowledgements. We would like to thank Maria Bergemann and Matthew Raymond Gent for a preliminary analysis of the stellar spectrum. This research has made use of the services of the ESO Science Archive Facility. This research was made possible through the use of the AAVSO Photometric All-Sky Survey (APASS), funded by the Robert Martin Ayers Sciences Fund. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. This paper includes data collected by the K2 mission. Funding for the K2 mission is provided by the NASA Science Mission directorate. This research has made use of the Exoplanet Follow-up Observation Program website, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. This research has made use of NASA's Astrophysics Data System Bibliographic Services. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. This research has made use of the VizieR catalogue access tool, CDS, Strasbourg, France. The original description of the VizieR service was published in Ochsenbein et al. (2000). T.M. acknowledges financial support from Belpo for contract PRODEX PLATO mission development. We acknowledge the anonymous referee whose comments helped improve and clarify this manuscript.

References

- Acuña, L., Deleuil, M., Mousis, O., et al. 2021, *A&A*, 647, A53
 Adibekyan, V., Dorn, C., Sousa, S. G., et al. 2021, *Science*, 374, 330
 Agol, E., Dorn, C., Grimm, S. L., et al. 2021, *Planet. Sci. J.*, 2, 1
 Aguichine, A., Mousis, O., Devouard, B., & Ronnet, T. 2020, *ApJ*, 901, 97
 Aguichine, A., Mousis, O., Deleuil, M., & Marcq, E. 2021, *ApJ*, 914, 84
 Allard, F., Homeier, D., & Freytag, B. 2012, *Philos. Trans. R. Soc. Lond. A*, 370, 2765
 Asphaug, E., & Reufer, A. 2014, *Nat. Geosci.*, 7, 564
 Bedell, M., Meléndez, J., Bean, J. L., et al. 2014, *ApJ*, 795, 23
 Benneke, B., Wong, I., Piaulet, C., et al. 2019, *ApJ*, 887, L14
 Benz, W., Slattery, W. L., & Cameron, A. G. W. 1988, *Icarus*, 74, 516
 Benz, W., Anic, A., Horner, J., & Whitby, J. A. 2007, *Space Sci. Rev.*, 132, 189
 Biazzo, K., D'Orazi, V., Desidera, S., et al. 2012, *MNRAS*, 427, 2905
 Bonfanti, A., Fossati, L., Kubyshkina, D., & Cubillos, P. E. 2021, *A&A*, 656, A157
 Boyajian, T. S., von Braun, K., van Belle, G., et al. 2013, *ApJ*, 771, 40
 Bressan, A., Marigo, P., Girardi, L., et al. 2012, *MNRAS*, 427, 127
 Brewer, J. M., & Fischer, D. A. 2018, *ApJS*, 237, 38
 Brewer, J. M., Fischer, D. A., Valenti, J. A., & Piskunov, N. 2016, *ApJS*, 225, 32
 Brooks, S. P., Giudici, P., & Philippe, A. 2003, *J. Comput. Graph. Stat.*, 12, 1
 Brugger, B., Mousis, O., Deleuil, M., & Deschamps, F. 2017, *ApJ*, 850, 93
 Cameron, A. G. W. 1985, *Icarus*, 64, 285
 Carter, J. A., Agol, E., Chaplin, W. J., et al. 2012, *Science*, 337, 556
 Castelli, F., & Kurucz, R. L. 2003, in *Modelling of Stellar Atmospheres*, eds. N. Piskunov, W. W. Weiss, & D. F. Gray, 210, A20
 Christiansen, J. L., Crossfield, I. J. M., Barentsen, G., et al. 2018, *AJ*, 155, 57
 Claret, A., & Bloemen, S. 2011, *A&A*, 529, A75
 Cutri, R. M. & et al. 2014, VizieR Online Data Catalog: II/328
 Díaz, R. F., Almenara, J. M., Santerne, A., et al. 2014, *MNRAS*, 441, 983
 Dorn, C., & Heng, K. 2018, *ApJ*, 853, 64
 Dorn, C., Khan, A., Heng, K., et al. 2015, *A&A*, 577, A83
 Dorn, C., Venturini, J., Khan, A., et al. 2017, *A&A*, 597, A37
 Doyle, A. P., Smalley, B., Faedi, F., Pollacco, D., & Gómez Maqueo Chew, Y. 2017, *MNRAS*, 469, 4850
 Feltzing, S., & Gonzalez, G. 2001, *A&A*, 367, 253
 Ginzburg, S., Schlichting, H. E., & Sari, R. 2016, *ApJ*, 825, 29
 Gupta, A., & Schlichting, H. E. 2021, *MNRAS*, 504, 4634
 Gustafsson, B., Edvardsson, B., Eriksson, K., et al. 2008, *A&A*, 486, 951
 Haywood, R. D., Collier Cameron, A., Unruh, Y. C., et al. 2016, *MNRAS*, 457, 3637
 Hedin, A. E. 1983, *J. Geophys. Res.*, 88, 10170
 Heiter, U., Jofré, P., Gustafsson, B., et al. 2015, *A&A*, 582, A49
 Henden, A. A., Levine, S., Terrell, D., & Welch, D. L. 2015, in *Am. Astron. Soc. Meeting Abstracts*, 225, 336.16
 Ivans, I. I., Simmerer, J., Sneden, C., et al. 2006, *ApJ*, 645, 613
 Izidoro, A., Ogihara, M., Raymond, S. N., et al. 2017, *MNRAS*, 470, 1750
 Jeans, J. 1925, Cambridge at the University Press
 Jofré, P., Heiter, U., Soubiran, C., et al. 2014, *A&A*, 564, A133
 Kervella, P., Bigot, L., Gallenne, A., & Thévenin, F. 2017, *A&A*, 597, A137
 Kipping, D. M. 2010, *MNRAS*, 408, 1758
 Kipping, D. M. 2014, *MNRAS*, 440, 2164
 Kjeldsen, H., Bedding, T. R., Arentoft, T., et al. 2008, *ApJ*, 682, 1370
 Leleu, A., Alibert, Y., Hara, N. C., et al. 2021, *A&A*, 649, A26
 Lissauer, J. J., Fabrycky, D. C., Ford, E. B., et al. 2011, *Nature*, 470, 53
 Lopez, E. D., & Fortney, J. J. 2014, *ApJ*, 792, 1
 Lopez, T. A., Barros, S. C. C., Santerne, A., et al. 2019, *A&A*, 631, A90
 MacDonald, M. G., Ragozzine, D., Fabrycky, D. C., et al. 2016, *AJ*, 152, 105
 MacDonald, M. G., Shakespeare, C. J., & Ragozzine, D. 2021, *AJ*, 162, 114
 Madhusudhan, N., Nixon, M. C., Welbanks, L., Piette, A. A. A., & Booth, R. A. 2020, *ApJ*, 891, L7
 Marcq, E., Salvador, A., Massol, H., & Davaille, A. 2017, *J. Geophys. Res. (Planets)*, 122, 1539
 Marcy, G. W., Isaacson, H., Howard, A. W., et al. 2014, *ApJS*, 210, 20
 McKay, A. J., DiSanti, M. A., Kelley, M. S. P., et al. 2019, *AJ*, 158, 128
 Meléndez, J., Bergemann, M., Cohen, J. G., et al. 2012, *A&A*, 543, A29
 Meléndez, J., Ramírez, I., Karakas, A. I., et al. 2014, *ApJ*, 791, 14
 Melwani Daswani, M., Vance, S. D., Mayne, M. J., & Glein, C. R. 2021, *Geophys. Res. Lett.*, 48, e94143
 Morel, T. 2018, *A&A*, 615, A172
 Morel, T., Miglio, A., Lagarde, N., et al. 2014, *A&A*, 564, A119
 Mousis, O., Deleuil, M., Aguichine, A., et al. 2020, *ApJ*, 896, L22
 Mousis, O., Aguichine, A., Bouquet, A., et al. 2021, *Planet. Sci. J.*, 2, 72
 Munari, U., Henden, A., Frigo, A., et al. 2014, *AJ*, 148, 81

- Ochsenbein, F., Bauer, P., & Marcout, J. 2000, *A&AS*, 143, 23
- Otegi, J. F., Dorn, C., Helled, R., et al. 2020, *A&A*, 640, A135
- Owen, J. E., & Jackson, A. P. 2012, *MNRAS*, 425, 2931
- Plotnykov, M., & Valencia, D. 2020, *MNRAS*, 499, 932
- Pluriel, W., Marcq, E., & Turbet, M. 2019, *Icarus*, 317, 583
- Ramos, X. S., Charalambous, C., Benítez-Llambay, P., & Beaugé, C. 2017, *A&A*, 602, A101
- Reddy, B. E., Tomkin, J., Lambert, D. L., & Allende Prieto, C. 2003, *MNRAS*, 340, 304
- Rivkin, A. S., Howell, E. S., Vilas, F., & Lebofsky, L. A. 2002, *Hydrated Minerals on Asteroids: The Astronomical Record*, 235
- Ruden, S. P. 1999, in *NATO Advanced Study Institute (ASI) Series C*, 540, *The Origin of Stars and Planetary Systems*, eds. C. J. Lada & N. D. Kylafis, 643
- Sanz-Forcada, J., Micela, G., Ribas, I., et al. 2011, *A&A*, 532, A6
- Scora, J., Valencia, D., Morbidelli, A., & Jacobson, S. 2020, *MNRAS*, 493, 4910
- Snedden, C. A. 1973, PhD thesis, The University of Texas at Austin, USA
- Sotin, C., Grasset, O., & Mocquet, A. 2007, *Icarus*, 191, 337
- Southworth, J. 2008, *MNRAS*, 386, 1644
- Terquem, C., & Papaloizou, J. C. B. 2007, *ApJ*, 654, 1110
- Thompson, M. A., Telus, M., Schaefer, L., et al. 2021, *Nat. Astron.*, 5, 575
- Tsantaki, M., Santos, N. C., Sousa, S. G., et al. 2019, *MNRAS*, 485, 2772
- Tsiaras, A., Waldmann, I. P., Tinetti, G., Tennyson, J., & Yurchenko, S. N. 2019, *Nat. Astron.*, 3, 1086
- Turbet, M., Bolmont, E., Ehrenreich, D., et al. 2020, *A&A*, 638, A41
- Vernazza, P., Marsset, M., Beck, P., et al. 2015, *ApJ*, 806, 204
- Wang, H. S., Morel, T., Quanz, S. P., & Mojzsis, S. J. 2020, *A&A*, 644, A19
- Yu, X., Moses, J. I., Fortney, J. J., & Zhang, X. 2021, *ApJ*, 914, 38
- Zeng, L., Jacobsen, S. B., Sasselov, D. D., et al. 2019, *Proc. Natl. Acad. Sci. U.S.A.*, 116, 9723

Appendix A: System parameters**Table A.1.** List of parameters used in the analysis. The priors are provided together with the posteriors. The posterior values represent the median and 68.3% credible interval. Derived values that might be useful for follow-up work are also reported.

Parameter	Prior	Posterior
<i>Stellar Parameters</i>		
Effective temperature T_{eff} [K]	$\mathcal{N}(5275.0, 50.0)$	$5354.7^{+27.9}_{-21.2}$
Surface gravity $\log g$ [cgs]	$\mathcal{N}(4.5, 0.11)$	$4.55^{+0.02}_{-0.02}$
Iron abundance [Fe/H] [dex]	$\mathcal{N}(0.08, 0.05)$	0.07 ± 0.05
Distance to Earth D [pc]	$\mathcal{N}(201.54, 1.97)$	201.5 ± 1.9
Interstellar extinction $E(B - V)$ [mag]	$\mathcal{U}(0.0, 1.0)$	$0.006^{+0.009}_{-0.005}$
Systemic radial velocity γ [km s ⁻¹]	$\mathcal{U}(-10.0, 10.0)$	$0.6392^{+0.0012}_{-0.0013}$
Linear limb-darkening coefficient u_a	(derived)	$0.4906^{+0.0075}_{-0.0071}$
Quadratic limb-darkening coefficient u_b	(derived)	$0.2084^{+0.0045}_{-0.0047}$
Stellar density $\rho_{\star}/\rho_{\odot}$	(derived)	$1.534^{+0.081}_{-0.090}$
Stellar mass M_{\star} [M_{\odot}]	(derived)	$0.891^{+0.017}_{-0.027}$
Stellar radius R_{\star} [R_{\odot}]	(derived)	$0.834^{+0.011}_{-0.01}$
Stellar age τ [Gyr]	(derived)	$3.3^{+2.4}_{-3.2}$
<i>Planet b Parameters</i>		
Orbital Period P_b [d]	$\mathcal{N}(2.35322, 0.01)$	$2.35308^{+0.00022}_{-0.00023}$
Transit epoch $T_{0,b}$ [BJD - 2450000]	$\mathcal{N}(7773.317, 0.001)$	$7773.31682^{+0.00092}_{-0.00090}$
Radial velocity semi-amplitude K_b [km s ⁻¹]	$\mathcal{U}(0.0, 0.1)$	$0.00146^{+0.00049}_{-0.00050}$
Orbital inclination i_b [°]	$\mathcal{S}(70.0, 90.0)$	$87.9^{+1.3}_{-1.1}$
Planet-to-star radius ratio k_b	$\mathcal{U}(0.0, 1.0)$	$0.01586^{+0.00072}_{-0.00066}$
Orbital eccentricity e_b	$\mathcal{T}(0.0, 0.083, 0.0, 1.0)$	$0.047^{+0.050}_{-0.033}$
Argument of periastron ω_b [°]	$\mathcal{U}(0.0, 360.0)$	169^{+93}_{-109}
System scale a_b/R_{\star}	(derived)	$8.6^{+0.1}_{-0.2}$
Impact parameter b_b	(derived)	$0.305^{+0.175}_{-0.191}$
Transit duration $T_{14,b}$ [h]	(derived)	$2.00^{+0.09}_{-0.11}$
Semi-major axis a_b [AU]	(derived)	$0.03332^{+0.00021}_{-0.00034}$
Planet mass M_b [M_{\oplus}]	(derived)	$2.80^{+0.94}_{-0.96}$
Planet radius R_b [R_{\oplus}]	(derived)	$1.442^{+0.071}_{-0.063}$
Planet bulk density ρ_b [g cm ⁻³]	(derived)	$5.1^{+2.0}_{-1.8}$
<i>Planet c Parameters</i>		
Orbital Period P_c [d]	$\mathcal{N}(3.55987, 0.01)$	$3.56004^{+0.00012}_{-0.00011}$
Transit epoch $T_{0,c}$ [BJD - 2450000]	$\mathcal{N}(7740.3223, 0.001)$	$7740.32185^{+0.00087}_{-0.00090}$
Radial velocity semi-amplitude K_c [km s ⁻¹]	$\mathcal{U}(0.0, 0.1)$	$0.00270^{+0.00052}_{-0.00051}$
Orbital inclination i_c [°]	$\mathcal{S}(70.0, 90.0)$	$88.7^{+0.8}_{-0.7}$
Planet-to-star radius ratio k_c	$\mathcal{U}(0.0, 1.0)$	$0.02418^{+0.00056}_{-0.00051}$
Orbital eccentricity e_c	$\mathcal{T}(0.0, 0.083, 0.0, 1.0)$	$0.037^{+0.041}_{-0.025}$
Argument of periastron ω_c [°]	$\mathcal{U}(0.0, 360.0)$	171^{+129}_{-78}
System scale a_c/R_{\star}	(derived)	11.3 ± 0.2
Impact parameter b_c	(derived)	$0.254^{+0.148}_{-0.160}$
Transit duration $T_{14,c}$ [h]	(derived)	$2.37^{+0.05}_{-0.06}$
Semi-major axis a_c [AU]	(derived)	$0.04391^{+0.00028}_{-0.00045}$
Planet mass M_c [M_{\oplus}]	(derived)	$5.95^{+1.17}_{-1.12}$
Planet radius R_c [R_{\oplus}]	(derived)	$2.198^{+0.066}_{-0.054}$
Planet bulk density ρ_c [g cm ⁻³]	(derived)	$3.1^{+0.7}_{-0.6}$
<i>Planet d Parameters</i>		
Continued on next page		

Table A.1. – continued from previous page

Parameter	Prior	Posterior
Orbital Period P_d [d]	$\mathcal{N}(5.40478, 0.01)$	5.40479 ± 0.00021
Transit epoch $T_{0,d}$ [BJD - 2450000]	$\mathcal{N}(7743.1607, 0.001)$	$7743.15984^{+0.00095}_{-0.00093}$
Radial velocity semi-amplitude K_d [km s $^{-1}$]	$\mathcal{U}(0.0, 0.1)$	0.00285 ± 0.00055
Orbital inclination i_d [°]	$\mathcal{S}(70.0, 90.0)$	$88.9^{+0.6}_{-0.5}$
Planet-to-star radius ratio k_d	$\mathcal{U}(0.0, 1.0)$	$0.02540^{+0.00069}_{-0.00065}$
Orbital eccentricity e_d	$\mathcal{T}(0.0, 0.083, 0.0, 1.0)$	$0.039^{+0.045}_{-0.027}$
Argument of periastron ω_d [°]	$\mathcal{U}(0.0, 360.0)$	207^{+69}_{-138}
System scale a_d/R_\star	(derived)	15.0 ± 0.3
Impact parameter b_d	(derived)	$0.297^{+0.145}_{-0.170}$
Transit duration $T_{14,d}$ [h]	(derived)	$2.71^{+0.07}_{-0.08}$
Semi-major axis a_d [AU]	(derived)	$0.05800^{+0.00037}_{-0.00059}$
Planet mass M_d [M_\oplus]	(derived)	$7.20^{+1.39}_{-1.40}$
Planet radius R_d [R_\oplus]	(derived)	$2.310^{+0.077}_{-0.068}$
Planet bulk density ρ_d [g cm $^{-3}$]	(derived)	3.2 ± 0.7
<i>Planet e Parameters</i>		
Orbital Period P_e [d]	$\mathcal{N}(8.26144, 0.01)$	$8.26146^{+0.00022}_{-0.00021}$
Transit epoch $T_{0,e}$ [BJD - 2450000]	$\mathcal{N}(7740.6451, 0.001)$	$7740.64563^{+0.00085}_{-0.00087}$
Radial velocity semi-amplitude K_e [km s $^{-1}$]	$\mathcal{U}(0.0, 0.1)$	$0.00387^{+0.00094}_{-0.00093}$
Orbital inclination i_e [°]	$\mathcal{S}(70.0, 90.0)$	$88.7^{+0.3}_{-0.2}$
Planet-to-star radius ratio k_e	$\mathcal{U}(0.0, 1.0)$	$0.03604^{+0.00074}_{-0.00072}$
Orbital eccentricity e_e	$\mathcal{T}(0.0, 0.083, 0.0, 1.0)$	$0.049^{+0.048}_{-0.034}$
Argument of periastron ω_e [°]	$\mathcal{U}(0.0, 360.0)$	223^{+67}_{-123}
System scale a_e/R_\star	(derived)	$19.8^{+0.3}_{-0.4}$
Impact parameter b_e	(derived)	$0.474^{+0.081}_{-0.115}$
Transit duration $T_{14,e}$ [h]	(derived)	2.97 ± 0.05
Semi-major axis a_e [AU]	(derived)	$0.07697^{+0.00050}_{-0.00079}$
Planet mass M_e [M_\oplus]	(derived)	$11.28^{+2.78}_{-2.72}$
Planet radius R_e [R_\oplus]	(derived)	$3.276^{+0.095}_{-0.082}$
Planet bulk density ρ_e [g cm $^{-3}$]	(derived)	$1.8^{+0.5}_{-0.4}$
<i>Planet f Parameters</i>		
Orbital Period P_f [d]	$\mathcal{N}(12.75759, 0.01)$	$12.75760^{+0.00051}_{-0.00048}$
Transit epoch $T_{0,f}$ [BJD - 2450000]	$\mathcal{N}(7738.7019, 0.001)$	$7738.70226^{+0.00093}_{-0.00092}$
Radial velocity semi-amplitude K_f [km s $^{-1}$]	$\mathcal{U}(0.0, 0.1)$	$0.00072^{+0.00091}_{-0.00052}$
Orbital inclination i_f [°]	$\mathcal{S}(70.0, 90.0)$	$88.8^{+0.2}_{-0.1}$
Planet-to-star radius ratio k_f	$\mathcal{U}(0.0, 1.0)$	$0.03065^{+0.00085}_{-0.00083}$
Orbital eccentricity e_f	$\mathcal{T}(0.0, 0.083, 0.0, 1.0)$	$0.057^{+0.059}_{-0.040}$
Argument of periastron ω_f [°]	$\mathcal{U}(0.0, 360.0)$	172^{+117}_{-112}
System scale a_f/R_\star	(derived)	26.5 ± 0.5
Impact parameter b_f	(derived)	$0.541^{+0.073}_{-0.109}$
Transit duration $T_{14,f}$ [h]	(derived)	3.20 ± 0.08
Semi-major axis a_f [AU]	(derived)	$0.10283^{+0.00066}_{-0.00105}$
Planet mass M_f [M_\oplus]	(derived)	$2.43^{+3.05}_{-1.75}$
Planet radius R_f [R_\oplus]	(derived)	$2.787^{+0.093}_{-0.085}$
Planet bulk density ρ_f [g cm $^{-3}$]	(derived)	$0.6^{+0.8}_{-0.4}$
<i>Planet g Parameters</i>		
Orbital Period P_g [d]	$\mathcal{N}(41.97, 0.1)$	$41.96822^{+0.00817}_{-0.00774}$
Transit epoch $T_{0,g}$ [BJD - 2450000]	$\mathcal{N}(7773.76, 2457773.93)$	$7773.86006^{+0.01931}_{-0.03522}$
Radial velocity semi-amplitude K_g [km s $^{-1}$]	$\mathcal{U}(0.0, 1.0)$	$0.00049^{+0.00058}_{-0.00035}$

Continued on next page

Table A.1. – continued from previous page

Parameter	Prior	Posterior
Orbital inclination i_g [°]	$\mathcal{S}(70.0, 90.0)$	$89.5^{+0.4}_{-0.3}$
Planet-to-star radius ratio k_g	$\mathcal{U}(0.0, 1.0)$	$0.03199^{+0.00327}_{-0.00248}$
Orbital eccentricity e_g	$\mathcal{T}(0.0, 0.083, 0.0, 1.0)$	$0.054^{+0.060}_{-0.038}$
Argument of periastron ω_g [°]	$\mathcal{U}(0.0, 360.0)$	164^{+148}_{-104}
System scale a_g/R_\star	(derived)	$58.6^{+1.0}_{-1.2}$
Impact parameter b_g	(derived)	$0.550^{+0.319}_{-0.365}$
Transit duration $T_{14,g}$ [h]	(derived)	$4.71^{+0.79}_{-1.63}$
Semi-major axis a_g [AU]	(derived)	$0.22745^{+0.00146}_{-0.00233}$
Planet mass M_g [M_\oplus]	(derived)	$2.45^{+2.92}_{-1.74}$
Planet radius R_g [R_\oplus]	(derived)	$2.911^{+0.305}_{-0.230}$
Planet bulk density ρ_g [g cm^{-3}]	(derived)	$0.5^{+0.7}_{-0.4}$
<i>Instrument-related Parameters</i>		
HARPS jitter $\sigma_{j, \text{RV}}$ [km s^{-1}]	$\mathcal{U}(0.0, 0.1)$	$0.00146^{+0.00068}_{-0.00077}$
K2 contamination [%]	$\mathcal{T}(0.0, 0.005, 0.0, 1.0)$	$0.003^{+0.004}_{-0.002}$
K2 jitter $\sigma_{j, \text{K2}}$ [ppm]	$\mathcal{U}(0.0, 10^5)$	185.9 ± 2.7
K2 out-of-transit flux	$\mathcal{U}(0.99, 1.01)$	$1.0000058^{+0.0000037}_{-0.0000038}$
SED jitter [mag]	$\mathcal{U}(0.0, 0.1)$	$0.02^{+0.017}_{-0.013}$

Notes:

- $\mathcal{N}(\mu, \sigma^2)$: Normal distribution with mean μ and width σ^2
- $\mathcal{U}(a, b)$: Uniform distribution between a and b
- $\mathcal{S}(a, b)$: Sine distribution between a and b
- $\mathcal{T}(\mu, \sigma^2, a, b)$: Truncated normal distribution with mean μ and width σ^2 , between a and b

A.3. Interior-atmosphere modelling to assess the observability of rocky planets with JWST (Acuña et al. Submitted)

L. Acuña, M. Deleuil, O. Mousis

This corresponds to the third first-author publication where I present the results of this thesis. The version of the model I use in this work involves our own atmospheric model (see Chapter 4), and the adaptive MCMC (see Sect. 5.2), in addition to the supercritical water layer in the interior model, and the interior-atmosphere coupling employed in the previous two first-author publications (see Sections A.1 and A.2).

The work corresponding to this publication has been presented in the following conferences:

- 'Celebrating JWST's first six months of exoplanet data', as a contributed talk, in November 2022, in Ringberg Castle, Kreuth (Germany)
- Europlanet Science Conference (EPSC) 2022, as a contributed talk, in September 2022, in Granada (Spain)
- Semaine de la Astrophysique Française (SF2A) 2022, as a contributed talk, in June 2022, in Besançon (France)

Interior-atmosphere modelling to assess the observability of rocky planets with JWST

L. Acuña¹, M. Deleuil¹, and O. Mousis^{1,2}

¹ Aix-Marseille Université, CNRS, CNES, Institut Origines, LAM, Marseille, France
e-mail: acuna@mpia.de

² Institut universitaire de France (IUF), France

Received XX; accepted XX

ABSTRACT

Context. Super-Earths present compositions dominated by refractory materials. However, there is a degeneracy in their interior structure between a planet with no atmosphere and a small Fe content, and a planet with a thin atmosphere and a higher core mass fraction. To break this degeneracy, atmospheric characterization observations are required.

Aims. We present a self-consistent interior-atmosphere model to constrain the volatile mass fraction, surface pressure, and temperature of rocky planets with water and CO₂ atmospheres. These parameters obtained in our analysis can then be used to predict observations in emission spectroscopy and photometry with JWST, which can determine the presence of an atmosphere or its composition.

Methods. We couple a 1D interior model with a supercritical water layer with a non-grey atmospheric model. To obtain the bolometric emission and Bond albedo for an atmosphere in radiative-convective equilibrium, we present the k-uncorrelated approximation for fast computations within our retrieval on planetary mass, radius and host stellar abundances. For the generation of emission spectra, we use our k-correlated atmospheric model. An adaptive MCMC is used for an efficient sampling of the parameter space at low volatile mass fractions.

Results. We show how to use our modelling approach to predict observations with JWST for TRAPPIST-1 c and 55 Cancri e, which have been proposed in Cycle 1. TRAPPIST-1 c's most likely scenario is a bare surface, although the presence of an atmosphere cannot be ruled out. If the emission in the MIRI F1500 filter is 731 ppm or higher, there would be a water-rich atmosphere. For fluxes between 730 and 400 ppm, no atmosphere is present, while low emission fluxes (300 ppm) indicate a CO₂-dominated atmosphere. In the case of 55 Cancri e, a combined spectrum with NiRCam and MIRI LRS may present high uncertainties at wavelengths between 3 and 3.7 μ m. There are no water or CO₂ features in this range, so their identification in the emission spectrum is not affected by high uncertainties.

Key words. Planets and satellites: atmospheres – Planets and satellites: interiors – Planets and satellites: composition – Planets and satellites: individual: TRAPPIST-1 d – Planets and satellites: individual: 55 Cnc e – Methods: statistical – Methods: numerical

1. Introduction

Low-mass exoplanets ($M < 20 M_{\oplus}$) have two different sub-populations based on their radius and density: super-Earths and sub-Neptunes. Super-Earths have radii of $R = 1.3 R_{\oplus}$, while the radius of sub-Neptunes corresponds to $R = 2.4 R_{\oplus}$ (Fulton et al. 2017; Fulton & Petigura 2018). If we compare these radii with planet interior and evolution models, super-Earths are mostly composed of Fe and Si-bearing rocks, whereas sub-Neptunes have a significant volatile (H/He, water) content. Despite having an idea of the main component for these planets from their mass and radius data, and interior structure models, we do not know their exact interior composition due to degeneracies.

In the case of super-Earths, we still have the question of whether the planet could have a thin atmosphere, or if it presents a bare rock surface. Atmospheres containing H/He are discarded, since a very small fraction of H/He entails a minimum radius of $\approx 1.6 R_{\oplus}$ (Lopez & Fortney 2014). Therefore, an atmosphere composed of water formed from ice pebbles accreted beyond or in the vicinity of the water ice line (Mousis et al. 2019; Krissansen-Totton et al. 2021; Kimura & Ikoma 2022), a secondary atmosphere built-up by outgassing (Ortenzi et al. 2020; Baumeister et al. 2021; Liggins et al. 2022), or a silicate at-

mosphere (Zahnle et al. 2009) are the most likely scenarios for super-Earths. This produces a degeneracy in the internal structure of super-Earths and Earth-sized planets: the same planetary mass and radius can be explained with a planet with no atmosphere and a low-Fe content rocky bulk (Madhusudhan 2012; Dorn et al. 2017), or a planet with a thin atmosphere and a core mass fraction similar to that of Earth (32% of core mass fraction).

This degeneracy in interior structure can only be broken with the support of atmospheric characterization data. The presence of an atmosphere has been confirmed in the hot super-Earth π Mensae c, whose detection of C II ions indicates atmospheric escape of a high-molecular atmosphere (García Muñoz et al. 2021). Phase curves have also been used to determine the existence of a silicate atmosphere in K2-141 b (Zieba et al. 2022), as well as transmission spectroscopy for the terrestrial planet LHS 3844 b (Diamond-Lowe et al. 2020). Moreover, Kreidberg et al. (2019) use the phase curves to confirm the absence of an atmosphere, as well as to constrain which material constitutes the planetary surface.

JWST (Gardner et al. 2006) will observe several super-Earths to confirm the presence of an atmosphere, or even narrow their possible atmospheric compositions. In this study, we

present a self-consistent interior-atmosphere model, Marseille’s Super-Earth Interior model (MSEI), to perform retrievals from estimated mass, radius, and stellar host abundances. As a result, we determine the posterior distribution functions (PDFs) of the atmospheric mass, and surface pressure and temperature of water and CO₂-dominated atmospheres. These atmospheric parameters obtained from our retrieval analysis can be used as input for an atmospheric model to produce spectra. We set an example of this with our k-correlated atmospheric model, MSEIRADTRAN, to generate emission spectra to predict observations with JWST with MIRI photometric filters, and NirCam and MIRI Low-Resolution Spectrometer (MIRI LRS). Combined interior and atmospheric models have been used to constrain the water mass fractions (WMF) of rocky planets as WMF < 10⁻³ (Agol et al. 2021). However, the consistent exploration of the parameter space in the region close to WMF = 0 is necessary to obtain the PDFs of the water mass fraction and the surface pressure accurately. For this reason, we employ an adaptive Markov chain Monte Carlo (MCMC) (Director et al. 2017) to explore the low surface pressures for possible water and CO₂ atmospheres in rocky planets.

We recall the basics of our interior model, MSEI, in Sect. 2. We explain the updates implemented in our atmosphere model, MSEIRADTRAN, with respect to similar previous k-correlated models (Marcq et al. 2017; Pluriel et al. 2019) in Sect. 3, as well as introduce the non-grey, k-uncorrelated approximation for fast computations of the bolometric flux and Bond albedo within our MCMC framework. In Sect. 4, we detail the implementation of the adaptive MCMC, and show an example of the retrieval with it and a non-adaptive one. We assess the observability of two planets that have been proposed for observations in Cycle 1 of JWST with our model: TRAPPIST-1 c (Gillon et al. 2016; Grimm et al. 2018) and 55 Cancri e (Ehrenreich et al. 2012; Bourrier et al. 2018). The planet and instrument parameters that we use as input for our interior-atmospheric analyses and Pandexo (Batalha et al. 2020) to predict uncertainties in JWST observations are summarised in Sect. 5. We present our results and conclusions in Sect. 6 and Sect. 7, respectively.

2. Interior structure model

We remind the reader the fundamental principles on which our interior structure model is based. The input of the interior structure model are the total mass and two compositional parameters, which are the core mass fraction (CMF), and the water mass fraction (WMF). The CMF and the WMF are defined as the mass of the Fe-rich core and the hydrosphere divided by the total planetary mass, respectively. In the 1D interior model, the planetary radius, r , is represented by a one-dimensional grid. Along this grid, the pressure, $P(r)$, the temperature, $T(r)$, the gravity acceleration, $g(r)$, and the density, $\rho(r)$, are calculated at each point. These four variables are obtained by solving the equation that defines it. The pressure is computed by integrating the equation of hydrostatic equilibrium (see Eq. 1), while the temperature requires integrating the adiabatic gradient profile (Eq. 2). In Eq. 2, γ and ϕ correspond to the Grüneisen and seismic parameters, respectively. The former describes the behaviour of the temperature in a crystal to its density. The latter parameter provides the speed at which seismic waves propagate in the same crystalline structure. Their formal definitions are shown in Eq. 3, where we can see that the seismic parameter is related to the slope of the density at constant pressure, while the Grüneisen parameter depends on the derivative of the pressure with respect to the internal energy, E . The acceleration of gravity is obtained by solving

the integral that results from Gauss’s theorem (Eq. 4), where G is the gravitational constant, and m corresponds to the mass at a given radius, r .

$$\frac{dP}{dr} = -\rho g \quad (1)$$

$$\frac{dT}{dr} = -g \frac{\gamma T}{\phi} \quad (2)$$

$$\begin{cases} \phi = \frac{dP}{d\rho} \\ \gamma = V \left(\frac{dP}{dE} \right)_V \end{cases} \quad (3)$$

$$\frac{dg}{dr} = 4\pi G\rho - \frac{2Gm}{r^3} \quad (4)$$

The density, $\rho(r)$, is computed with the Equation of State (EOS), which provides the density as a function of temperature and pressure. The interior structure model is divided three separate layers: a Fe-rich core, a mantle rich in silicates, and a water layer. We use a different EOS to calculate the density for each of these layers. We adopt the Vinet EOS (Vinet et al. 1989), with a thermal correction, for the core and the mantle. More details about this EOS and its reference parameter values for the core and mantle can be found in Brugger et al. (2016, 2017). For the hydrosphere, we use the EOS and specific internal energy of Mazevet et al. (2019) for supercritical and plasma phases of water, which is valid within the pressure and temperature regime ($P > 300$ bar, $T > 700$ K) covered by our interior structure model. We discuss the validity ranges of different water EOS for this regime in Acuña et al. (2021), while a detailed comparison of different EOS for high-pressure and high-temperature water and its effects on the total radius of the planet can be found in Aguichine et al. (2021).

The final input for our interior structure model are the surface temperature and pressure. These are the boundary conditions, together with the gravitational acceleration at the centre of the planet, whose value is zero, $g(r = 0) = 0$. Finally, the mass of each planetary layer is obtained by integrating the equation of conservation of mass (Eq. 5). The total planetary mass is the sum of the individual mass of the layers. When the total input mass and the initial boundary conditions are met, the model reaches convergence.

$$\frac{dm}{dr} = 4\pi r^2 \rho \quad (5)$$

2.1. Interior-atmosphere coupling

The surface pressure for the interior model depends on the mass of the atmospheric mass on top of the outermost interface of the interior model. For envelopes whose bottom pressure is greater or equal to $P = 300$ bar, the interior model’s surface pressure is set constant to 300 bar, which is the interface at which the interior and the atmosphere are coupled. Then the supercritical water layer extends from this interface to the boundary between the hydrosphere and the silicate mantle at higher pressures. For atmospheres whose surface pressure is less than 300 bar, the interior and atmosphere are coupled at the atmosphere-mantle interface,

having the water envelope in vapour phase only. The water mass fraction takes into account the mass of the atmosphere, M_{atm} . The atmospheric mass is calculated as in Eq. 6, where P_{base} is the pressure at the base of the atmosphere (atmosphere-interior interface), R_{bulk} is the radius comprised between the center of the planet and the base of the atmosphere, and g_{surf} is the acceleration of gravity at this interface. The coupling between the interior and the atmosphere models at a maximum pressure of 300 bar is sufficiently close to the critical point ($P = 220$ bar) of water to prevent the atmospheric model from taking over pressures at which convection dominates over radiation.

$$M_{atm} = \frac{P_{base} 4\pi R_{bulk}^2}{g_{surf}} \quad (6)$$

The atmospheric model calculates the outgoing longwave radiation (OLR) and the Bond albedo, A_B , given as a function of bulk mass, radius and temperature at the bottom of the atmosphere. If an atmosphere is in radiative equilibrium, its absorbed flux, F_{abs} must be equal to its emitted radiation, which is the OLR. The absorbed flux depends on the Bond albedo via Eqs. 7 and 8, where σ is the Stefan-Boltzmann constant, and T_{eq} is the planetary equilibrium temperature. This requires to know the semi-major axis of the planet, a_d , as well as the stellar radius and effective temperature, R_\star and T_\star , respectively.

$$F_{abs} = \sigma T_{eq}^4 \quad (7)$$

$$T_{eq} = (1 - A_B)^{0.25} \left(0.5 \frac{R_\star}{a_d} \right)^{0.5} T_\star \quad (8)$$

For a constant planetary mass and radius, we find the temperature at the base of the atmosphere by solving $OLR(T_{base}) - F_{abs}(T_{base}) = 0$ with a root-finding method, such as the bisection method. Then this is the input boundary condition for the interior structure model. The radius calculated by the interior structure model (from the centre of the planet up to the base of the atmosphere) is an input for the atmospheric model, while the temperature at the bottom of the atmosphere is an input for both the interior and the atmospheric model. Therefore, the self-consistent coupling of both models is not straightforward, and requires an iterative algorithm that checks that convergence is reached for the total radius and surface temperature. The total radius are computed as the sum of the bulk radius calculated by the interior model, and the atmospheric thickness obtained by the atmospheric model. We refer the reader to Acuña et al. (2021) for a detailed description of this algorithm.

3. Atmospheric model

The interior-atmosphere coupling presented in our previous work (Mousis et al. 2020; Acuña et al. 2021; Acuña et al. 2022) was done by using grids of data generated by the atmospheric model of Pluriel et al. (2019). These grids provide the OLR, Bond albedo and atmospheric thickness for a given set of mass, radius and surface temperature assuming a constant surface pressure. However, the use of these grids presents the following disadvantage: the grids do not enable us to generate emission spectra that could be used to simulate observations. Therefore, we develop our own atmospheric model, MSEIRADTRAN. We started the development of MSEIRADTRAN by modifying the

atmosphere model presented in Marcq et al. (2017)¹ to include up-to-date opacity and EOS data. In the following, we recall the basic structure and principles of MSEIRADTRAN and the atmospheric models presented in Marcq et al. (2017); Pluriel et al. (2019).

The 1D atmospheric model first proposes a pressure-temperature (PT) profile. This profile consists of a near-surface, dry convective layer, followed by a wet convective region where condensation takes place, and an isothermal mesosphere on top. If the surface temperature is cold enough to allow for condensation of water, the dry troposphere will not exist. For the isothermal mesosphere, we assume a constant temperature of 200 K (Marcq 2012; Marcq et al. 2017). The OLR is not very dependent on the temperature of an upper mesosphere (Kasting 1988). In addition, we do not take into account mesospheric stellar heating, which could increase significantly the temperature of the mesosphere. Therefore, adopting a low mesospheric temperature yields similar thermal profiles to self-consistent atmospheric calculations (Lupu et al. 2014). The one-dimensional grid that represents the pressure contains 512 computational layers. The adiabatic gradient used to calculate the temperature in each of these points in the convective regions depends on whether it is located in the dry or wet convective layer. The details of the computation of the wet and dry adiabatic gradients are recalled in Sect. 3.1.

The calculations of the emission spectrum and the Bond albedo are performed by bands. We divide the spectrum from 0 to 10100 cm^{-1} (equivalent to $\geq 1 \mu\text{m}$ in wavelength) in 36 bands to obtain the OLR, similar to Pluriel et al. (2019). For each band, we calculate the total optical depth in each computational layer, which has four different contributions. These contributions are the optical depth due to collision-induced absorption (CIA), and line opacity (see Sect. 3.4), Rayleigh scattering, and clouds. We treat Rayleigh scattering as in Pluriel et al. (2019), where the Rayleigh scattering opacity is related to wavelength, λ , following Eq. 9. The parameters κ_0 and λ_0 are adopted from Kopparapu et al. (2013) and Sneep & Ubachs (2005) for H_2O , and CO_2 , respectively. The opacity of clouds is considered for the atmospheric layers where condensation takes place. Similar to Marcq et al. (2017) and Pluriel et al. (2019), the cloud opacity is parameterized after Kasting (1988), who assumes a cloud opacity proportional to the extinction coefficient, Q_{ext} (see Eq. 10). The dependence of the extinction coefficient on wavelength (Eq. 11) is similar to that of water clouds on Earth (Kasting 1988; Marcq et al. 2017).

$$\kappa_{\text{Rayleigh}}(\lambda) = \kappa_0 \left(\frac{\lambda_0}{\lambda} \right)^4 \quad (9)$$

$$\kappa_{\text{clouds}}(\lambda) = 130 Q_{ext}(\lambda) \quad (10)$$

$$Q_{ext} = \begin{cases} 1 & \lambda \leq 20 \mu\text{m} \\ 3.26 \cdot \lambda^{-0.4} & \lambda > 20 \mu\text{m} \end{cases} \quad (11)$$

The total optical depth (Eq. 12), together with the PT profile, are the input for the radiative transfer equation solver, DISORT (Stamnes et al. 2017). DISORT obtains the emitted upward flux at the top of the atmosphere (TOA). The TOA flux is calculated for all 36 bands, which are then summed to obtain the bolometric, wavelength-integrated TOA flux, or OLR. Immediately after

¹ <http://marcq.page.laetmos.ipsl.fr/radconv1d.html>

the OLR is computed, we start the calculation of reflection spectra and the Bond albedo, described in Sect. 3.5.

$$\tau_{total} = \tau_{line} + \tau_{Rayleigh} + \tau_{CIA} + \tau_{clouds} \quad (12)$$

3.1. Atmospheric EOS

The thermal structure of the atmosphere is divided in two main layers. The near-surface layer is adiabatic, which means convection takes place, while the layer on top is a mesosphere with constant temperature. This is set to 200 K, which is representative of the cool temperatures that hot low-mass planets present in their mesospheres (Lupu et al. 2014; Leconte et al. 2013). The adiabatic layer is divided in two sub-layers: a dry convective layer and a wet convective one. Condensation may take place depending on the value of the pressure compared to the water saturation pressure, because we assume an atmosphere with one condensable species, H₂O, and one non-condensable gas, CO₂. To obtain the temperature in a computational layer, i , we consider two approximations. The first one is that the change in temperature and pressure within an atmospheric layer is small enough to approximate $\left(\frac{\partial T}{\partial P}\right)_S \simeq \frac{T_{i-1} - T_i}{P_{i-1} - P_i}$. The second approximation is $\Delta P \sim P_i \Delta \ln(P)$, because $\frac{d \ln(P)}{dP} \sim \frac{\Delta \ln(P)}{\Delta P} = \frac{1}{P_i}$. Then the temperature of a given computational layer, T_i can be calculated as a function of the pressure in that layer, P_i , and the derivative of the temperature with respect to pressure at constant entropy, $(\partial T / \partial P)_S$, as seen in Eq. 13.

$$T_i = T_{i-1} - \left(\frac{\partial T}{\partial P}\right)_S P_i \Delta \ln(P_i) \quad (13)$$

If the pressure of water vapour is below the gas saturation pressure, $P_v < P_{sat}(T)$, or if its temperature is above the temperature of the critical point of water, $T > T_{crit}$, we are under dry convection (Marcq 2012; Marcq et al. 2017). We calculate the derivative $(dT/dP)_S$ in the dry case following Marcq et al. (2017) (Eq. 14). ρ_v and ρ_c are the densities of water vapour and CO₂, respectively; and $C_{p,v}$ and $C_{p,c}$, their heat capacities. $V_v = 1/\rho_v$ is the specific volume of water vapour.

$$\left(\frac{\partial T}{\partial P}\right)_{S, dry} = \frac{\rho_v T (\partial V_v / \partial T)_P}{\rho_v C_{p,v} + \rho_c C_{p,c}} \quad (14)$$

When the atmospheric pressure reaches the water saturation curve, $P = P_{sat}(T)$, water vapour starts to condense out and clouds form. Since the phase change requires energy in the form of latent heat, the wet adiabatic coefficient is different from the dry adiabatic one (Eq. 14). The expression for the derivative $(dT/dP)_S$ in the wet case with CO₂ as the only non-condensable gas is provided in Eq. 15 (Marcq et al. 2017). M_c is the molecular weight of carbon dioxide, and $C_{v,c}$ is the specific heat capacity at constant volume of CO₂. R is the ideal gas constant, while α_v is the mixing ratio of water vapour density relative to CO₂, $\alpha_v = \rho_v / \rho_c$. Its derivative, $\partial \ln(\alpha_v) / \partial \ln(T)$ (Eq. 16), needs to be calculated as well (Kasting 1988; Marcq 2012; Marcq et al. 2017).

$$\left(\frac{\partial T}{\partial P}\right)_{S, wet} = \frac{1}{(dP_{sat}/dT) + \rho_c R / M_c (1 + \partial \ln(\rho_v) / \partial \ln(T) - \partial \ln(\alpha_v) / \partial \ln(T))}$$

Article number, page 4 of 13

$$\frac{\partial \ln(\alpha_v)}{\partial \ln(T)} = \frac{R/M_c (\partial \ln(\rho_v) / \partial \ln(T)) - C_{v,c} - \alpha_v (\partial s_v / \partial \ln(T))}{\alpha_v (s_v - s_c) + R/M_c} \quad (16)$$

The density and heat capacity of water were previously obtained by using the steam tables of Haar et al. (1984). These tables treat water as a non-ideal gas, although they are not valid for $T > 2500$ K. Therefore, for temperatures higher than 2500 K, we use the EOS tables from Haldemann et al. (2020) to calculate the thermodynamic properties of water. These tables are a compilation of different EOSs, where each EOS is applied in its validity region of the water phase diagram. There are two EOSs that are used in the region relevant for the atmospheres of low-mass, highly-irradiated planets. The first EOS is the IAPWS95 (Wagner & Pruß 2002), whose validity range for the high-pressure supercritical regime corresponds to 251 to 1273 K in temperature, and up to 1 GPa in pressure. Haldemann et al. (2020) transition to an EOS valid at low pressures and high temperatures at 1200 K. This second EOS is the CEA (Chemical Equilibrium with Applications) package (Gordon 1994; McBride 1996). This package incorporates the effects of single ionisation and thermal dissociation, which are processes that occur only at high temperatures in gas phase. Fig. 1 shows the dry adiabatic coefficient as a function of pressure and temperature in the region of the water phase diagram relevant for hot planetary atmospheres. The temperature derivative $(dT/dP)_S$ is closely related to the adiabatic coefficient, κ_{ad} (Eq. 17). The reduction of the dry adiabatic coefficient at $T = 1000$ to 2500 K is due to thermal dissociation, whereas the decrease at higher temperatures ($T \geq 6000$ K) is caused by thermal ionisation (Haldemann et al. 2020).

$$\left(\frac{\partial T}{\partial P}\right)_S = \frac{T}{P} \kappa_{ad}(P, T) \quad (17)$$

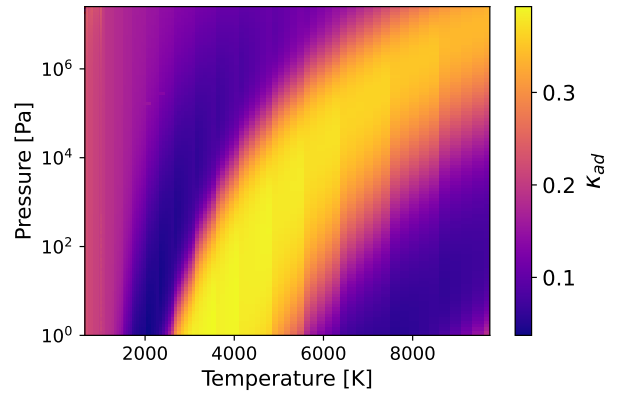


Fig. 1: Dry adiabatic coefficient for water, κ_{ad} , as a function of pressure and temperature. The displayed region covers the cold and hot gas phase of water up to pressures close to the critical point.

We assume that CO₂ is an ideal gas to calculate its density, ρ_c . As we treat CO₂ as an ideal gas, we can calculate its heat capacity $C_{v,c} = C_{p,c} - R/M_c$. We calculate the specific heat of CO₂

by using Shomate’s equation (Eq. 18), whose tabulated coefficients A to E are provided by Chase (1998). This is part of the NIST (National Institute of Standards and Technology) chemistry webbook database², and it is valid for temperatures up to 6000 K.

$$C_{p,c}(T) = A + B T + C T^2 + D T^3 + E/T^2 \quad (18)$$

Finally, we determine the atmospheric thickness, z_{atm} under the assumption of hydrostatic equilibrium (Eq. 1). In the case of the atmospheric model, the pressure is integrated over altitude, z , instead of internal radius, r . The altitude of the computational atmospheric layer i is shown in Eq. 19, where $\rho_{total,i}$ is the total mass density at altitude z_i , $\rho_{total,i} = \rho_{v,i} + \rho_{c,i}$. This expression is derived by approximating $(dP/dz) \sim \Delta P/z_{i-1} - z_i$. The final atmospheric thickness is obtained by evaluating z at the transit pressure, 20 mbar.

$$z_i = z_{i-1} + \frac{P_i \Delta \ln(P)}{g_i \rho_{total,i}} \quad (19)$$

3.2. K-correlated method implementation

We employ the k-correlated method (Liou 1980) to compute the emission spectrum. We discretize the dependence of the opacity on wavenumber, ν , by dividing the spectrum in bands. The spectral transmittance of a spectral band b , \mathcal{T}_b , is defined as the exponential sum of the opacity of the band, κ^b , times the column density m , which is only dependent on pressure and temperature (Sanchez-Lavega 2011). The exponential sum is performed over G -points, which are the abscissa values chosen to discretize the cumulative probability function of the opacity, $G(\kappa)$. Each G -point, G_i , has an associated weight in the exponential sum in Eq. 20, w_i . The discrete opacity value is not only dependent on the spectral band, but also on the G -point, which is noted explicitly in Eq. 20 as κ_i^b . N_G corresponds to the total number of G -points.

$$\mathcal{T}^b(m) = \int_0^1 e^{-\kappa^b(G) m} dG \simeq \sum_{i=1}^{N_G} e^{-\kappa_i^b m(P,T)} w_i \quad (20)$$

$$F_{TOA,b}^\uparrow = \int_0^1 F^b(G) dG = \sum_{i=1}^{N_G} F_{G_i}^b w_i \quad (21)$$

In the case of an arbitrary set of G -points, the weights are equal to $w_i = \Delta G_i$, which are the widths of the bins in G -space. In each atmospheric layer, the pressure and temperature are considered constant. Under this condition and within each spectral band, we can exchange wavenumber with G (Mollière 2017). Consequently, we integrate the upward flux over G to obtain the emitted flux within each band (Eq. 21). The upward top flux per bin and per G -point, $F_{G_i}^b$, is obtained by invoking the radiative solver for a total optical depth whose line optical depth is $\tau_{line} = \tau_i^b$ (Eq. 22).

$$\tau_i^b = \kappa_i^b \rho \Delta z \quad (22)$$

Following Malik et al. (2017, 2019), we define our discrete G -points as the roots of the Legendre polynomial, $G_{LG,i}$ (Eq. 23).

² <https://webbook.nist.gov/cgi/cbook.cgi?ID=C124389&Mask=1>

The corresponding weights, w_i , are the Legendre-Gaussian (LG) weights associated to the Legendre polynomial of N_G th order, P_{N_G} (Eq. 24). The LG weights are calculated from the i th root of the N_G th order Legendre polynomial, y_i , as well as from the polynomial’s derivative, P'_{N_G} (Abramowitz & Stegun 1972; Malik et al. 2017). We choose to have $N_G = 16$ G -points.

$$G_{LG,i} = \frac{(1 + y_i)}{2} \quad (23)$$

$$w_{LG,i} = \frac{2}{(1 - y_i^2) P'_{N_G}(y_i)^2} \quad (24)$$

To generate emission spectra to assess the observability with JWST, we use the original spectral resolution of the opacity k-table data (see Sect. 3.4) in our k-correlated model, which corresponds to $R = 200$ to 300 in the spectral range $\lambda = 1$ to 20 μm .

3.3. K-uncorrelated approximation

The k-correlated method requires solving the radiative transfer equation 16 times per band, with a different line optical depth (see Eq. 22). The radiative transfer solver is the most computationally expensive subroutine in our 1D atmospheric model, so we adopt a new approximation that we refer to as the k-uncorrelated method, which require less calls to DISORT per spectral bin. The difference between the uncorrelated-k and the correlated-k methods is where the iteration over G -points occurs. In the k-correlated method, the spectral lines are assumed to be correlated from an atmospheric layer to the next one, making it necessary to propagate the fluxes through the atmosphere for all G -points before performing the integral in Eq. 20. In the uncorrelated-k approach, the iteration over G -points ends before inputting the optical depth to the RT solver. The final line optical depth that is the input for the RT solver in the k-uncorrelated method assumes that all the contributions from the G -points are included before the sum in Eq. 20, being calculated as $\tau^b = -\ln(\mathcal{T}^b)$. Therefore, the radiative solver is invoked once per spectral bin. This reduces the computation time of the emission flux in one spectral band from 0.40 seconds to 0.22 seconds. In addition, we reduce the spectral resolution to $R = \lambda/\Delta\lambda = 10$ for the emission spectrum and Bond albedo, which is fast enough to perform a high number of interior-atmosphere models in a retrieval analysis (see Sect. 4) within a few days.

We use the version of our atmospheric model with the k-uncorrelated approximation to generate the OLR and Bond albedo necessary to calculate the boundary conditions for the interior model. To compare the effect of the k-uncorrelated approximation on the total radius of planets with water-rich envelopes, we compute two sets of mass-radius relationships (Fig. 2, upper panel). The first set is obtained by coupling the interior model with our k-uncorrelated model (dotted lines), and the second one with the k-correlated model of Pluriel et al. (2019). The k-uncorrelated approximation causes differences of less than 1% in radius compared to the k-correlated model of Pluriel et al. (2019). This difference stems from the difference in the temperature at the interior-atmosphere interface ($P = 300$ bar) between MSEIRADTRAN and Pluriel et al. (2019). The difference in temperature is between 200 to 230 K (Fig. 2, lower panel). Note that the constant surface temperature at masses below $5 M_\oplus$ in the Pluriel et al. (2019) model are due to the limit in the data

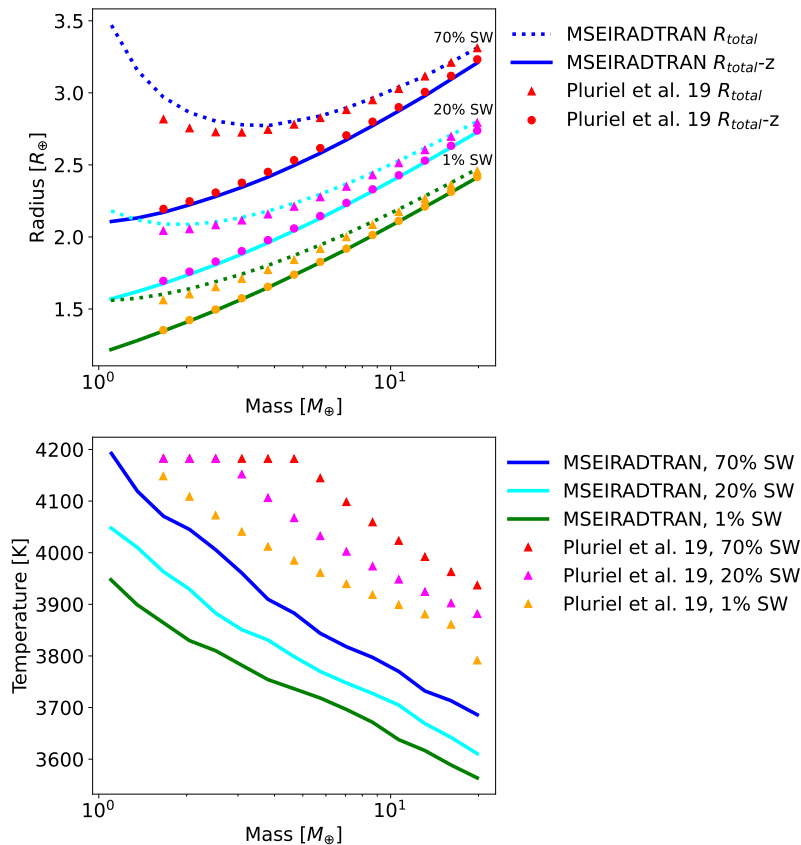


Fig. 2: Upper panel: Mass-radius relationships for a planet with a water-dominated atmosphere orbiting a Sun-like star at $a_d = 0.05$ AU. Dashed lines indicate the total radius calculated by the k-unrelated version of MSEIRADTRAN, while the solid line corresponds to the interior radius, which comprises the core, mantle, and supercritical water (SW). Triangles and circles indicate the total radius and the interior radius obtained when the interior model is coupled with the atmospheric model of Pluriel et al. (2019), respectively. Lower panel: Temperature at the 300 bar interface as a function of planetary mass.

grid we use, having no physical interpretation. In Fig. 3 we show the OLR and Bond albedo as a function of bottom temperature for our atmospheric model and that of Pluriel et al. (2019). The trends of the Bond albedo are very similar for both models, obtaining a similar absorbed flux (dashed lines). Consequently, the discrepancy in bottom temperature is caused by the difference in OLR between the two models, which is approximately a factor of 2.

We conclude that the k-unrelated MSEI may be used to calculate the radius of water-rich planets within our MCMC retrieval framework, since fast computations of the OLR and Bond albedo are necessary. The k-unrelated approach is an alternative to grey models, whose difference in OLR with k-correlated models is greater than that obtained with the k-unrelated model. However, the k-unrelated approximation should not be used to compute spectra at higher resolution since the planetary emission differs by a factor of 2 compared to a k-correlated method. This entails a significant difference in predictions of the emission spectrum for observations.

3.4. Opacity data

The total optical depth computed in the atmospheric model include contributions from collision-induced absorption (CIA)

and line absorption. CIA absorption is particularly important in dense gases, such as steam and CO_2 at high pressures, specially if the line opacity is weak (Pluriel et al. 2019). We adopt CIA absorption data for $\text{H}_2\text{O}-\text{CO}_2$ and $\text{H}_2\text{O}-\text{H}_2\text{O}$ collisions provided by Ma & Tipping (1992) and Tran et al. (2018)³, respectively. CO_2-CO_2 CIA opacities are read from a look-up table obtained by Bézard et al. (2011) and Marcq et al. (2008), which is also used in the atmospheric model by Marcq et al. (2017). Our $\text{H}_2\text{O}-\text{H}_2\text{O}$ CIA table covers the complete spectral range where we calculate both our emission and reflection spectra, while the $\text{H}_2\text{O}-\text{CO}_2$ CIA table covers the bands with wavelength $\geq 1 \mu\text{m}$, which corresponds to the bands necessary for the emission spectrum only. For the bands whose wavelength is outside the spectral range of the CIA table, we assume a constant CIA opacity value equal to the opacity at the limit band of the table.

Grimm & Heng (2015) provide a database⁴ of pre-calculated opacity k-tables for different species and line lists. For water and CO_2 -dominated atmospheres, we adopt the POKAZATEL (Polyansky et al. 2018) and HITEMP2010 (Rothman et al. 2010) opacity data, respectively. POKAZATEL presents the widest va-

³ https://www.lmd.jussieu.fr/~lmdz/planets/LMDZ.GENERIC/datagcm/continuum_data/

⁴ <https://chaldene.unibe.ch/data/Opacity3/>

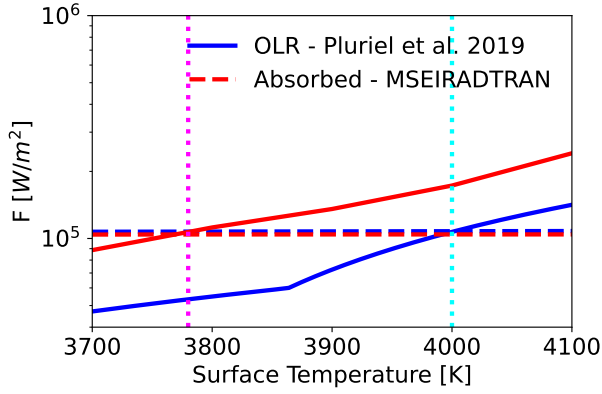


Fig. 3: Outgoing Longwave Radiation (OLR, solid lines) and absorbed flux (dashed lines) as a function of bottom atmospheric temperature, calculated with our atmospheric model, MSEIRADTRAN (red), and Pluriel et al. (2019) (blue). Dotted lines indicate the temperature at which the OLR and the absorbed flux are equal, meaning the atmosphere is in radiative-convective equilibrium.

lidity range in temperature for water in planetary atmospheres, with a maximum temperature of 5000 K, while HITEMP maximum temperature is 4000 K. Following the procedure described in Leconte (2021), we bin the k-correlated opacities to the same spectral bins of Marcq et al. (2017) and Pluriel et al. (2019). We calculate the k-coefficients for our water-CO₂ mixture by assuming that the spectral features of the individual gases are correlated (Malik et al. 2017). In the correlated approximation, the mixed opacity is estimated as indicated in Eq. 25, where χ_j is the mixing ratio by mass of the j th gas, $\kappa_{j,i}$ is the k-coefficient of the j th gas evaluated at the G_i point. The mixing ratio by mass is defined as $\chi_j = \frac{X_j MW_j}{\mu}$, where $X_j = \frac{P_j}{P}$ is the volume mixing ratio of the j th species, MW_j is its molecular weight, and μ is the mean molecular weight of the mixture.

$$\kappa_{mix,i} = \sum_{j=1}^{N_{gases}} \chi_j \kappa_{j,i} \quad (25)$$

3.5. Reflection spectra and Bond albedo

Once the bolometric OLR is obtained, we initiate the calculation of the reflectivity in 30 bands, from 5 to 0.29 μm , to obtain the Bond albedo (Pluriel et al. 2019). The bands for which we calculate both the emission flux and the reflectivity (from 1 to 5 μm) cannot have the two quantities calculated simultaneously since DISORT requires different input settings to calculate them. For the emission, we assume zero illumination from the top of the atmosphere, as well as an upward flux that forms 90 degrees with the surface of the planet, which corresponds to a polar angle equal to zero. To obtain the reflectivity, we assume an isotropically-incident source of radiation at the top of the atmosphere, while turning off all thermal emission sources (Pluriel et al. 2019). DISORT calculates the reflectivity of the atmosphere as a function of incident beam angle, which corresponds to the solar zenith angle (SZA). The SZA, θ , is the angle that the incident light forms with the normal of the incident surface. Once we obtain the dependence of the reflectivity

with SZA, we can average it as indicated in Eq. 26 (Simonelli & Veverka 1988).

$$A_B(\nu) = 2 \int_0^{\pi/2} A_B(\nu, \theta) \cos(\theta) \sin(\theta) d\theta \quad (26)$$

After averaging the reflectivity over SZA, we obtain the reflection spectrum, which is the dependence of the albedo as a function of wavenumber. To obtain the bolometric Bond albedo, we integrate Eq. 27 (Pluriel et al. 2019). $A_B(\nu)$ is the reflectivity as a function of wavenumber, $B_\nu(T_\star)$ is Planck's function for a temperature equal to the effective temperature of the host star, and σ is the Stefan-Boltzmann constant.

$$A_{Bond, bol} = \frac{\pi \int_0^\infty A_B(\nu) B_\nu(T_\star) d\nu}{\sigma T_\star^4} \quad (27)$$

The Bond albedo is a parameter particularly sensitive to the choice of phase function. For atmospheric layers that present clouds, the gas contributes to scattering with a Rayleigh phase function, while we assume the Henyey-Greenstein phase function for clouds. DISORT requires the calculation of the Legendre moments of the combined phase function, which we estimate as the weighted average of the moments of the two individual phase functions (Liou 1980; Boucher 1998). The weights are calculated as the ratio of the optical depth due to Rayleigh scattering or clouds divided by the total optical depth, $\tau_{Rayleigh} + \tau_{clouds}$, for Rayleigh and Henyey-Greenstein phase function moments, respectively. For clear atmospheric layers, the only contribution to scattering is Rayleigh scattering due to the gas, so the total phase function moment corresponds to that of Rayleigh scattering (Marcq et al. 2017).

DISORT also requires as input the single scattering albedo of each atmospheric layer. The single scattering albedo is defined as the ratio of scattering efficiency to total extinction efficiency. The total extinction is a sum of both extinction by scattering and extinction by absorption. Therefore, a single scattering albedo of 1 indicates that all extinction is due to scattering, whereas a value of zero means that absorption dominates. Similar to the moments of the phase function, we estimate the combined single scattering albedo from gas (Rayleigh) and clouds with their weighted average. The single scattering albedo due to Rayleigh scattering is calculated as the Rayleigh optical depth divided by the total optical depth, $\tau_{clear} + \tau_{Rayleigh}$. The clear optical depth is the sum of the line and CIA optical depths (Sect. 3.4). The single scattering albedo due to clouds is calculated as the ratio of the clouds optical depth divided by $\tau_{clear} + \tau_{cloud}$, times the cloud single scattering albedo defined in Kasting (1988):

$$\varpi_0 = \begin{cases} 1 & \lambda \leq 2 \mu\text{m} \\ 1.24 \cdot \lambda^{-0.32} & \lambda > 2 \mu\text{m} \end{cases} \quad (28)$$

4. Markov chain Monte Carlo (MCMC)

We use the MCMC Bayesian algorithm described in Dorn et al. (2015) and later adapted by Acuña et al. (2021) to our forward interior-atmosphere model. We recall that the model parameters are the planetary mass, M , the CMF, x_{core} , and the WMF, x_{H_2O} . Therefore, one single model is determined by these three parameters as $\mathbf{m} = \{M, x_{core}, x_{H_2O}\}$. The available data are the total mass M , the total radius R , and the Fe/Si abundance, $\mathbf{d} = \{M_{obs}, R_{obs}, Fe/Si_{obs}\}$, whose observational errors

are $\sigma(M_{obs})$, $\sigma(R_{obs})$, $\sigma(Fe/Si_{obs})$, respectively. When the Fe/Si mole ratio is not considered in the inverse problem, the data is reduced to only the total planetary mass and radius, $\mathbf{d} = \{M_{obs}, R_{obs}\}$. The prior information consists on a Gaussian distribution centered on the mean value of the observed mass, with a standard deviation equal to the observational uncertainty. For the CMF and WMF, we consider as priors uniform distributions ranging from 0 to 1. The MCMC scheme first starts by drawing a value for each of the model parameters from their prior distributions, which we denote as $\mathbf{m}_1 = \{M_1, x_{core,1}, x_{H2O,1}\}$. The interior model calculates the planetary radius and Fe/Si abundance that corresponds to these model parameters, which is $\mathbf{g}(\mathbf{m}_1) = \{R_1, M_1, Fe/Si_1\}$. The likelihood of this model is computed (Eqs. 29 and 30) and a new model is drawn from the prior distributions, \mathbf{m}_2 .

$$L(\mathbf{m}_i | \mathbf{d}) = C \exp \left(-\frac{1}{2} \left[\left(\frac{(R_i - R_{obs})}{\sigma(R_{obs})} \right)^2 + \left(\frac{(M_i - M_{obs})}{\sigma(M_{obs})} \right)^2 + \left(\frac{(Fe/Si_i - Fe/Si_{obs})}{\sigma(Fe/Si_{obs})} \right)^2 \right] \right), \quad (29)$$

$$C = \frac{1}{(2\pi)^{3/2} [\sigma^2(M_{obs}) \cdot \sigma^2(R_{obs}) \cdot \sigma^2(Fe/Si_{obs})]^{1/2}} \quad (30)$$

The log-likelihoods, $l(\mathbf{m}_i | \mathbf{d}) = \log(L(\mathbf{m}_i | \mathbf{d}))$, of both models are used to estimate the acceptance probability (Eq. 31). Consecutively, a random number is drawn from a uniform distribution between 0 and 1. If P_{accept} is greater than this random number, \mathbf{m}_2 is accepted and the chain moves to this set of model parameters, starting the following chain $n + 1$. Otherwise, the chain remains in \mathbf{m}_1 and a different set of model parameters is proposed, \mathbf{m}_3 . The accepted models are stored, and values of their parameters conform the PDF that will enable us to estimate their mean and uncertainties.

$$P_{accept} = \min \left\{ 1, e^{(l(\mathbf{m}_{new,i}(\mathbf{d})) - l(\mathbf{m}_{old}(\mathbf{d})))} \right\} \quad (31)$$

4.1. Adaptive MCMC

In our initial implementation of the MCMC (Acuña et al. 2021), the random walker uses a uniform distribution to choose the next state where it is going to move in the parameter space of the CMF and WMF. This is called a naive walk (Mosegaard & Tarantola 1995), in which all points in the parameter space have a probability of being chosen proportional to their number of neighbours. This poses the following problem: for the states whose CMF or WMF are close to 0 or 1, they are less likely to be sampled in the random walk, because they have less neighbours than the central values. A model with WMF = 1 is not physical, although many highly-irradiated rocky planets might present low-mass atmospheres, which correspond to a WMF close to 0. To compensate the lower probability of being chosen in the limiting states of the prior, we use an adaptive step size in the walker. This consists on adapting the maximum size of the perturbation used to generate a new model instead of using a fixed value everywhere in the parameter space. This adaptive step size will decrease in the limiting areas of the prior (i.e low WMF states) and have its greatest value in the centre of the prior (WMF = 0.5). The self-adjusting step size is carried by a transformation of the parameter space, which ranges from exponential to spherical transforms

(Chaudhry et al. 2021). In this work, we choose to implement the self-adjusting logit transform (SALT), proposed by Director et al. (2017). The SALT transform is publicly available in the *SALTSampler* R package⁵, which eases its implementation in Python for our own model.

We compare the non-adaptive and adaptive MCMC for one planet, TOI-220 b (Hoyer et al. 2021). We consider as input data the total mass and radius, as well as the Fe/Si mole ratio, which has been calculated with the stellar abundances of the host star. No maximum limits have been established for the CMF or the WMF. TOI-220 b has an equilibrium temperature of 806 K, which means that it is strongly irradiated and could present steam and supercritical phases. Table 1 presents the input data, and the retrieved parameters of the non-adaptive and adaptive MCMCs. All three agree within uncertainties for mass, radius and Fe/Si. The uncertainties of the mass and radius in the non-adaptive MCMC are smaller than the input data. This difference in uncertainties is significant in the case of the total mass. This indicates that the non-adaptive MCMC is not as effective as the adaptive MCMC at sampling all possible $\{x_{core}, x_{H2O}\}$ pairs that could reproduce the mass and radius data. As a consequence, the uncertainties of the WMF are being underestimated in the non-adaptive MCMC, while the adaptive MCMC produces a greater confidence interval for the WMF, and retrieves the exact uncertainties of the mass and radius.

	Data (Hoyer et al. 2021)	Non-adaptive	Adaptive
$M [M_{\oplus}]$	13.8±1.0	13.8±0.7	13.7±1.0
$R [R_{\oplus}]$	3.03±0.15	3.06±0.12	2.98±0.15
Fe/Si	0.65±0.09	0.64±0.11	0.64±0.10
x_{core}		0.08±0.03	0.09±0.03
x_{H2O}		0.62±0.10	0.58±0.14

Table 1: TOI-220 b MCMC input (Data), and output mean values and 1σ uncertainties for the non-adaptive and adaptive MCMCs.

In Fig. 4, we show the sampled 2D PDFs for the CMF and the WMF in the ternary diagram. In addition to the same area of the ternary diagram as the non-adaptive algorithm, the adaptive MCMC explores an area at lower WMF along the Fe/Si = 0.65 isoline, going down to WMF = 0.10 in the driest simulations. This is a consequence of the ability of the adaptive MCMC to sample better the extremes of the prior distribution of the WMF, in comparison to the non-adaptive MCMC. Furthermore, the acceptance rate is also improved in the adaptive case, having an acceptance rate of 53% in comparison to the original acceptance rate of 35% of the non-adaptive case within the same time.

5. Planetary and observation parameters

Our MCMC analysis requires as input the planetary mass, radius and Fe/Si mole ratio for the two planets we consider as test-case in this work, TRAPPIST-1 c and 55 Cancri e. Their values and references are shown in Table 2. To determine the surface temperature at which the atmosphere is in radiative-convective equilibrium, the stellar effective temperature, stellar radius, and semi-major axis are needed, as seen in Eq. 8. The values we adopt and their references are given in Table 2.

To simulate the observations with JWST in photometry, we assume the atmospheric parameters retrieved in our adaptive

⁵ <https://rdrr.io/cran/SALTSampler/man/SALTSampler-package.html>

Table 2: Planetary parameters for TRAPPIST-1 c and 55 Cancri e: masses, radii, Fe/Si mole ratios, semi-major axes, and host stellar effective temperatures and radii.

	$M [M_{\oplus}]$	$R [R_{\oplus}]$	Fe/Si	a_d [AU]	T_{\star} [K]	$R_{\star} [R_{\odot}]$	References
TRAPPIST-1 c	1.308 ± 0.056	$1.097^{+0.014}_{-0.012}$	0.76 ± 0.12	1.58×10^{-2}	2566	0.119	1, 2
55 Cancri e	$7.99^{+0.32}_{-0.33}$	1.875 ± 0.029	0.60 ± 0.14	1.54×10^{-2}	5172	0.943	3, 4

References. (1) Agol et al. (2021); (2) Unterborn et al. (2018); (3) Bourrier et al. (2018); (4) Luck (2016)

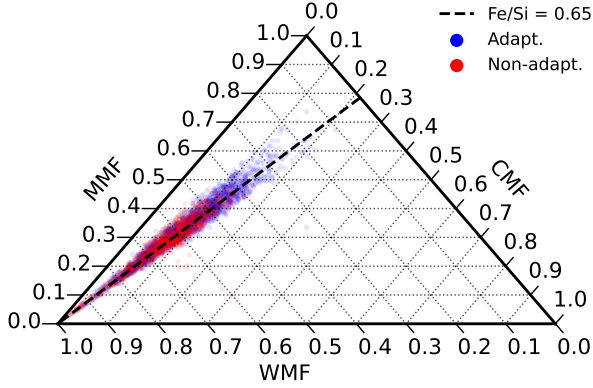


Fig. 4: Sampled 2D PDFs of the CMF and WMF in the ternary diagram for TOI-220 b for non-adaptive (red) and adaptive (blue) MCMCs. The mean value of the input Fe/Si mole ratio is indicated in a dashed black line.

MCMC analysis and generate emission spectra with their respective temperature-pressure profiles. Consecutively, we bin the emission spectrum using the response functions of each of the MIRI photometry filters⁶ (Glasse et al. 2015; Piette et al. 2022). The mean flux, $\langle f_{\lambda} \rangle$, of an emission spectrum, $f(\lambda)$, observed with a filter with transmission function $R(\lambda)$, is defined in Eq. 32 (Stolker et al. 2020). We consider random Gaussian noise of 100 ppm for each filter to derive the uncertainties of the mean flux (Lustig-Yaeger et al. 2019; Piette et al. 2022).

$$\langle f_{\lambda} \rangle = \frac{\int f(\lambda) R(\lambda) d\lambda}{\int R(\lambda) d\lambda} \quad (32)$$

For the observation of the emission spectrum of 55 Cancri e, we use Pandexo (Batalha et al. 2020) to simulate the expected noise. Our input includes the stellar effective temperature, and stellar and planet radius (see Table 2). Additional input parameters are found in the database accessible by Pandexo and ExoMast, which are shown in Table 3. We adopt observation and instrumentation variables from Hu et al. (2021). We adopt stellar spectrum templates generated by petitRADTRANS (Mollière et al. 2019) to calculate the planet-to-star flux ratio with our planetary emission spectra.

Parameter	Value
<i>Star</i>	
Metallicity, $\log[\text{Fe}/\text{H}]$	0.35
Gravity, $\log g$ [cgs]	4.43
J Magnitude	4.59
<i>Planet</i>	
Transit duration [d]	0.0647
<i>Observation</i>	
Baseline [h]	3.2
Number of eclipses	2
Instrument	MIRI LRS NIRCam
Mode	Slitless F444W, subgrism 64
Saturation limit [full well]	80%
Constant minimum noise	100 ppm

Table 3: Input parameters for Pandexo to simulate observations of the emission spectrum of 55 Cancri e with JWST’s MIRI Low Resolution Spectroscopy (LRS) and NIRCam instruments.

6. Interior composition and simulated spectra

6.1. TRAPPIST-1 c

TRAPPIST-1 c has been proposed to be observed in thermal emission with the Mid-infrared instrument (MIRI) in JWST Cycle 1 (Kreidberg et al. 2021). It will be observed with the F1500W filter during 4 eclipses, which is the filter centered at $\lambda = 15 \mu\text{m}$. We analyse TRAPPIST-1 c with our adaptive MCMC with a water and a CO_2 -dominated atmosphere. Fig. 5 shows the marginal posterior distributions (PDF) in 1D and 2D for this analysis. The CMF agrees well with the value obtained previously with our non-adaptive MCMC in Acuña et al. (2021). In contrast, the distribution of the WMF derived by the adaptive MCMC is wider than the non-adaptive one, with mean and standard deviation values of $\text{WMF}_{\text{adapt}} = (11.0 \pm 5.6) \times 10^{-6}$, and $\text{WMF}_{\text{non-adap}} = (0.0^{+2.7}_{-0.0}) \times 10^{-6}$, respectively. This is because the adaptive MCMC is more effective at exploring the corner region of the compositional parameter space, where the WMFs are close to zero. More models in this region are accepted in the posterior distribution, and therefore it becomes wider with larger standard deviation than the non-adaptive posterior distribution. The corresponding surface pressure derived by the adaptive MCMC is $P_{\text{surf}} = 15 \pm 7$ bar for a water-dominated envelope. This 1σ confidence interval must be taken carefully since the PDF of the surface pressure does not present a Gaussian distribution shape.

A WMF of zero is also compatible with the density of TRAPPIST-1 c. Consequently, we can conclude that TRAPPIST-1 c could have a H_2O atmosphere of up to ≈ 25 bar of surface pressure, or no atmosphere at all. The analyses performed

⁶ <http://svo2.cab.inta-csic.es/svo/theory/fps3/index.php?id=JWST>

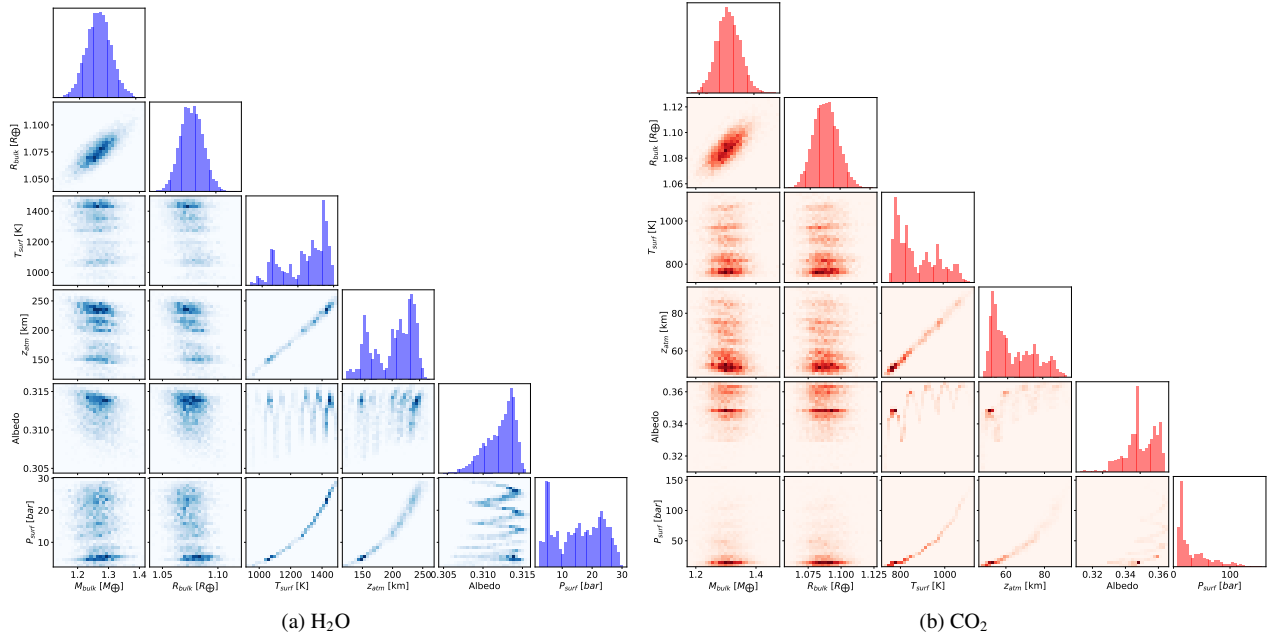


Fig. 5: PDF pyramid plot of the atmospheric parameters of TRAPPIST-1 c for H₂O (left panel) and CO₂ (right panel) atmospheres. The atmospheric parameters include the surface temperature (T_{surf}), atmospheric thickness (z_{atm}), Bond albedo, and the surface pressure (P_{surf}).

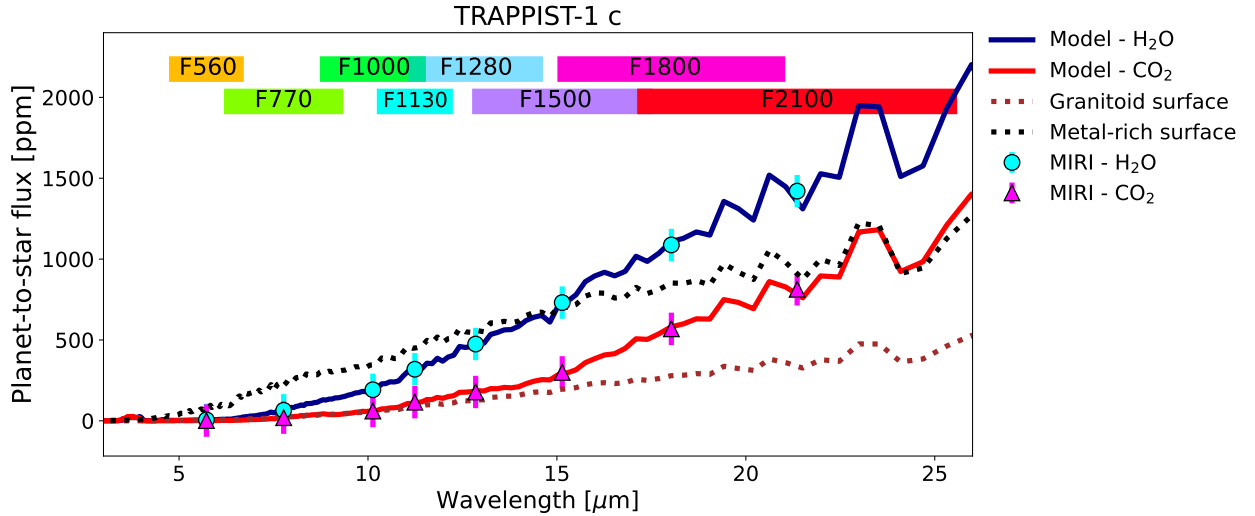


Fig. 6: Simulated emission spectra of TRAPPIST-1 c with MIRI photometric filter mean fluxes for water and CO₂ atmospheres, obtained with the high-resolution, k-correlated MSEIRADTRAN model. Color boxes indicate the wavelength ranges of the different filters.

with the k-correlated atmospheric model and the k-uncorrelated approximation agree that a H₂O atmosphere in TRAPPIST-1 c would have a surface temperature between 1000 and 1500 K, and an atmospheric thickness of 150 to 250 km. The retrieved Bond albedos differ by 0.05 due to the use of different atmospheric opacity data.

For a CO₂-dominated atmosphere, we retrieve a similar CMF to the water case, although the volatile mass fraction increases by

one order of magnitude, $VMF = (2.49 \pm 2.07) \times 10^{-5}$. The molecular weight of CO₂ is higher than that of water vapour, producing a more compressed atmosphere for a similar surface pressure and temperature. In addition, the radiative properties (i.e opacity) of CO₂ yields a lower surface temperature for the same irradiation conditions in comparison to a water-dominated envelope, which contributes to a lower atmospheric thickness. As a consequence, the models with a CO₂ envelope can accommodate a more mas-

sive atmosphere for TRAPPIST-1 c than the water models, retrieving a surface pressure of $P_{surf} = 35 \pm 29$ bar. The surface temperature and atmospheric thickness are $T_{surf} = 807 \pm 102$ K, and $z_{atm} = 63 \pm 12$ km, respectively.

We assume the atmospheric parameters retrieved in our adaptive MCMC analysis and generate emission spectra with their respective temperature-pressure profiles as explained in Sect. 5. Fig. 6 shows the complete emission spectra and mean filter fluxes for TRAPPIST-1 c. We observe that for the nIR filters ($\lambda = 5.60$ to $11.30 \mu\text{m}$), both atmosphere models have very similar fluxes that are compatible within uncertainties, which makes it not possible to distinguish between the two compositions in these wavelengths. For the filters F1500, F1800 and F2100, the mean flux uncertainties of the water and the CO_2 atmospheres do not overlap, allowing the different compositions to be distinguished. Therefore, observing TRAPPIST-1 c in emission with filter F1500, as proposed by Kreidberg et al. (2021), is well-adapted to differentiate between a water or a CO_2 -dominated atmosphere. We also consider the possibility of a bare surface in TRAPPIST-1 c, given the high probability obtained in our MCMC analysis for a volatile mass fraction equal to zero. Hu et al. (2012) obtain the emission spectra of bare terrestrial surfaces for different minerals. We estimate the brightness temperature for the irradiation conditions of TRAPPIST-1 c for two minerals, a metal-rich surface and a granitoid one, since these are the two surfaces with the highest and lowest emission for the same irradiation conditions, respectively. We approximate the emission spectrum of these surfaces to that of a black body with a temperature equal to the estimated brightness temperature. The emission flux in the F1500 filter for a water atmosphere is slightly higher than of the metal-rich surface. If TRAPPIST-1 c emission flux in this filter is 731 ppm or higher, it is indicative of a thin, water-rich atmosphere. For very low emission fluxes (≈ 300 ppm), TRAPPIST-1 c would present a CO_2 -dominated atmosphere. For fluxes between 730 to 400 ppm, TRAPPIST-1 c would have no atmosphere, with an emission that corresponds to a bare surface. The surfaces with the lowest emission are granitoid, feldspathic, or clay (Hu et al. 2012).

6.2. 55 Cancri e

55 Cancri e is a super-Earth in a close orbit ($P = 0.66$ days) to a bright star, for which several interior and atmospheric hypothesis have been proposed. Madhusudhan (2012) explored a carbon-rich interior given the high C/O ratio found for its host star, showing that in this case the planetary bulk density would be lower than that of a silicate-rich mantle planet, such as Earth. They concluded that a volatile layer would not be necessary to account for its density. On the other hand, a classical Fe-rich core and a silicate mantle are compatible with a volatile envelope rich in secondary atmosphere species. Furthermore, the absence of a H/He-dominated envelope seems likely due to the lack of hydrogen and helium emission and absorption lines in the spectrum (Ehrenreich et al. 2012; Zhang et al. 2021). The presence of a secondary atmosphere is supported by phase curve data from the *Spitzer* Space Telescope (Angelo & Hu 2017). A fully H_2O -dominated atmosphere has been discarded, since it will require the presence of water and hydrogen simultaneously in the atmosphere due to water dissociation. Therefore, the most likely composition for the atmosphere of 55 Cancri e is a mixture of silicate compounds (Keles et al. 2022), such as HCN, detected by Tsiaras et al. (2016), with traces of water (detected by Esteves et al. 2017), or CO_2 . Despite the water and CO_2 -dominated atmospheres being discarded by the current data, it is still inter-

esting to carry out our interior-atmosphere analysis for 55 Cancri e, since the scale height of a silicate high-molecular weight atmosphere might be similar to that of an envelope with the two compositions we consider. The scale height influences, together with the abundances, the intensity of a line in the spectrum of an exoplanet.

Hu et al. (2021) have proposed to observe 55 Cancri e in emission spectroscopy combining NIRCам F444W filter ($3\text{--}5 \mu\text{m}$), and MIRI's Low Resolution Spectrograph (MIRI LRS; $5\text{--}14 \mu\text{m}$). In Table 4, we observe that a water-dominated atmosphere reproduces well the observed data, with a surface pressure higher than 200 bar, whereas the CO_2 envelope is not extended enough to match the density of 55 Cancri e, yielding a more dense interior. At temperatures higher than 4000 K, CO_2 would not be the dominant species in a C-rich atmosphere, but CO. This changes the emission of the atmosphere as CO is a different absorber from CO_2 . A CO-rich atmosphere could also explain the low-density of 55 Cancri e in this scenario, since CO has a lower molecular weight than CO_2 , yielding a larger atmospheric scale height. When H/He is not included in the interior modelling, water as a trace species is necessary to explain the low density of 55 Cancri e, since a purely dry silicate atmosphere will have a smaller thickness than a CO_2 atmosphere due to their heavier molecular weights under similar atmospheric surface conditions. Adding silicate absorbers decreases the total planetary radius in H/He envelopes (Misener & Schlichting 2022). However, more modelling work is necessary to explore the effect of silicates in atmospheres that have lost their primordial H/He. A planet with no volatiles could match the low planetary density if the core and the mantle were less dense than that of an Earth-like interior, pointing to a carbon-rich mantle as suggested by Madhusudhan (2012). The emission spectrum in this scenario would be that of a bare surface, requiring terrestrial surface models such as the ones presented in Hu et al. (2012).

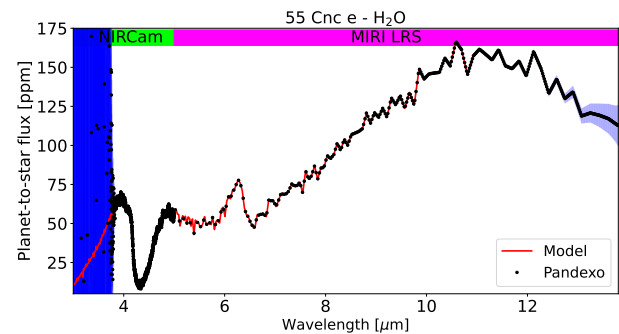


Fig. 7: Predicted emission spectrum with the k-correlated, high-resolution MSEIRADTRAN model for a water-rich atmosphere in 55 Cancri e with NIRCам and MIRI LRS.

Fig. 7 shows the complete predicted emission spectrum of 55 Cancri e from 3 to $14 \mu\text{m}$. We notice that for wavelengths below $\approx 3.7 \mu\text{m}$ the noise is quite high, although in the range $\lambda = 3$ to $3.7 \mu\text{m}$ there are no water or CO_2 spectral lines. Nonetheless, the rest of the spectral coverage of the proposed observations have low noise level, which makes the spectral features of water easy to identify with JWST in the high-molecular weight atmosphere scenario of 55 Cancri e.

	Data	H ₂ O	CO ₂
$M [M_{\oplus}]$	$7.99^{+0.32}_{-0.33}$	$7.99^{+0.29}_{-0.32}$	8.17 ± 0.29
$R [R_{\oplus}]$	1.875 ± 0.029	1.877 ± 0.020	1.850 ± 0.021
Fe/Si	0.60 ± 0.14	0.62 ± 0.14	0.48 ± 0.13
CMF		0.20 ± 0.05	0.15 ± 0.05
VMF		$(6.7^{+7.4}_{-5.9}) \times 10^{-5}$	$(5.1^{+5.9}_{-4.6}) \times 10^{-5}$
P_{surf} [bar]		209 ± 93	> 300
T_{surf} [K]		4161 ± 199	4035 ± 597 (at 300 bar)
z_{atm} [km]		522 ± 46	152 ± 33 (from 300 bar)
Bond albedo		0.191 ± 0.001	0.351 ± 0.004

Table 4: MCMC retrieved mean value and 1σ uncertainties for observable (mass, radius and Fe/Si mole ratio), compositional (core and volatile mass fractions) and atmospheric (surface pressure and temperature, atmospheric thickness and Bond albedo) parameters of 55 Cancri e.

7. Conclusions

In this work, we present a self-consistent model built to estimate the internal compositions and structures of low-mass planets with water and CO₂ atmospheres given their observed mass, radius, and their host stellar abundances. We couple self-consistently the interior and the atmosphere to obtain the boundary conditions at the top of a supercritical water layer or a silicate mantle given the irradiation conditions of the planet. This is done by calculating the bolometric emission flux and the Bond albedo, to compute the flux emitted and absorbed by an atmosphere in radiative-convective equilibrium. For fast computations of the radiative quantities within our MCMC retrieval approach, we present the k-uncorrelated approximation. This is as fast as grey models, but does not underestimate the outgoing longwave radiation as much as these compared to a k-correlated model. We show that the k-uncorrelated approximation presents sufficient accuracy to obtain the radius of an irradiated water planet within a 1% error with k-correlated models. This error increases with higher water mass fraction (WMF > 70%) and lower planetary mass. We also demonstrate that using a constant step size when sampling the prior distribution in a MCMC scheme is not efficient in exploring the parameter space in interior modelling. This causes an underestimation of the uncertainties of the compositional parameters. Therefore, it is necessary to use an adaptive MCMC when performing retrieval with interior models, especially for planets whose compositional parameters can reach the maximum or minimum possible values. This is the case of rocky Earth-sized planets and super-Earths, whose WMFs are close to zero, but are nonetheless important to determine their surface pressure.

Moreover, we use the surface pressure and temperature conditions retrieved with our interior-atmosphere model to generate emission spectra with our k-correlated atmospheric model, MSEIRADTRAN. We showcase how to use interior and atmospheric modelling simultaneously to predict observations for two rocky planets, TRAPPIST-1 c and 55 Cancri e, which have been proposed for observations in emission photometry and spectroscopy with JWST. We bin our emission spectra according to the response functions of the MIRI filters to predict emission fluxes for TRAPPIST-1 c for different scenarios, while for 55 Cancri e we input our emission spectra to Pandexo to predict observational uncertainties.

The most likely scenario for TRAPPIST-1 c is the lack of an atmosphere (WMF = 0). Nonetheless, the presence of a secondary atmosphere cannot be ruled out. In this scenario, TRAPPIST-1 c could have a H₂O-dominated atmosphere of up to 25 bar of surface pressure. A CO₂-rich envelope would have

higher surface pressures, and therefore it would be more massive than the water case, with a maximum surface pressure of 64 bar. We present emission flux estimates for the filter centered at 15 μ m, F1500, that can be compared with observations by Kreidberg et al. (2021) to determine if TRAPPIST-1 c has a bare surface. For 55 Cancri e, a massive envelope with more than 300 bar of surface pressure that contains water is necessary to fit its low density. We determine that a combined spectrum with NirCam and MIRI LRS, as proposed by Hu et al. (2021), may present a high noise level at wavelengths between 3 and 3.7 μ m. However, this part of the spectrum does not contain any spectral lines of water or CO₂, which are essential to determine the abundances in the envelope.

In our modelling approach, we have considered pure water and CO₂ envelopes. However, the atmospheres of low-mass planets are more diverse than these two compositional scenarios. The atmospheric compositions of sub-Neptunes are proving to be a mixture of H/He, water and other compounds via observations and models (Madhusudhan et al. 2020; Bézard et al. 2020; Guzmán-Mesa et al. 2022), while super-Earths can have more exotic atmospheric compositions, such as mineral atmospheres (Keles et al. 2022). Therefore, the scope of future work will be to include more gases in the atmospheric model, as well as the calculation of transmission spectra in addition to the existing implementation of emission and reflection spectra. Our interior-atmosphere model, MSEI, serves as a precedent to develop models with more diverse envelope compositions to prepare proposals for JWST and future atmospheric characterization facilities, such as Ariel (Tinetti et al. 2018). Our model can also be used within retrieval frameworks to interpret mass, radius and upcoming JWST emission spectral data simultaneously to break degeneracies in exoplanet composition.

Acknowledgements. M.D. and O.M. acknowledge support from CNES.

References

- Abramowitz, M. & Stegun, I. A. 1972, Handbook of Mathematical Functions
 Acuña, L., Deleuil, Magali, Mousis, Olivier, et al. 2021, A&A, 647, A53
 Acuña, L., Lopez, T. A., Morel, T., et al. 2022, A&A, 660, A102
 Agol, E., Dorn, C., Grimm, S. L., et al. 2021, The Planetary Science Journal, 2, 1
 Aguchine, A., Mousis, O., Deleuil, M., & Marcq, E. 2021, The Astrophysical Journal, 914, 84
 Angelo, I. & Hu, R. 2017, The Astronomical Journal, 154, 232
 Batalha, N., Fix, M., Fraine, J., et al. 2020, natashabatalha/PandExo: Release 1.5
 Baumeister, P., Tosi, N., Grenfell, J. L., & MacKenzie, J. 2021, in Plato Mission Conference 2021. Presentations and posters of the online PLATO Mission Conference 2021, 87
 Bézard, B., Charnay, B., & Blain, D. 2020, arXiv e-prints, arXiv:2011.10424

- Bézar, B., Fedorova, A., Bertaux, J.-L., Rodin, A., & Korabiev, O. 2011, *Icarus*, 216, 173
- Boucher, O. 1998, *Journal of the Atmospheric Sciences*, 55, 128
- Bourrier, Dumusque, X., Dorn, C., et al. 2018, *A&A*, 619, A1
- Brugger, B., Mousis, O., Deleuil, M., & Deschamps, F. 2017, *The Astrophysical Journal*, 850, 93
- Brugger, B., Mousis, O., Deleuil, M., & Lunine, J. I. 2016, *The Astrophysical Journal*, 831, L16
- Chase, M. W., J. 1998, *J. Phys. Chem. Ref. Data, Monograph 9*.
- Chaudhry, S., Lautzenheiser, D., & Ghosh, K. 2021, *arXiv e-prints*, arXiv:2110.10840
- Diamond-Lowe, H., Charbonneau, D., Malik, M., Kempton, E. M. R., & Beletsky, Y. 2020, *AJ*, 160, 188
- Director, H. M., Gattiker, J., Lawrence, E., & Wiel, S. V. 2017, *Journal of Statistical Computation and Simulation*, 87, 3521
- Dorn, C., Khan, A., Heng, K., et al. 2015, *A&A*, 577, A83
- Dorn, C., Venturini, J., Khan, A., et al. 2017, *A&A*, 597, A37
- Ehrenreich, Bourrier, V., Bonfils, X., et al. 2012, *A&A*, 547, A18
- Esteves, L. J., de Mooij, E. J. W., Jayawardhana, R., Watson, C., & de Kok, R. 2017, *The Astronomical Journal*, 153, 268
- Fulton, B. J. & Petigura, E. A. 2018, *AJ*, 156, 264
- Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, *The Astronomical Journal*, 154, 109
- García Muñoz, A., Fossati, L., Youngblood, A., et al. 2021, *ApJ*, 907, L36
- Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, *Space Sci. Rev.*, 123, 485
- Gillon, M., Jehin, E., Lederer, S. M., et al. 2016, *Nature*, 533, 221
- Glasse, A., Rieke, G. H., Bauwens, E., et al. 2015, *Publications of the Astronomical Society of the Pacific*, 127, 686
- Gordon, S. 1994, *Tech. rep. NASA Lewis Research Center*.
- Grimm, S. L., Demory, Brice-Olivier, Gillon, Michaël, et al. 2018, *A&A*, 613, A68
- Grimm, S. L. & Heng, K. 2015, *The Astrophysical Journal*, 808, 182
- Guzmán-Mesa, A., Kitzmann, D., Mordasini, C., & Heng, K. 2022, *MNRAS*, 513, 4015
- Haar, L., Gallagher, J. S., & Kell, G. S. 1984
- Haldemann, J., Alibert, Y., Mordasini, C., & Benz, W. 2020, *A&A*, 643, A105
- Hoyer, S., Gandolfi, D., Armstrong, D. J., et al. 2021, *Monthly Notices of the Royal Astronomical Society*, 505, 3361
- Hu, R., Brandeker, A., Damiano, M., et al. 2021, *Determining the Atmospheric Composition of the Super-Earth 55 Cancri e, JWST Proposal. Cycle 1, ID. #1952*
- Hu, R., Ehlmann, B. L., & Seager, S. 2012, *The Astrophysical Journal*, 752, 7
- Kasting, J. F. 1988, *Icarus*, 74, 472
- Keles, E., Mallonn, M., Kitzmann, D., et al. 2022, *MNRAS*, 513, 1544
- Kimura, T. & Ikoma, M. 2022, *Nature Astronomy*, 6, 1296
- Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, *The Astrophysical Journal*, 765, 131
- Kreidberg, L., Agol, E., Bolmont, E., et al. 2021, *Hot Take on a Cool World: Does Trappist-1c Have an Atmosphere?, JWST Proposal. Cycle 1, ID. #2304*
- Kreidberg, L., Koll, D. D. B., Morley, C., et al. 2019, *Nature*, 573, 87
- Krissansen-Totton, J., Galloway, M. L., Wogan, N., Dhaliwal, J. K., & Fortney, J. J. 2021, *ApJ*, 913, 107
- Leconte, J. 2021, *A&A*, 645, A20
- Leconte, J., Forget, F., Charnay, B., Wordsworth, R., & Pottier, A. 2013, *Nature*, 504, 268
- Liggins, P., Jordan, S., Rimmer, P. B., & Shorttle, O. 2022, *Journal of Geophysical Research (Planets)*, 127, e07123
- Liou, K. N. 1980, *An introduction to atmospheric radiation*.
- Lopez, E. D. & Fortney, J. J. 2014, *ApJ*, 792, 1
- Luck, R. E. 2016, *The Astronomical Journal*, 153, 21
- Lupu, R. E., Zahnle, K., Marley, M. S., et al. 2014, *The Astrophysical Journal*, 784, 27
- Lustig-Yaeger, J., Meadows, V. S., & Lincowski, A. P. 2019, *The Astronomical Journal*, 158, 27
- Ma, Q. & Tipping, R. H. 1992, *The Journal of Chemical Physics*, 96, 8655
- Madhusudhan, N. 2012, *The Astrophysical Journal*, 758, 36
- Madhusudhan, N., Nixon, M. C., Welbanks, L., Piette, A. A. A., & Booth, R. A. 2020, *ApJ*, 891, L7
- Malik, M., Grosheintz, L., Mendonça, J. M., et al. 2017, *The Astrophysical Journal*, 153, 56
- Malik, M., Kitzmann, D., Mendonça, J. M., et al. 2019, *The Astrophysical Journal*, 157, 170
- Marcq, E. 2012, *Journal of Geophysical Research: Planets*, 117
- Marcq, E., Bézar, B., Drossart, P., et al. 2008, *Journal of Geophysical Research (Planets)*, 113, E00B07
- Marcq, E., Salvador, A., Massol, H., & Davaille, A. 2017, *Journal of Geophysical Research (Planets)*, 122, 1539
- Mazevet, S., Licari, A., Chabrier, G., & Potekhin, A. Y. 2019, *A&A*, 621, A128
- McBride, B. J. G. 1996, *Tech. rep. NASA Lewis Research Center*.
- Misener, W. & Schlichting, H. E. 2022, *MNRAS*, 514, 6025
- Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, *A&A*, 627, A67
- Mollière, P. M. 2017, *PhD thesis, Heidelberg University*
- Mosegaard, K. & Tarantola, A. 1995, *Journal of Geophysical Research (Solid Earth)*, 100, 12,431
- Mousis, O., Deleuil, M., Aguichine, A., et al. 2020, *The Astrophysical Journal Letters*, 896, L22
- Mousis, O., Ronnet, T., & Lunine, J. I. 2019, *ApJ*, 875, 9
- Ortenzi, G., Noack, L., Sohl, F., et al. 2020, *Scientific Reports*, 10, 10907
- Piette, A. A. A., Madhusudhan, N., & Mandell, A. M. 2022, *MNRAS*, 511, 2565
- Pluriel, W., Marcq, E., & Turbet, M. 2019, *Icarus*, 317, 583
- Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 480, 2597
- Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111, 2139
- Sanchez-Lavega, A. 2011, *Contemporary Physics*, 52, 487
- Simonelli, D. P. & Veverka, J. 1988, *Icarus*, 74, 240
- Sneep, M. & Ubachs, W. 2005, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 92, 293
- Stamnes, K., Tsay, S. C., Jayaweera, K., et al. 2017, *DISORT: DIScrete Ordinate Radiative Transfer*
- Stolker, Quanz, S. P., Todorov, K. O., et al. 2020, *A&A*, 635, A182
- Tinetti, G., Drossart, P., Eccleston, P., et al. 2018, *Experimental Astronomy*, 46, 135
- Tran, H., Turbet, M., Chelin, P., & Landsheere, X. 2018, *Icarus*, 306, 116
- Tsiaras, A., Rocchetto, M., Waldmann, I. P., et al. 2016, *The Astrophysical Journal*, 820, 99
- Untertorn, C. T., Desch, S. J., Hinkel, N. R., & Lorenzo, A. 2018, *Nature Astronomy*, 2, 297
- Vinet, P., Rose, J. H., Ferrante, J., & Smith, J. R. 1989, *Journal of Physics: Condensed Matter*, 1, 1941
- Wagner, W. & Pruß, A. 2002, *Journal of Physical and Chemical Reference Data*, 31, 387
- Zahnle, K., Marley, M. S., Freedman, R. S., Lodders, K., & Fortney, J. J. 2009, *ApJ*, 701, L20
- Zhang, M., Knutson, H. A., Wang, L., et al. 2021, *The Astronomical Journal*, 161, 181
- Zieba, S., Zilinskas, M., Kreidberg, L., et al. 2022, *arXiv e-prints*, arXiv:2203.00370

B. Other publications and contributions

Planetary system LHS 1140 revisited with ESPRESSO and TESS (Lillo-Box et al. 2020)

J. Lillo-Box, P. Figueira, A. Leleu, L. Acuña, J. P. Faria, N. Hara, N. C. Santos, A. C. M. Correia, P. Robutel, M. Deleuil, D. Barrado, S. Sousa, X. Bonfils, O. Mousis, J. M. Almenara, N. Astudillo-Defru, E. Marcq, S. Udry, C. Lovis, and F. Pepe

The modelling of LHS 1140 b was done with the version of the model described in [Brugger et al. \(2017\)](#) and Sect. 2.1, since this planet is found in the habitable zone and it could sustain liquid water. To retrieve its CMF and WMF, I embedded this version of the model into the MCMC algorithm (see Sect 5.1). For the innermost planet, LHS 1140 c, the mass-radius relationships in Fig. 16 of this publication were derived with the coupled interior-atmosphere model. In this case, I used the grids with atmospheric data generated with the model from [Pluriel et al. \(2019\)](#), as explained in Sect. 3.1.

TOI-220 b: a warm sub-Neptune discovered by TESS (Hoyer et al. 2021)

S. Hoyer, D. Gandolfi, D. J. Armstrong, M. Deleuil, L. Acuña, J. R. de Medeiros, E. Goffo, J. Lillo-Box, E. Delgado Mena, T. A. Lopez, A. Santerne, S. Sousa, M. Fridlund, V. Adibekyan, K. A. Collins, L. M. Serrano, P. Cortés-Zuleta, S. B. Howell, H. Deeg, A. Aguichine, O. Barragán, E. M. Bryant, B. L. Canto Martins, K. I. Collins, B. F. Cooke, R. F. Díaz, M. Esposito, E. Furlan, S. Hojjatpanah, J. Jackman, J. M. Jenkins, E. L. N. Jensen, D. W. Latham, I. C. Leão, R. A. Matson, L. D. Nielsen, A. Osborn, J. F. Otegi, F. Rodler, S. Sabotta, N. J. Scott, S. Seager, C. Stockdale, P. A. Strøm, R. Vanderspek, V. Van Eylen, P. J. Wheatley, J. N. Winn, J. M. Almenara, D. Barrado, S. C. C. Barros, D. Bayliss, F. Bouchy, P. T. Boyd, J. Cabrera, W. D. Cochran, O. Demangeon, J. P. Doty, X. Dumusque, P. Figueira, W. Fong, S. Grziwa, A. P. Hatzes, P. Kabáth, E. Knudstrup, J. Korth, J. H. Livingston, R. Luque, O. Mousis, S. E. Mullally, H. P. Osborn, E. Pallé, C. M. Persson, S. Redfield, N. C. Santos, J. Smith, J. Šubjak, J. D. Twicken, S. Udry, and D. A. Yahalomi

TOI-220 b is highly irradiated, so in this publication, I used the coupled interior-atmosphere model within the non-adaptive MCMC method to retrieve the planetary composition. In the interior-atmosphere coupling, I used the grids of atmospheric data produced with the atmospheric model of [Pluriel et al. \(2019\)](#).

The results of this publication were also presented as a poster in the TESS Science conference II, in August 2021 (virtual).

TOI-2196 b: Rare planet in the hot Neptune desert transiting a G-type star (Persson et al. 2022)

C. M. Persson, I. Y. Georgieva, D. Gandolfi, L. Acuña, A. Aguichine, A. Muresan, E. Guenther, J. Livingston, K. A. Collins, M. Fridlund, E. Goffo, J. S. Jenkins, P. Kabáth, J. Korth, A. M. Levine, L. M. Serrano, J. Vines, O. Barragán, I. Carleo, K. D. Colon, W. D. Cochran, J. L. Christiansen, H. J. Deeg, M. Deleuil, D. Dragomir, M. Esposito, T. Gan, S. Grziwa, A. P. Hatzes, K. Hesse, K. Horne, J. M. Jenkins, J. F. Kielkopf, P. Klagyivik, K. W. F. Lam, D. W. Latham, R. Luque, J. Orell-Miquel, A. Mortier, O. Mousis, N. Narita, H. L. M. Osborne, E. Pallé, R. Papini, G. R. Ricker, H. Schmerling, S. Seager, K. G. Stassun, V. Van Eylen, R. Vanderspek, G. Wang, J. N. Winn, B. Wohler, R. Zambelli, and C. Ziegler

TOI-2196 b is too inflated to have a water-rich atmosphere, so I modelled this particular case with our preliminary H/He model. This model was built by combining our interior's core and mantle with the mass-radius relationships for H/He envelopes from Zeng et al. (2019), as described in Sect. 6.3. The version employed in this publication of the MCMC method is the adaptative one (see Sect. 5.2).

Characterization of the HD 108236 system with CHEOPS and TESS. Confirmation of a fifth transiting planet. (Hoyer et al. 2022)

S. Hoyer, A. Bonfanti, A. Leleu, L. Acuña, L. M. Serrano, M. Deleuil, A. Bekkelien, C. Broeg, H. G. Florén, D. Quelo, T. G. Wilson, S. G. Sousa, M. J. Hooton, V. Adibekyan, Y. Alibert, R. Alonso, G. Anglada, J. Asquier, T. Bárczy, D. Barrado, S. C. C. Barros, W. Baumjohann, M. Beck, T. Beck, W. Benz, N. Billot, F. Biondi, X. Bonfils, A. Brandeker, J. Cabrera, S. Charnoz, A. Collier Cameron, Sz. Csizmadia, M. B. Davies, L. Delrez, O. D. S. Demangeon, B.-O. Demory, D. Ehrenreich, A. Erikson, A. Fortier, L. Fossati, M. Fridlund, D. Gandolfi, M. Gillon, M. Güdel, N. Hara, K. Heng, K. G. Isaak, J. M. Jenkins, L. L. Kiss, J. Laskar, D. W. Latham, A. Lecavelier des Etangs, M. Lendl, C. Lovis, A. Luntzer, D. Magrin, P. F. L. Maxted, V. Nascimbeni, G. Olofsson, R. Ottensamer, I. Pagano, E. Pallé, C. M. Persson, G. Peter, D. Piazza, G. Piotto, D. Pollacco, R. Ragazzoni, N. Rando, H. Rauer, I. Ribas, G. R. Ricker, S. Salmon, N. C. Santos, G. Scandariato, S. Seager, D. Ségransan, A. E. Simon, A. M. S. Smith, M. Steller, Gy. M. Szabó, N. Thomas, J. D. Twicken, S. Udry, V. Van Grootel, R. K. Vanderspek, N. A. Walton, K. Westerdorff, J. N. Winn

The mass-radius relationships shown in Fig. 9 of this publication were computed with the same version of the interior-atmosphere model used in Appendix A.1. This means that the interior was coupled with the atmosphere by using the grids with atmospheric data (Pluriel et al. 2019). This version of the model is combined with the non-adaptive version of the MCMC to retrieve the core and water mass fractions of HD 108236 b and c.

HD207897 b: A dense sub-Neptune transiting a nearby and bright K-type star ([Heidari et al. 2022](#))

N. Heidari, I. Boisse, J. Orell-Miquel, G. Hébrard, L. Acuña, N. C. Hara, J. Lillo-Box, J. D. Eastman, L. Arnold, N. Astudillo-Defru, V. Adibekyan, A. Bieryla, X. Bonfils, F. Bouchy, T. Barclay, C. E. Brasseur, S. Borgniet, V. Bourrier, L. Buchhave, A. Behmard, C. Beard, N. M. Batalha, B. Courcol, P. Cortés-Zuleta, K. Collins, A. Carmona, I. J. M. Crossfield, A. Chontos, X. Delfosse, S. Dalal, M. Deleuil, O. D. S. Demangeon, R. F. Díaz, X. Dumusque, T. Daylan, D. Dragomir, E. Delgado Mena, C. Dressing, F. Dai, P. A. Dalba, D. Ehrenreich, T. Forveille, B. Fulton, T. Fetherolf, G. Gaisné, S. Giacalone, N. Riazi, S. Hoyer, M. J. Hobson, A. W. Howard, D. Huber, M. L. Hill, L. A. Hirsch, H. Isaacson, J. Jenkins, S. R. Kane, F. Kiefer, R. Luque, D. W. Latham, J. Lubin, T. Lopez, O. Mousis, C. Moutou, G. Montagnier, L. Mignon, A. Mayo, T. Močnik, J. M. A. Murphy, E. Pallé, F. Pepe, E. A. Petigura, J. Rey, G. Ricker, P. Robertson, A. Roy, R. A. Rubenzahl, L. J. Rosenthal, A. Santerne, N. C. Santos, S. G. Sousa, K. G. Stassun, M. Stalport, N. Scarsdale, P. A. Strøm, S. Seager, D. Segransan, P. Tenenbaum, R. Tronsgaard, S. Udry, R. Vanderspek, F. Vakili, J. Winn, and L. M. Weiss

In this work, I combined the non-adaptive version of the MCMC method with the interior-atmosphere model to retrieve the composition of HD 207897 b. In the coupling between the interior and the atmosphere, I used the grids of atmospheric data provided by the atmospheric model of [Pluriel et al. \(2019\)](#).

Masses for the seven planets in K2-32 and K2-233. Four diverse planets in resonant chain and the first young rocky worlds ([Lillo-Box et al. 2020](#))

J. Lillo-Box, T. A. Lopez, A. Santerne, L. D. Nielsen, S. C. C. Barros, M. Deleuil, L. Acuña, O. Mousis, S. G. Sousa, V. Adibekyan, D. J. Armstrong, D. Barrado, D. Bayliss, D. J. A. Brown, X. Dumusque, P. Figueira, S. Hojjatpanah, H. P. Osborn, N. C. Santos, O. D. S. Demangeon, and S. Udry

In this work, I compared the masses and radii of the three planets with mass-radius relationships provided by [Mousis et al. \(2020\)](#), and I concluded that they were likely to be dry, rocky super-Earths. I combined the version of the interior model detailed in Sect. 2.1 and [Brugger et al. \(2017\)](#) with the non-adaptive MCMC algorithm to estimate the core mass fractions with uncertainties for K2-32 e, and K2-233 b and c. The water mass fraction was held constant and equal to zero.

A Reanalysis of the Composition of K2-106b: an Ultra-short Period Super-Mercury Candidate ([Rodríguez Martínez et al. 2022](#))

R. Rodríguez Martínez, B. S. Gaudi, J. G. Schulze, L. Acuña, J. Kolecki, J. A. Johnson, A. P. Asnodkar, K. M. Boley, M. Deleuil, O. Mousis, W. R. Panero, and J. Wang

In this publication, I combined our interior-atmosphere model with the adaptive MCMC to estimate the composition of K2-106 b (see Sect. 5.2). The interior was coupled with the latest version of the atmospheric model (see Chapter 4) for the scenarios that allow for the presence of an envelope (scenarios 2 and 3).

Irradiated Ocean Planets Bridge Super-Earth and Sub-Neptune Populations (Mousis et al. 2020)

O. Mousis, M. Deleuil, A. Aguichine, E. Marcq, J. Naar, L. Acuña Aguirre, B. Brugger, and T. Gonçalves

I contributed to this publication by revising and correcting the implementation of the supercritical water layer in the interior structure model done by J. Naar. I also helped in the generation of plots, and the writing of the interior model section.

TOI-969: a late-K dwarf with a hot mini-Neptune in the desert and an eccentric cold Jupiter (Lillo-Box et al. In rev)

J. Lillo-Box, D. Gandolfi, D.J. Armstrong, K. A. Collins, L. D. Nielsen, R. Luque, J. Korth, S. G. Sousa, S. Quinn, L. Acuña, S. B. Howell, G. Morello, C. Hellier, S. Giacalone, S. Hoyer, K. Stassun, E. Pallé, V. Adibekyan, T. Azevedo Silva, D. Barrado, M. Deleuil, J. D. Eastman, F. Hawthorn, J. M. Irwin, J. M. Jenkins, D. W. Latham, A. Muresan, C.M. Persson, A. Santerne, N. C. Santos, A. Savel, H. P. Osborn, J. Teske, P. J. Wheatley, J. N. Winn, S. C. C. Barros, D. A. Caldwell, D. Charbonneau, R. Cloutier, J. Crane, O. D. S. Demangeon, R. F. Díaz, X. Dumusque, M. Esposito, B. Falk, H. Gill, S. Hojjatpanah, L. Kreidberg, I. Mireles, A. Osborn, G. R. Ricker, J. Rodriguez, R. P. Schwarz, S. Seager, J. Serrano Bell, S. A. Shectman, A. Shporer, M. Vezie, S. X. Wang, G. Zhou

The version of the interior-atmosphere model for water envelopes is the same as in Appendix A.1, where the grids with atmospheric data from Pluriel et al. (2019) are employed. The non-adaptive version of the MCMC is combined with this version of the model to estimate the core and water mass fraction. In addition, I also embed the H/He envelope model described in Sect. 6.3 with the non-adaptive MCMC to retrieve the volatile mass fraction assuming a H/He atmosphere for TOI-969 b.

C. Résumé en français

C.1. Introduction

La première exoplanète avérée, 51 Peg b, a été découverte par [Mayor and Queloz \(1995\)](#), élargissant le champ de la science planétaire au-delà du système solaire. Depuis lors, plus de 5000 exoplanètes¹ ont été détectées, avec des caractéristiques fort différentes de celles du système solaire. En effet, si celles-ci étaient communément classées en trois groupes distinctifs: planètes telluriques (Mercure, Vénus, la Terre et Mars), géantes gazeuses (Jupiter et Saturne), et les géantes glacées (Uranus et Neptune), la population des exoplanètes révèle une grande diversité en termes de masses, de rayons et de conditions d'irradiation. En particulier, la population d'exoplanètes de faible masse (moins de 20 masses terrestres, $20 M_{\oplus}$) suggère l'existence d'au moins deux sous-populations sans équivalents dans notre système solaire: les super-Terres, dont la densité moyenne est similaire à celle de la Terre, et les mini-Neptunes, dont la densité moyenne est compatible avec une composition riche en volatiles. Déterminer la composition de ces différentes familles de planètes est donc nécessaire pour mieux comprendre leur mécanisme de formation et leur évolution.

La composition des exoplanètes ne peut pas être mesurée directement comme nous l'avons fait pour l'intérieur de la Terre. Au lieu de cela, nous en sommes souvent réduits à comparer le rayon et la masse, obtenus via nos techniques de détection, avec des modèles numériques, connus sous le nom de modèles de structure interne. Ces modèles déterminent la façon dont la matière peut se répartir à l'intérieur d'une planète de masse et rayon donnés, en tenant compte de la manière dont les différents matériaux se comportent sous l'effet de la pression et de la température, à différentes profondeurs. Dans le cas des exoplanètes possédant une atmosphère étendue, les télescopes actuels ont permis les premières détections de certains composants clés, tels que l'eau, apportant de premières estimations de leur composition atmosphérique. Ainsi, outre la structure interne, il devient également nécessaire de prendre en compte les effets d'une atmosphère éventuelle au-dessus du manteau, voire d'une couche d'eau à la surface. Depuis les travaux de [Valencia et al. \(2006\)](#) et [Seager et al. \(2007\)](#), plusieurs modèles intérieurs ont été développés pour les super-Terres et les sub-Neptunes ([Lopez and Fortney 2014](#); [Dorn et al. 2015](#); [Brugger et al. 2016](#); [Unterborn et al. 2018](#); [Zeng et al. 2019](#); [Turbet et al. 2020](#); [Marounina and Rogers 2020](#)), et chacun d'entre eux a des implémentations, des hypothèses et des objectifs différents.

Dans cette thèse, nous présentons le développement d'un modèle d'intérieur qui reproduit les propriétés des super-Terre et des sub-Neptunes. Le but de ce modèle est de représenter de manière auto-cohérente les effets de l'atmosphère sur l'intérieur, et vice versa, permettant ainsi une meilleure évaluation de la composition de ces planètes, même dans le cas où celles-ci sont fortement irradiées. Dans l'introduction, nous décrivons le modèle de structure interne (Sect. [C.2](#)) et le modèle d'atmosphère (Sect. [3](#)) avec lequel il a été couplé au moment du démarrage de la thèse, ainsi que ses originalités. La sous-section [C.2.2](#) explique ensuite plus spécifiquement les développements qu'il a fallu réaliser pour coupler de façon cohérente l'intérieur et l'atmosphère. Etant donné que les paramètres planétaires fondamentaux, masse et rayon, sont mesurés avec une

¹NASA Exoplanet Archive, mis à jour le 1er août 2022 (<https://exoplanetarchive.ipac.caltech.edu>).

précision limitée, il était requis d'en tenir compte pour décider, parmi les modèles possibles, quels paramètres décrivent au mieux la composition d'une planète donnée. Nous avons donc développé un algorithme de Monte Carlo qui permet d'évaluer l'erreur associée à chacun des paramètres estimés par le modèle, et ce, dans un temps de calcul raisonnable (Sect. C.2.3). L'application du modèle à différents systèmes planétaires est présentée dans la section C.3: ces applications ont porté sur l'analyse complète de systèmes multiplanètes, tel que TRAPPIST-1, bien connu pour abriter des planètes qui pourraient maintenir de l'eau liquide à leur surface, mais également sur un large échantillon de planètes en transit, caractérisées par des programmes de suivi en vitesses radiales.

Enfin, dans la section C.4, nous discutons des conclusions et des implications des résultats de cette thèse sur les processus de formation planétaires.

C.2. Modèle

C.2.1. Intérieur

Le point de départ de ce travail est le modèle d'intérieur développé par Bastien Brugger (Brugger et al. 2016, 2017). Construit en prenant pour référence les modèles développés pour l'intérieur de la Terre, ce modèle suppose une planète différenciée, composée d'un noyau de fer, d'un manteau riche en silicates, et une couche d'eau structurée en un océan liquide supérieur avec de la glace à haute pression en dessous. C'est donc un modèle bien adapté, au premier ordre, au cas des planètes terrestres.

L'idée d'une exoplanète possédant un océan global pouvant atteindre 100 km (la plus grande profondeur d'océan de la Terre est de 11 km) a été proposée au début des années 2000 par (Léger et al. 2004) qui ont ainsi proposé la dénomination de "planète océan". Ce modèle n'était cependant applicable stricto sensu qu'aux planètes tempérées, la température et la pression de surface devant être compatibles avec l'eau à l'état liquide. Cependant, de nombreuses planètes que nous découvrons et caractérisons encore actuellement, sont très proches de leur étoile, ce qui augmente leur température de surface au-dessus du point d'ébullition de l'eau. Si un océan se trouvait à la surface de ces planètes proches de leur étoile et donc fortement irradiées, il entrerait en ébullition jusqu'à ce que tout son contenu en eau soit passé en phase gazeuse, formant ainsi une atmosphère de vapeur. Par conséquent, il était indispensable d'adapter le modèle original au cas des planètes irradiées. La première étape de cette adaptation a consisté à rajouter pour la couche d'eau, la phase correspondant à des pressions et températures élevées. Cette phase est la phase dite "supercritique", qui se trouve au-dessus du point critique de l'eau. Dans ce régime, il n'y a plus de distinction claire entre les phases liquide et gazeuse. En fait, l'eau se comporte comme un fluide, et présente simultanément les propriétés d'un liquide et d'un gaz. Par exemple, l'eau supercritique peut traverser des solides comme un gaz, et dissoudre des matériaux comme un liquide. En outre, la densité de l'eau supercritique change radicalement par rapport à celle de l'eau liquide. Dans les modèles d'intérieurs, la fonction qui exprime la densité d'un matériau en tant que variable dépendant de la pression et de la température, est appelée "équation d'état" (*Equation of State*, EOS). Nous avons donc remplacé l'EOS de l'eau liquide et de la glace dans le modèle original de Brugger et al. (2017) par l'EOS de l'eau supercritique, obtenue par des expériences en laboratoire et des simulations de mécanique quantique par Mazevet et al. (2019). La température à l'intérieur de la planète est calculée en supposant que la chaleur est conservée lorsqu'elle est transportée par convection, ce qui signifie qu'il s'agit d'un processus adiabatique. Le profil adiabatique est calculé

en utilisant un paramètre nommé paramètre de Grüneisen, dont la valeur dépend du matériau et de sa phase. Nous avons également adopté ce paramètre tel que proposé par [Mazevet et al. \(2019\)](#) par souci de cohérence. Nous avons comparé les effets de la modification de l'EOS et du paramètre de Grüneisen obtenus par d'autres travaux qui sont basés sur des expériences ou sur des calculs théoriques seuls, sur la densité planétaire totale. Nous avons constaté que ces derniers peuvent produire des différences allant jusqu'à 20% du rayon, et que ces autres formulations telles que celles de [\(Duan and Zhang 2006\)](#) surestiment le rayon, produisant des planètes avec intérieurs moins denses, en particulier pour les planètes dont la fraction de masse d'eau est supérieure à 20%.

C.2.2. Atmosphère

Le modèle d'intérieur peut calculer la température, la densité, la masse enfermée dans un certain rayon et l'épaisseur de la couche supercritique jusqu'à un minimum de 300 bars (la pression à la surface de la Terre est de 1 bar). À des pressions inférieures, l'eau commence à devenir transparente à la lumière stellaire, ce qui modifie la manière dont l'eau gazeuse (qui n'est plus supercritique) transporte la chaleur. Le processus qui définit la façon dont la chaleur est distribuée dans l'atmosphère est le transfert radiatif, et les calculs nécessaires pour résoudre l'équation du transfert radiatif sont effectués par un modèle d'atmosphère.

Dans un premier temps, le modèle de structure interne a été couplé de façon cohérente avec une grille de modèles d'atmosphère développés par E. Marcq et son équipe ([Marcq 2012](#); [Marcq et al. 2017](#); [Pluriel et al. 2019](#)). Ce couplage a permis de déterminer des relations masse-rayon réalistes pour des planètes riches en eau et fortement irradiées ([Mousis et al. 2020](#)). Néanmoins, l'utilisation de ces grilles présentait l'inconvénient de devoir gérer de grands ensembles de données et de ne pas offrir une flexibilité suffisante dans le modèle de planète pour permettre de simuler des planètes à l'atmosphère peu épaisse. Pour cette raison, j'ai développé un modèle d'atmosphère qui peut générer des spectres d'émission et calculer l'albédo de Bond simultanément aux calculs du modèle d'intérieur, modèle basé sur le modèle atmosphérique de [Marcq et al. \(2017\)](#).

Notre modèle atmosphérique prend en entrée la pression et la température de surface, qui sont nécessaires pour calculer d'abord la dépendance de la température à la pression à différentes altitudes de l'atmosphère. Ceci est obtenu en supposant que la région proche du bas de l'atmosphère est convective, comme la troposphère de la Terre, où la convection est responsable de la circulation des flux d'air et du transport des nuages. Si la température est suffisamment basse, la convection se produit en même temps que la condensation, formant des nuages. Au-dessus de la couche convective, nous modélisons une mésosphère à température constante, où le transport de chaleur est dominé par la radiation au lieu de la convection. Immédiatement après avoir calculé le profil pression-température, l'épaisseur de l'atmosphère est estimée en résolvant l'équation de l'équilibre hydrostatique. Cette équation décrit l'équilibre entre la force gravitationnelle que l'intérieur de la planète exerce sur l'atmosphère, et la pression interne du gaz atmosphérique. Un point discret (ou une sous-couche) de l'atmosphère à pression et température constantes, absorbe, émet et diffuse la lumière via différents processus radiatifs. Ces mécanismes sont l'absorption induite par collision (*Collision-induced absorption*, CIA), l'absorption des raies, la diffusion Rayleigh et la diffusion des nuages.

Absorption induite par collision (CIA). Elle est causée par les collisions inélastiques des molécules d'un gaz. Dans les collisions inélastiques, l'énergie totale du système, qui est transportée par le moment des molécules avant leur collision, n'est pas conservée dans le processus. Cela signifie qu'un pourcentage de l'énergie est perdu lors de la collision et émis sous forme de chaleur. Cette chaleur modifie la transparence (ou l'opacité) du gaz aux radiations externes, ce qui doit être pris

en compte lors du calcul de son opacité totale, en particulier dans les gaz denses comme l'eau. Le CIA contribue au continuum du spectre d'émission.

Absorption des raies. Les atomes et les molécules émettent et absorbent de l'énergie à différentes fréquences en fonction des transitions quantiques de leurs électrons. Par conséquent, les raies spectrales peuvent être perçues comme une empreinte digitale unique d'un gaz particulier. Les raies spectrales de chaque gaz individuel doivent être prises en compte lors du calcul de l'opacité finale d'un mélange de gaz. Dans notre modèle, ces deux gaz sont l'eau et le dioxyde de carbone (CO₂).

Diffusion de Rayleigh et des nuages. Lorsque les atomes et les molécules de l'atmosphère reçoivent une radiation dont la longueur d'onde est supérieure à la taille des particules du gaz, la lumière est diffusée presque uniformément dans toutes les directions. Dans le cas de l'atmosphère terrestre, la taille des particules de l'atmosphère fait que la lumière optique bleue est la longueur d'onde la plus affectée par la diffusion de Rayleigh, donnant au ciel sa couleur bleue caractéristique. De même, les particules des nuages diffusent également la lumière, mais comme leur taille est plus grande que celle des molécules de gaz, les directions vers lesquelles la lumière est dispersée changent. Dans le cas des nuages, la plupart des radiations sont dispersées dans la même direction que le rayon lumineux d'origine.

Tous ces processus sont pris en compte lors du calcul de l'opacité du gaz qui compose l'atmosphère, qui est l'entrée du solveur d'équations de transfert radiatif (Stamnes et al. 2017). La solution finale des équations de transfert radiatif est le spectre d'émission et le spectre de réflexion. Si nous intégrons sur la longueur d'onde le spectre d'émission, nous obtenons la radiation sortant aux grandes longueurs d'onde (*Outgoing longwave radiation*, OLR), qui est l'énergie totale émise par l'atmosphère. D'autre part, l'intégration sur la longueur d'onde du spectre de réflexion donne l'albédo de Bond, qui est le pourcentage de la lumière stellaire que la planète reflète. Celle-ci est égale à 1 si toute la lumière est réfléchie, et à zéro si toute la lumière est absorbée. Si l'énergie émise est égale à l'énergie absorbée, le flux d'énergie net de l'atmosphère est nul, et l'atmosphère est en équilibre radiatif. Par conséquent, pour une quantité donnée d'énergie stellaire reçue au sommet de l'atmosphère, l'OLR et l'albédo de Bond dépendent de la température et de la pression de surface, qui peuvent être réglées avec précision pour établir l'équilibre radiatif dans l'atmosphère.

Dans le cas d'une atmosphère riche en eau, la densité de la couche d'eau supercritique dans le modèle d'intérieur est très sensible aux changements de température et de pression. Cela signifie que la température et la pression au sommet de la couche supercritique doivent être calculées au préalable par le modèle d'atmosphère afin que le rayon final de la planète soit cohérent. J'ai conçu un algorithme qui couple le modèle d'intérieur et le modèle atmosphérique de manière autoconsistante pour résoudre ce problème.

Le modèle couplé d'intérieur-atmosphère prend en entrée la masse totale, les pourcentages de masse qui composent le noyau et la couche d'eau, que l'on appelle fraction de masse du noyau (*Core mass fraction*, CMF), et fraction de masse de l'eau (*Water mass fraction*, WMF) respectivement. La luminosité stellaire et la distance de l'étoile sont aussi des paramètres d'entrée pour estimer l'énergie stellaire reçue au sommet de l'atmosphère. Les données finales, calculées par le modèle, sont le rayon total (et donc la densité), l'énergie émise et réfléchie par l'atmosphère, ainsi que la pression et la température de surface. Il est important de noter que la pression de surface est liée à la masse de l'atmosphère, qui contribue à la fraction de masse totale de l'eau.

C.2.3. Monte-Carlo par chaînes de Markov

Le modèle intérieur-atmosphère que j'ai décrit dans les sections précédentes est un modèle direct (*forward model* en anglais), ce qui signifie qu'il prend un ensemble de variables d'entrée qui ne sont pas nécessairement observables. Afin de déterminer les erreurs robustes pour les paramètres que le modèle ajuste, nous avons employé la méthode de statistique bayésienne de Monte-Carlo par chaînes de Markov (*Markov Chain Monte Carlo*, MCMC) (Mosegaard and Tarantola 1995; Dorn et al. 2015). Dans le cas spécifique de notre modèle, les observables sont la masse et le rayon de la planète, et les paramètres estimés par le modèle sont ceux qui décrivent la composition interne: CMF et WMF, et les paramètres atmosphériques: conditions de surface, épaisseur de l'atmosphère et albédo de Bond.

Dans un premier temps, et pour le cas du seul modèle de structure interne, nous avons suivi une approche similaire à celle décrite par Dorn et al. (2015). La première étape de l'algorithme consiste à proposer un ensemble de paramètres d'entrée (masse, fractions de masse du noyau et de l'eau) et à calculer le modèle correspondant pour obtenir le rayon correspondant. Pour quantifier dans quelle mesure cet ensemble de paramètres d'entrée correspond aux paramètres observés de la planète, nous calculons la fonction de vraisemblance logarithmique. Cette fonction est inversement proportionnelle à la somme des résidus quadratiques divisée par les erreurs. En d'autres termes, cette fonction diminuera si la différence entre le rayon et la masse observés et calculés augmente. Un nouvel ensemble de paramètres d'entrée est ensuite tiré au hasard et leur log-vraisemblance est comparée de nouveau à celle de l'ensemble initial de paramètres d'entrée. Intuitivement, si la vraisemblance du second ensemble est plus grande que celle de l'ensemble original, l'algorithme adopte comme nouvel état le deuxième ensemble; tandis que s'il est inférieur, le deuxième ensemble est écarté et un nouvel ensemble de paramètres d'entrée est à nouveau tiré au hasard pour être comparé à l'ensemble original. Dans le premier cas, on dit que l'on accepte le modèle proposé, en avançant d'un pas dans la chaîne. Dans le second cas, le modèle proposé est rejeté. Cependant, dans les algorithmes MCMC, la différence entre les deux vraisemblances est obtenue et comparée à un nombre aléatoire tiré d'une distribution uniforme indépendante. Si la différence entre les vraisemblances est supérieure à ce nombre aléatoire, la proposition est acceptée. Cela signifie que la proposition n'est pas toujours acceptée lorsqu'elle a une plus grande vraisemblance que l'état initial, puisque l'algorithme MCMC introduit un élément aléatoire dans la décision. Le nombre de modèles dont les vraisemblances doivent être évaluées pour être comparées à un seul état original, constitue les étapes d'une chaîne. Lorsqu'une proposition est acceptée, une nouvelle chaîne est lancée. Les étapes exposées ci-dessus sont répétées jusqu'à ce qu'un nombre fixe de propositions (ou chaînes) acceptées soit atteint. Après un nombre suffisamment important de chaînes, on obtient l'histogramme des propositions acceptées et les distributions de probabilité a posteriori (*Posterior distribution functions*, PDF) des variables ajustées par le modèle. Par conséquent, étant donné les valeurs moyennes et les barres d'erreur de la masse et du rayon de la planète, nous pouvons estimer les valeurs moyennes et les incertitudes des fractions de masse du noyau et de l'eau, comme la moyenne et l'écart-type de leurs PDF respectives.

Dans l'implémentation la plus simple du MCMC, le modèle proposé est généré en ajoutant une petite augmentation à la valeur originale de l'entrée du modèle, explorant l'espace des paramètres à partir d'une distribution uniforme. Dans la première implémentation du MCMC que j'ai réalisée, cette distribution était toujours la même, quelle que soit la position dans l'espace des paramètres. Cependant, la probabilité de choisir une valeur dans l'espace des paramètres est directement proportionnelle à son nombre de voisins. En d'autres termes, une fraction de masse d'eau de 0.01 (1%) a moins de chances d'être choisie qu'une fraction de masse d'eau de 0.50 (50%), qui se trouve

en plein centre de la plage des valeurs possibles de la fraction de masse d'eau. Cela conduit à ce que les zones de plus faible probabilité, à savoir les valeurs proches des limites 0 et 1, ne soient pas échantillonnées efficacement, ce qui sous-estime l'écart-type des PDFs et biaise les estimations. Pour résoudre ce problème, les limites de la distribution uniforme d'où est tiré le pas itératif sont adaptées en fonction de l'état initial: ces limites sont diminuées quand on est proche de $WMF = 0$, et augmentées si on se trouve dans des valeurs intermédiaires $WMF = 0.30 - 0.70$. C'est ce qu'on appelle un MCMC adaptatif (Director et al. 2017), et la technique est particulièrement utile pour les planètes rocheuses de la taille de la Terre et les super-Terres, dont les WMF sont proches de zéro, mais néanmoins importantes pour déterminer leurs pressions de surface (< 300 bar).

C.3. Analyse de systèmes exoplanétaires

Tout au long de son développement, le code a été utilisé pour analyser différents systèmes planétaires, dans le cadre d'études spécifiques ou de l'analyse de planètes caractérisées par des programmes d'observation sol ou spatiaux. Certains de ces systèmes sont des systèmes compacts dans lesquels plusieurs petites planètes transitent leur étoile. D'autres ont un nombre de planètes connues restreint. L'ensemble des planètes que nous avons analysées couvre toute la gamme de la population des petites planètes, et les étoiles hôtes ont des caractéristiques très différentes. Après avoir présenté les analyses de ces planètes, nous verrons ce qu'elles nous apportent comme informations quant aux mécanismes de formation et d'évolution de la population des petites planètes.

C.3.1. TRAPPIST-1

TRAPPIST-1 est une étoile naine rouge ultra-froide, et peu lumineuse, qui héberge sept planètes de la taille de la Terre (Gillon et al. 2016, 2017), dont trois se trouvent dans la zone habitable de l'étoile (*Habitable zone*, HZ). Celle-ci correspond à la plage de distances de l'étoile hôte à laquelle, si la planète possède de l'eau à sa surface, celle-ci pourrait être à l'état liquide. Toutes les planètes du système de TRAPPIST-1 ont des masses et des rayons mesurés avec une bonne précision (Agol et al. 2021), ce qui fait de ce système un cas bien adapté pour déterminer leur composition. Cette analyse peut apporter des informations sur les mécanismes de formation et d'évolution des planètes de faible masse dans le même environnement initial.

Dans notre analyse, nous avons considéré deux scénarios: dans le scénario 1, les paramètres observés sont la masse et le rayon de chaque planète; dans le scénario 2, nous avons également considéré comme observable le rapport Fe/Si de l'étoile. Ce paramètre, déduit à partir de l'analyse spectrale détaillée de l'étoile, peut être utilisé pour mieux contraindre la masse du noyau par rapport au manteau, car le fer et le silicium se trouvent principalement dans le noyau et le manteau, respectivement. A défaut de mesures directes dans les planètes, l'utilisation des abondances stellaires dans la modélisation de l'intérieur est justifiée par des simulations de formation de planètes, incluant la chimie de la nébuleuse protoplanétaire, et qui montrent que les abondances en Fe et Si de l'étoile et de la nébuleuse sont conservées à l'intérieur de la planète (Thiabaud et al. 2015). Dans le cas de planètes rocheuses avec des masses et des rayons précis, le rapport Fe/Si est utile pour réduire l'ensemble des compositions d'intérieur possibles.

Pour le système de TRAPPIST-1, nous constatons que les planètes les plus denses du système ont des densités similaires à celle de la Terre. Cela signifie que les planètes n'ont pas perdu une partie de leur manteau. Un exemple de perte de manteau dans le système solaire est Mercure, dont

l'abondance Fe/Si est plus élevée que pour les autres planètes telluriques (à savoir la Terre et Vénus). La perte du manteau peut être causée par des collisions qui arrachent la couche la plus externe de l'intérieur après la formation de la planète, ou par l'évaporation du manteau, dans un processus lors duquel la radiation de l'étoile chauffe le manteau et l'évapore. Dans le scénario 1, où le rapport Fe/Si est déduit par le modèle, à partir des rapports Fe/Si individuels nous avons estimé un Fe/Si commun pour toutes les planètes du système de 0.45 - 0.97, valeur qui est également en accord avec le rapport Fe/Si mesuré de l'étoile hôte. Ce rapport Fe/Si correspond à une fraction de masse des noyaux planétaires du système de 23% à 40%, intervalle qui inclut la valeur de la fraction de masse de la Terre (32%).

Pour la fraction de masse d'eau estimée pour chaque planète du système, nous avons trouvé que celle-ci augmente au fur et à mesure que l'on s'éloigne de l'étoile pour les planètes internes (b à e). Cette augmentation est suivie d'un palier, avec une fraction de masse d'eau constante pour les planètes les plus externes (f à h). TRAPPIST-1 d semble être la seule exception à cette tendance au sein du système, puisqu'elle présente un contenu en eau légèrement supérieur à celui de la planète e dans notre analyse initiale, où nous avons considéré que son hydrosphère est une couche d'eau condensée pure. Cependant, on peut souligner que nous avons constaté que la densité et le flux incident de TRAPPIST-1 d sont également compatibles avec une atmosphère dominée par le CO₂ en équilibre, avec une pression de surface de 300 bar. Ceci suggère que l'eau pourrait être présente sous forme gazeuse dans une atmosphère secondaire, donnant une fraction de masse volatile inférieure à celle de la planète e.

Compte tenu de sa densité et de ses conditions d'irradiation, il est finalement peu probable que la planète d présente un océan d'eau liquide, bien que ce scénario ne puisse être complètement écarté tant que la composition de l'atmosphère ne sera pas connue. Les planètes e, f et g ont de meilleures perspectives en termes d'habitabilité. L'hydrosphère de ces planètes pourrait être stratifiée en une couche superficielle de glace ou d'eau liquide. Dans le cas des glaces à basse pression, leur base pourrait être fondue par la chaleur fournie par le manteau, et former une couche océanique inférieure (Noack et al. 2016). Pour une fraction de masse d'eau supérieure à 10%, il y a 50% de chances qu'une couche habitable d'eau liquide se forme sous la surface. Pour une fraction de masse d'eau supérieure à 14%, l'hydrosphère est inhabitable. Dans notre scénario le plus conservateur (scénario 1), les planètes e à g atteignent ces valeurs dans leurs barres d'erreur, bien que leurs valeurs minimales s'étendent jusqu'à 0-3% d'eau, ce qui les placerait dans le régime habitable.

C.3.2. Systèmes multiplanétaires

Nous avons effectué une analyse similaire avec un échantillon d'autres systèmes multiplanétaires afin de déterminer leur composition, et vérifier si d'autres systèmes présentent une tendance similaire à celle du système de TRAPPIST-1 pour leur fraction de masse d'eau.

Cette étude nous a permis de montrer que les systèmes multiplanétaires présentent une nette dichotomie entre les planètes intérieures et sèches, et les planètes extérieures, riches en matières volatiles. Elle a également mis en évidence des disparités de composition d'un système à l'autre.

Le système K2-138 présente la même tendance d'un gradient plus plateau que nous avons observée dans TRAPPIST-1. Kepler-102 pourrait potentiellement présenter une fraction de masse d'eau constante dans la partie extérieure du système également, étant donné les incertitudes sur les paramètres fondamentaux de ses planètes les plus extérieures. D'autres systèmes ne présentent qu'un gradient comme tendance de leur contenu en eau, comme Kepler-11, Kepler-80 et les planètes intérieures de TOI-178.

Les systèmes multiplanétaires présentent une diversité dans leurs contenus en volatiles. Dans

certaines de ces systèmes, une ou deux planètes semblent faire exception à une tendance générale dans le système en question. Nous sommes en mesure d'expliquer les exceptions à la tendance particulière de chaque système par l'échappement atmosphérique de Jeans et/ou la photoévaporation due aux radiations X et ultraviolet (UV). Ces deux mécanismes conduisent en effet la planète à perdre une quantité significative de son atmosphère. Dans l'échappement atmosphérique de Jeans, l'énergie cinétique des atomes et des molécules dépasse l'énergie gravitationnelle de la planète. Les molécules ayant suffisamment d'énergie cinétique s'échappent du fait de leur agitation thermique (Jeans 1925). Dans la photoévaporation XUV, les couches supérieures de l'atmosphère sont chauffées par le rayonnement stellaire incident. Les gaz atmosphériques ainsi chauffés se dilatent, ce qui entraîne l'accélération des molécules et leur fuite vers l'espace (Owen and Jackson 2012).

Enfin, bien qu'à ce jour notre modèle d'atmosphère soit encore limité aux atmosphères riches en eau ou CO₂, nous avons pu identifier les planètes de nos systèmes qui présentent vraisemblablement une atmosphère riche en H/He. Pour ce faire, nous comparons le rayon obtenu et celui observé: si le rayon moyen obtenu par le modèle est significativement inférieur au rayon observé, cela signifie que la densité réelle de la planète est inférieure à celle de notre planète avec hydrosphère telle que simulée. Par conséquent, son atmosphère doit être composée d'un élément moins dense et plus volatile que l'eau, comme l'hydrogène et l'hélium.

Enfin, nous notons également que les planètes sèches avec des fractions de masse de noyau élevées sont toujours situées dans la partie la plus interne du système planétaire. Ces planètes ont des fractions de masse de noyau similaires à celle de Mercure (70% de noyau) ou plus, ce qui signifie qu'elles ont pu subir une perte du manteau, ou qu'elles se sont formées dans des zones de la nébuleuse protoplanétaire où le fer était très abondant. Selon les modèles de formation de planètes, ces régions se trouveraient à proximité de la ligne de roche (Aguichine et al. 2020; Scora et al. 2020), qui est l'orbite à laquelle les éléments réfractaires (Fe et Si) atteignent leur température de condensation. Cela correspond à la température à laquelle les roches de silicate et de fer s'évaporent en gaz.

Ces résultats apportent des contraintes utiles pour mieux comprendre la formation et de l'évolution des planètes de faible masse dans les systèmes multiplanétaires. L'augmentation du contenu en eau avec la distance croissante à l'étoile pour les planètes interne de ces systèmes planétaires pourrait être due à la photoévaporation XUV ou à une perte de masse alimentée par le noyau (*core-powered mass loss*). Dans ce dernier mécanisme, le noyau de la planète émet de la chaleur qui provient de l'accrétion initiales de roches et de planétésimaux de la nébuleuse protoplanétaire. Le matériau accrété de la nébuleuse contient également de l'hydrogène, formant une atmosphère primordiale. Si cette enveloppe primordiale d'hydrogène se refroidit plus lentement que le noyau, l'énergie thermique émise par le noyau accentue la perte de masse atmosphérique de l'enveloppe (Ginzburg et al. 2016). Dans TRAPPIST-1 et K2-138, la fraction de masse d'eau constante des planètes extérieures peut être produite si les planètes se sont formées au voisinage de la ligne de glace d'eau. De même que pour le fer dans les lignes de roches, certains auteurs pensent que l'eau est abondante à l'orbite où l'eau commence à se condenser en glace (Mousis et al. 2019, 2021). Par la suite, ces planètes auraient pu migrer vers l'intérieur du système jusqu'à leur position actuelle, mécanisme qui est suggéré par les distances de leurs orbites, qui sont proportionnelles entre elles (Terquem and Papaloizou 2007; Izidoro et al. 2017; Ramos et al. 2017).

C.3.3. Cibles d'intérêt du JWST

Notre modèle d'atmosphère et d'intérieur peut estimer les paramètres atmosphériques de base pour les atmosphères d'eau et de CO₂, en fonction de la masse et du rayon de la planète. Les paramètres atmosphériques que nous sommes en mesure d'estimer sont la pression et la température de surface, l'épaisseur et la masse de l'atmosphère, et l'albédo de Bond. Connaissant les conditions de surface, on peut les utiliser pour générer des spectres d'émission, qui peuvent être ensuite utilisés avec des simulateurs d'observation pour les télescopes actuels, tels que le télescope spatial James Webb (Batalha et al. 2020). Pour générer des spectres d'émission pour le JWST, nous adaptons le modèle d'atmosphère pour augmenter la résolution spectrale. La résolution d'un spectre quantifie la capacité d'un instrument à distinguer deux raies spectrales dont les longueurs d'onde respectives sont très proches. J'ai appliqué cette version haute résolution de notre modèle atmosphérique à deux planètes que l'on propose d'observer avec le JWST: TRAPPIST-1 c et 55 Cancri e.

D'après notre analyse de la composition de ces planètes (Sect. C.3.1), nous constatons que TRAPPIST-1 c pourrait avoir une atmosphère dont la pression de surface est comprise entre 7 et 22 bars, ou pas d'atmosphère du tout. Le poids moléculaire du CO₂ est plus élevé que celui de la vapeur d'eau, ce qui produit une atmosphère plus comprimée pour une pression et une température de surface similaires. En outre, les modèles avec une enveloppe de CO₂ peuvent accommoder une atmosphère plus massive pour TRAPPIST-1 c que les modèles riches en eau. Nous notons que pour trois des filtres photométriques du JWST, les barres d'erreur du flux moyen des atmosphères d'eau et de CO₂ ne se chevauchent pas, ce qui montre que les observations dans ces filtres permettront de faire la différence entre ces deux compositions. Par conséquent, l'observation de TRAPPIST-1 c en émission avec un de ces filtres, comme proposé par Kreidberg et al. (2021), nous permettra de discerner entre les deux cas.

Pour la super-Terre chaude 55 Cancri e, nous observons qu'une enveloppe à 100% CO₂ n'est pas assez étendue pour expliquer sa faible densité, selon notre modélisation de son intérieur. Par conséquent, nous supposons une atmosphère dominée par l'eau puisque l'épaisseur de son atmosphère est similaire à celle d'une atmosphère de silicate, telle que proposée par (Keles et al. 2022) comme composition la plus probable. Les atmosphères silicatées (ou minérales) se forment lorsque la radiation de l'étoile hôte chauffe la surface d'une planète rocheuse sans atmosphère, évaporant sa surface rocheuse sans décaper le manteau, qui peut être composé de roches basaltiques (Kreidberg et al. 2019). Le cas d'une atmosphère dominée par l'eau reproduit bien la masse et le rayon observés, ce qui suggère que l'eau en tant qu'espèce à l'état de trace pourrait être nécessaire pour expliquer la faible densité de 55 Cancri e. D'autre part, la masse de silicates secs purs pourrait expliquer la faible densité planétaire si le noyau et le manteau étaient moins denses que ceux des intérieurs semblables à la Terre, ce qui indique un manteau riche en carbone comme le suggère (Madhusudhan 2012). Nous remarquons que pour les longueurs d'onde inférieures à 3,5 μm , le bruit des instruments du JWST est trop élevé pour distinguer des raies spectrales. Néanmoins, le reste de la gamme spectrale des observations proposées par Hu and Damiano (2021) a un faible niveau de bruit, ce qui rend les caractéristiques spectrales de l'eau faciles à identifier avec JWST pour 55 Cancri e.

C.3.4. Propriétés déduites de la composition des planètes

Nous profitons de l'échantillon complet de planètes que nous avons analysées avec le modèle d'intérieur-atmosphère durant son développement pour dériver des propriétés de la population générale des planètes de faible masse. Cet échantillon est constitué de 46 exoplanètes, comprenant

à la fois des planètes de taille proche de celle de la Terre, des super-Terres et des sub-Neptunes. Nous calculons l’histogramme des rayons observés de notre échantillon, que nous appelons la distribution des rayons. Dans celle-ci, les super-Terres correspondent aux planètes avec des fractions de masse d’eau inférieures à 10%, alors que la population des sub-Neptunes est constituée de planètes avec des fractions de masse d’eau de plus de 20%, et un maximum de 70%. La distribution de WMF pour les sub-Neptunes paraît essentiellement uniforme, avec une interruption possible pour des valeurs de la WMF entre 0.3 et 0.4 Cette absence de planètes avec une WMF entre 30 et 40% pourrait indiquer une transition entre les enveloppes riches en eau et celles riches en H/He. Nous voyons également que les planètes avec une masse inférieure à $2,5 M_{\oplus}$ ne présentent pas de fractions massiques d’eau supérieures à 10%, ce qui pourrait être dû à une fuite atmosphérique de Jeans les empêchant de conserver des enveloppes massives. Enfin, la distribution de la fraction de masse du noyau de notre échantillon contient à la fois des planètes de type terrestre avec des CMFs = 20 à 40%, et des super-Mercures avec un CMF > 80%. Lorsque la fraction de masse du noyau est obtenue en tenant compte du rapport planétaire Fe/Si à partir des abondances stellaires de l’hôte, sa distribution est gaussienne (forme de cloche), en accord avec les études précédentes (Plotnykov and Valencia 2020).

Notre échantillon présente également une large gamme d’incertitudes relatives (barres d’erreur) en masse et en rayon, ce qui nous permet d’analyser comment l’affinement de la précision des paramètres observables peut améliorer la détermination de la composition. Nos résultats sont en accord avec les travaux précédents (Otegi et al. 2020) et montrent que l’augmentation de la précision du rayon pour les sub-Neptunes améliore l’estimation de la fraction de masse de l’eau, mais n’a aucun effet sur la fraction de masse du noyau. De plus, la prise en compte du rapport Fe/Si comme donnée d’entrée permet de mieux contraindre la fraction de masse du noyau pour les sub-Neptunes, mais fournit des estimations très similaires pour la fraction de masse d’eau par rapport à la prise en compte des seules données de masse et de rayon. Dans le cas des super-Terres, l’incertitude du CMF est réduite lorsque la précision du rayon est améliorée et que le rapport Fe/Si est inclus dans la modélisation comme donnée d’entrée.

C.4. Perspectives

Le problème des intérieurs et des compositions planétaires présente des dégénérescences. Cela signifie que différentes solutions existent pour expliquer un même jeu d’observables. Dans notre cas, différentes compositions peuvent en effet donner la même masse, le même rayon et les mêmes densités. Nous avons vu précédemment que pour les super-Terres rocheuses avec des densités précises, le rapport de l’abondance Fe/Si de leur étoile hôte peut briser cette dégénérescence. Cependant, il nous reste une autre dégénérescence qui affecte certaines super-Terres et la plupart des sub-Neptunes. Dans notre modèle, nous avons toujours considéré que l’atmosphère de la planète a une composition soit riche en eau soit en CO₂, mais nous ne pouvons connaître avec certitude la composition de l’atmosphère qu’avec son spectre. Nous ne pouvons pas connaître la composition de l’atmosphère avec un modèle intérieur seul car une enveloppe d’eau et une atmosphère d’hydrogène peuvent donner une épaisseur atmosphérique similaire avec des masses atmosphériques différentes. Si la composition de l’atmosphère est connue, un modèle d’intérieur-atmosphère, tel que celui que nous avons développé dans cette thèse, pourrait être utilisé pour contraindre la pression et la température de surface. Les estimations actuelles de la composition atmosphérique de certaines sub-Neptunes, dérivées de différentes modélisations, suggèrent un mélange de H/He, d’eau et d’autres composés (Madhusudhan et al. 2020; Bézard

et al. 2020; Guzmán-Mesa et al. 2022), tandis que les super-Terres pourraient avoir des compositions atmosphériques plus exotiques, comme des atmosphères minérales (Keles et al. 2022). Par conséquent, les futurs développements du modèle d’atmosphère doivent inclure davantage de gaz légers comme l’hydrogène et l’hélium. En outre, l’algorithme MCMC du modèle pourrait être modifié pour prendre en compte non seulement la masse et le rayon observables de la planète, mais aussi le spectre observable de l’atmosphère de la planète. Cela permettrait de rompre la dégénérescence entre la composition de l’enveloppe et sa pression de surface.

Il a été démontré par différents auteurs que l’abondance d’une espèce chimique particulière dépend de la pression de surface de l’atmosphère d’une planète (Yu et al. 2021; Tsai et al. 2021; Hu et al. 2021). Cela est dû à un processus appelé photochimie: deux ou plusieurs atomes réagissent avec la lumière stellaire reçue au sommet de l’atmosphère pour produire certaines molécules. L’efficacité de la réaction, et donc l’abondance de la molécule finale, dépend de la température et de la pression auxquelles la réaction a lieu. Par conséquent, étant donné le spectre de l’atmosphère d’une planète donnée, nous pouvons modéliser ensemble l’intérieur et la photochimie pour contraindre les conditions de la couche la plus interne de l’atmosphère. Ceci est particulièrement intéressant pour les planètes rocheuses de la zone habitable, car si nous connaissons la densité de la planète et ses conditions de surface, nous pouvons savoir si elle présente un océan liquide où la vie pourrait éventuellement se développer. C’est pourquoi la détection et la caractérisation des planètes tempérées de taille terrestre sont une priorité pour les futures missions, telles que PLATO (Rauer et al. 2014). PLATO sera capable de détecter des planètes tempérées de taille terrestre autour d’étoiles de type solaire et de fournir des mesures précises de leur rayon pour caractériser leur densité. Il s’agira du premier programme sur les planètes habitables tempérées, qui fournira un échantillon de cibles dont on pourra par la suite, obtenir le spectre de l’atmosphère et, sur le plus long terme, rechercher des biosignatures. Les biosignatures sont des combinaisons d’espèces gazeuses à des abondances données qui ne sont pas attendues dans des atmosphères en équilibre chimique, mais qui peuvent être produites par l’action de la vie. Notons que les processus qui pourraient conduire l’atmosphère d’une planète terrestre hors de l’équilibre chimique pourraient également être géologiques. Il est donc essentiel de modéliser le déséquilibre chimique produit par les processus géologiques abiotiques (c’est-à-dire le volcanisme, la météorisation), pour éviter les faux positifs des biosignatures (Truong and Lunine 2021). Certains des instruments du Téléscope géant européen (*Extremely large telescope*, ELT), comme le spectrographe ANDES, permettra d’acquérir le spectre de l’atmosphère de planètes de faible masse pour rechercher des biosignatures (Snellen et al. 2013; Lopez-Morales et al. 2019). Dans les 20 à 30 prochaines années, le nombre de planètes de taille terrestre pour lesquelles nous disposerons de données atmosphériques pourrait augmenter grâce aussi à de futures missions spatiales comme Luvair (The LUVOIR Team 2019), HabEx (Gaudi et al. 2018) et LIFE (Konrad et al. 2021).

Enfin, mentionnons qu’une exploration par sonde in-situ des géantes de glace du système solaire est une priorité pour les 30 à 40 prochaines années. Les intérieurs et les atmosphères d’Uranus et de Neptune sont encore mal compris, car nous ne disposons que de données acquises à distance (Mousis et al. 2018; Fletcher et al. 2020). Une telle mission est essentielle pour briser les dégénérescences dont nous avons parlé plus haut pour les sub-Neptunes, car nos modèles d’intérieur et d’atmosphère pour cette classe de planètes sont basés sur nos connaissances des géantes gazeuses du système solaire, tout comme la Terre et Vénus sont les références pour les modèles de super-Terre.

Bibliography

- Acuña, L., Deleuil, Magali, Mousis, Olivier, Marcq, Emmanuel, Levesque, Maëva, Aguichine, Artyom, 2021. Characterisation of the hydrospheres of trappist-1 planets. *A&A* 647, A53.
- Acuña, L., Lopez, T. A., Morel, T., Deleuil, M., Mousis, O., Aguichine, A., Marcq, E., Santerne, A., 2022. Water content trends in K2-138 and other low-mass multi-planetary systems. *A&A* 660, A102.
- Adibekyan, V., et al., 2021. Composition of super-Earths, super-Mercuries, and their host stars. *Communications of the Byurakan Astrophysical Observatory* 68 (2), 447–453.
- Agol, E., et al., 2021. Refining the transit-timing and photometric analysis of TRAPPIST-1: Masses, radii, densities, dynamics, and ephemerides. *The Planetary Science Journal* 2 (1), 1.
- Agol, E., Steffen, J., Sari, R., Clarkson, W., 2005. On detecting terrestrial planets with timing of giant planet transits. *MNRAS* 359 (2), 567–579.
- Aguichine, A., Mousis, O., Deleuil, M., Marcq, E., 2021. Mass–radius relationships for irradiated ocean planets. *The Astrophysical Journal* 914 (2), 84.
- Aguichine, A., Mousis, O., Devouard, B., Ronnet, T., 2020. Rocklines as Cradles for Refractory Solids in the Protosolar Nebula. *ApJ* 901 (2), 97.
- Amundsen, D. S., Tremblin, P., Manners, J., Baraffe, I., Mayne, N. J., 2017. Treatment of overlapping gaseous absorption with the correlated-k method in hot Jupiter and brown dwarf atmosphere models. *A&A* 598, A97.
- Angelo, I., Hu, R., 2017. A case for an atmosphere on super-earth 55 cancri e. *The Astronomical Journal* 154 (6), 232.
- Asphaug, E., Reufer, A., 2014. Mercury and other iron-rich planetary bodies as relics of inefficient accretion. *Nature Geoscience* 7 (8), 564–568.
- Baraffe, I., Chabrier, G., Barman, T., 2008. Structure and evolution of super-Earth to super-Jupiter exoplanets. I. Heavy element enrichment in the interior. *A&A* 482 (1), 315–332.
- Barr, A. C., Dobos, V., Kiss, L. L., 2018. Interior structures and tidal heating in the TRAPPIST-1 planets. *A&A* 613, A37.
- Barros, S. C. C., et al., 2022. Detection of the tidal deformation of WASP-103b at 3σ with CHEOPS. *A&A* 657, A52.
- Batalha, N., Fix, M., Fraine, J., Zhou, Y., Schlawin, E., AarynnCarter, Cubillos, P., 2020. natasha-batalha/pandexo: Release 1.5.
- Bean, J. L., et al., 2021. Unlocking the Mysteries of the Archetype Sub-Neptune GJ1214b with a Full-Orbit Phase Curve. JWST Proposal. Cycle 1, ID. #1803.

- Benneke, B., et al., 2019. A sub-Neptune exoplanet with a low-metallicity methane-depleted atmosphere and Mie-scattering clouds. *Nature Astronomy* 3, 813–821.
- Benneke, B., et al., 2019. Water Vapor and Clouds on the Habitable-zone Sub-Neptune Exoplanet K2-18b. *ApJ* 887 (1), L14.
- Benz, W., et al., 2021. The CHEOPS mission. *Experimental Astronomy* 51 (1), 109–151.
- Benz, W., Slattery, W. L., Cameron, A. G. W., 1988. Collisional stripping of Mercury’s mantle. *Icarus* 74 (3), 516–528.
- Bézar, B., Charnay, B., Blain, D., 2020. Methane as a dominant absorber in the habitable-zone sub-Neptune K2-18 b. arXiv e-prints, arXiv:2011.10424.
- Bézar, B., Fedorova, A., Bertaux, J.-L., Rodin, A., Korablev, O., 2011. The 1.10- and 1.18- μm nightside windows of Venus observed by SPICAV-IR aboard Venus Express. *Icarus* 216 (1), 173–183.
- Bolmont, E., Selsis, F., Owen, J. E., Ribas, I., Raymond, S. N., Leconte, J., Gillon, M., 2017. Water loss from terrestrial planets orbiting ultracool dwarfs: implications for the planets of TRAPPIST-1. *MNRAS* 464 (3), 3728–3741.
- Bonfanti, A., et al., 2021. CHEOPS observations of the HD 108236 planetary system: a fifth planet, improved ephemerides, and planetary radii. *A&A* 646, A157.
- Borucki, W. J., et al., 2006. The Kepler Mission: A Transit-Photometry Mission to Discover Terrestrial Planets. *ISSI Scientific Reports Series* 6, 207–220.
- Boucher, O., 1998. On aerosol direct shortwave forcing and the henryey–greenstein phase function. *Journal of the Atmospheric Sciences* 55 (1), 128 – 134.
- Bourrier, V., et al., 2018. The 55 cancri system reassessed. *A&A* 619, A1.
- Brewer, J. M., Fischer, D. A., 2018. Spectral Properties of Cool Stars: Extended Abundance Analysis of Kepler Objects of Interest. *ApJS* 237 (2), 38.
- Brewer, J. M., Fischer, D. A., Valenti, J. A., Piskunov, N., 2016. Spectral Properties of Cool Stars: Extended Abundance Analysis of 1,617 Planet-search Stars. *ApJS* 225 (2), 32.
- Brugger, B., 2018. *Structure interne et minéralogie des exoplanètes terrestres de faible masse*. Ph. D. thesis. Thèse de doctorat dirigée par Mousis, Olivier et Deleuil, Magali Astrophysique et Cosmologie Aix-Marseille 2018.
- Brugger, B., Mousis, O., Deleuil, M., Deschamps, E., 2017. Constraints on super-earth interiors from stellar abundances. *The Astrophysical Journal* 850 (1), 93.
- Brugger, B., Mousis, O., Deleuil, M., Lunine, J. I., 2016. POSSIBLE INTERNAL STRUCTURES AND COMPOSITIONS OF PROXIMA CENTAURI b. *The Astrophysical Journal* 831 (2), L16.
- Buchner, J., 2021. Nested Sampling Methods. arXiv e-prints, arXiv:2101.09675.
- Buchner, J., Georgakakis, A., Nandra, K., Hsu, L., Rangel, C., Brightman, M., Merloni, A., Salvato, M., Donley, J., Kocevski, D., 2014. X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue. *A&A* 564, A125.

- Cameron, A. G. W., 1985. The partial volatilization of Mercury. *Icarus* 64 (2), 285–294.
- Cendes, Y., Williams, P. K. G., Berger, E., 2022. A Pilot Radio Search for Magnetic Activity in Directly Imaged Exoplanets. *AJ* 163 (1), 15.
- Chabrier, G., Mazevet, S., Soubiran, F., 2019. A new equation of state for dense hydrogen–helium mixtures. *The Astrophysical Journal* 872 (1), 51.
- Charbonneau, D., Brown, T. M., Noyes, R. W., Gilliland, R. L., 2002. Detection of an Extrasolar Planet Atmosphere. *ApJ* 568 (1), 377–384.
- Charnay, B., et al., 2022. A survey of exoplanet phase curves with Ariel. *Experimental Astronomy* 53 (2), 417–446.
- Chase, J., M. W., 1998. NIST-JANAF Thermochemical Tables, Fourth Edition. *J. Phys. Chem. Ref. Data, Monograph* 9.
- Chaudhry, S., Lautzenheiser, D., Ghosh, K., 2021. An Efficient Scheme for Sampling in Constrained Domains. *arXiv e-prints*, arXiv:2110.10840.
- Christiansen, J. L., et al., 2018. The K2-138 System: A Near-resonant Chain of Five Sub-Neptune Planets Discovered by Citizen Scientists. *AJ* 155 (2), 57.
- Coleman, G. A. L., Leleu, A., Alibert, Y., Benz, W., 2019. Pebbles versus planetesimals: the case of Trappist-1. *A&A* 631, A7.
- Connolly, J. A. D., 2005. Computation of phase equilibria by linear programming: A tool for geodynamic modeling and its application to subduction zone decarbonation. *Earth and Planetary Science Letters* 236 (1-2), 524–541.
- Connolly, J. A. D., 2009. The geodynamic equation of state: What and how. *Geochemistry, Geophysics, Geosystems* 10 (10), Q10014.
- Coustenis, A., Schneider, J., Bockelée-Morvan, D., Rauer, H., Wittemberg, R., Chassefière, E., Greene, T., Penny, A., Guillot, T., 1997. Spectroscopy of 51 Peg B: Search for Atmospheric Signatures. In: Soderblom, D. (Ed.), *Planets Beyond the Solar System and the Next Generation of Space Missions*, Volume 119 of *Astronomical Society of the Pacific Conference Series*, pp. 101.
- Csizmadia, S., Hellard, H., Smith, A. M. S., 2019. An estimate of the k_2 Love number of WASP-18Ab from its radial velocity measurements. *A&A* 623, A45.
- Cyr, K. E., Sears, W. D., Lunine, J. I., 1998. Distribution and Evolution of Water Ice in the Solar Nebula: Implications for Solar System Body Formation. *Icarus* 135 (2), 537–548.
- Dai, F., Masuda, K., Winn, J. N., Zeng, L., 2019. Homogeneous analysis of hot earths: Masses, sizes, and compositions. *The Astrophysical Journal* 883 (1), 79.
- David, T. J., et al., 2018. Three Small Planets Transiting the Bright Young Field Star K2-233. *AJ* 155 (5), 222.
- de Wit, J., et al., 2016. A combined transmission spectrum of the Earth-sized exoplanets TRAPPIST-1 b and c. *Nature* 537 (7618), 69–72.

- de Wit, J., et al., 2018. Atmospheric reconnaissance of the habitable-zone Earth-sized planets orbiting TRAPPIST-1. *Nature Astronomy* 2, 214–219.
- Demory, B.-O., et al., 2016. A map of the large day-night temperature gradient of a super-Earth exoplanet. *Nature* 532 (7598), 207–209.
- Des Marais, D. J., Harwit, M. O., Jucks, K. W., Kasting, J. F., Lin, D. N., Lunine, J. I., Schneider, J., Seager, S., Traub, W. A., Woolf, N. J., 2002. Remote sensing of planetary properties and biosignatures on extrasolar terrestrial planets. *Astrobiology* 2 (2), 153–181. PMID: 12469366.
- Diamond-Lowe, H., Berta-Thompson, Z., Charbonneau, D., Kempton, E. M. R., 2018. Ground-based Optical Transmission Spectroscopy of the Small, Rocky Exoplanet GJ 1132b. *AJ* 156 (2), 42.
- Diamond-Lowe, H., Charbonneau, D., Malik, M., Kempton, E. M. R., Beletsky, Y., 2020. Optical Transmission Spectroscopy of the Terrestrial Exoplanet LHS 3844b from 13 Ground-based Transit Observations. *AJ* 160 (4), 188.
- Director, H. M., Gattiker, J., Lawrence, E., Wiel, S. V., 2017. Efficient sampling on the simplex with a self-adjusting logit transform proposal. *Journal of Statistical Computation and Simulation* 87 (18), 3521–3536.
- Dittmann, J. A., et al., 2017. A temperate rocky super-Earth transiting a nearby cool star. *Nature* 544 (7650), 333–336.
- Dorn, C., Hinkel, N. R., Venturini, J., 2017. Bayesian analysis of interiors of hd 219134b, kepler-10b, kepler-93b, corot-7b, 55 cnc e, and hd 97658b using stellar abundance proxies. *A&A* 597, A38.
- Dorn, C., Khan, A., Heng, K., Connolly, J. A. D., Alibert, Y., Benz, W., Tackley, P., 2015. Can we constrain the interior structure of rocky exoplanets from mass and radius measurements? *A&A* 577, A83.
- Dorn, C., Lichtenberg, T., 2021. Hidden Water in Magma Ocean Exoplanets. *ApJ* 922 (1), L4.
- Dorn, C., Mosegaard, K., Grimm, S. L., Alibert, Y., 2018. Interior Characterization in Multiplanetary Systems: TRAPPIST-1. *The Astrophysical Journal* 865 (1), 20.
- Dorn, C., Noack, L., Rozel, A. B., 2018. Outgassing on stagnant-lid super-Earths. *A&A* 614, A18.
- Dorn, C., Venturini, Julia, Khan, Amir, Heng, Kevin, Alibert, Yann, Helled, Ravit, Rivoldini, Attilio, Benz, Willy, 2017. A generalized bayesian inference method for constraining the interiors of super earths and sub-neptunes. *A&A* 597, A37.
- Driscoll, P., Olson, P., 2011. Optimal dynamos in the cores of terrestrial exoplanets: Magnetic field generation and detectability. *Icarus* 213 (1), 12–23.
- Duan, Z., Møller, N., Weare, J. H., 1996. A general equation of state for supercritical fluid mixtures and molecular dynamics simulation of mixture pvtx properties. *Geochimica et Cosmochimica Acta* 60 (7), 1209–1216.
- Duan, Z., Zhang, Z., 2006. Equation of state of the h₂o, co₂, and h₂o–co₂ systems up to 10 gpa and 2573.15k: Molecular dynamics simulations with ab initio potential surface. *Geochimica et Cosmochimica Acta* 70 (9), 2311–2324.

- Ducrot, E., et al., 2020. TRAPPIST-1: Global results of the Spitzer Exploration Science Program Red Worlds. *A&A* 640, A112.
- Ducrot, E., et al., 2018. The 0.8-4.5 μm Broadband Transmission Spectra of TRAPPIST-1 Planets. *AJ* 156 (5), 218.
- Dunaeva, A. N., Antsyshkin, D. V., Kuskov, O. L., 2010. Phase diagram of H_2O : Thermodynamic functions of the phase transitions of high-pressure ices. *Solar System Research* 44 (3), 202–222.
- Edwards, B., Changeat, Q., Mori, M., Anisman, L. O., Morvan, M., Yip, K. H., Tsiaras, A., Al-Refai, A., Waldmann, I., Tinetti, G., 2021. Hubble WFC3 Spectroscopy of the Habitable-zone Super-Earth LHS 1140 b. *AJ* 161 (1), 44.
- Ehrenreich, D., et al., 2012. Hint of a transiting extended atmosphere on 55 cancri b. *A&A* 547, A18.
- Esteves, L. J., de Mooij, E. J. W., Jayawardhana, R., Watson, C., de Kok, R., 2017. A search for water in a super-earth atmosphere: High-resolution optical spectroscopy of 55cancri e. *The Astronomical Journal* 153 (6), 268.
- Fletcher, L. N., et al., 2020. Ice giant systems: The scientific potential of orbital missions to uranus and neptune. *Planetary and Space Science* 191, 105030.
- Foreman-Mackey, D., Hogg, D. W., Lang, D., Goodman, J., 2013. emcee: The MCMC Hammer. *PASP* 125 (925), 306.
- Fortney, J. J., Marley, M. S., Barnes, J. W., 2007. Planetary Radii across Five Orders of Magnitude in Mass and Stellar Insolation: Application to Transits. *ApJ* 659 (2), 1661–1672.
- Freedman, R. S., Lustig-Yaeger, J., Fortney, J. J., Lupu, R. E., Marley, M. S., Lodders, K., 2014. Gaseous Mean Opacities for Giant Planet and Ultracool Dwarf Atmospheres over a Range of Metallicities and Temperatures. *ApJS* 214 (2), 25.
- French, M., Desjarlais, M. P., Redmer, R., 2016. Ab initio calculation of thermodynamic potentials and entropies for superionic water. *Phys. Rev. E* 93 (2), 022140.
- Fulton, B. J., Petigura, E. A., 2018. The California-Kepler Survey. VII. Precise Planet Radii Leveraging Gaia DR2 Reveal the Stellar Mass Dependence of the Planet Radius Gap. *AJ* 156 (6), 264.
- Fulton, B. J., et al., 2017. The california-ikepler/isurvey. III. a gap in the radius distribution of small planets. *The Astronomical Journal* 154 (3), 109.
- Gao, P., Wakeford, H. R., Moran, S. E., Parmentier, V., 2021. Aerosols in exoplanet atmospheres. *Journal of Geophysical Research: Planets* 126 (4), e2020JE006655. e2020JE006655 2020JE006655.
- García Muñoz, A., Fossati, L., Youngblood, A., Nettelmann, N., Gandolfi, D., Cabrera, J., Rauer, H., 2021. A Heavy Molecular Weight Atmosphere for the Super-Earth π Men c. *ApJ* 907 (2), L36.
- Gaudi, B. S., et al., 2018. The Habitable Exoplanet Observatory (HabEx) Mission Concept Study Interim Report. arXiv e-prints, arXiv:1809.09674.
- Giacobbe, P., et al., 2021. Five carbon- and nitrogen-bearing species in a hot giant planet's atmosphere. *Nature* 592 (7853), 205–208.

- Gillmann, C., et al., 2022. The Long-Term Evolution of the Atmosphere of Venus: Processes and Feedback Mechanisms. *Space Sci. Rev.* 218 (7), 56.
- Gillon, M., et al., 2016. Temperate Earth-sized planets transiting a nearby ultracool dwarf star. *Nature* 533 (7602), 221–224.
- Gillon, M., et al., 2017. Seven temperate terrestrial planets around the nearby ultracool dwarf star TRAPPIST-1. *Nature* 542 (7642), 456–460.
- Ginzburg, S., Schlichting, H. E., Sari, R., 2016. Super-Earth Atmospheres: Self-consistent Gas Accretion and Retention. *ApJ* 825 (1), 29.
- Glasse, A., Rieke, G. H., Bauwens, E., García-Marín, M., Ressler, M. E., Rost, S., Tikkanen, T. V., Vandenbussche, B., Wright, G. S., 2015. The mid-infrared instrument for the James Webb Space Telescope/i, IX: Predicted sensitivity. *Publications of the Astronomical Society of the Pacific* 127 (953), 686–695.
- Goody, R., West, R., Chen, L., Crisp, D., 1989. The correlated-k method for radiation calculations in nonhomogeneous atmospheres. *Journal of Quantitative Spectroscopy and Radiative Transfer* 42, 539–550.
- Gordon, S., 1994. Computer program for calculation of complex chemical equilibrium compositions and applications. part 1: Analysis. Tech. rep. NASA Lewis Research Center.
- Gressier, Mori, M., Changeat, Q., Edwards, B., Beaulieu, J. P., Marcq, E., Charnay, B., 2022. Near-infrared transmission spectrum of trappist-1 h using hubble wfc3 g141 observations. *A&A* 658, A133.
- Grimm, S. L., et al., 2018. The nature of the trappist-1 exoplanets. *A&A* 613, A68.
- Grimm, S. L., Heng, K., 2015. HELIOS-K: An Ultrafast, Open-source Opacity Calculator for Radiative Transfer. *The Astrophysical Journal* 808 (2), 182.
- Guenther, E. W., et al., 2017. K2-106, a system containing a metal-rich planet and a planet of lower density. *A&A* 608, A93.
- Guillot, T., 2010. On the radiative equilibrium of irradiated planetary atmospheres. *A&A* 520, A27.
- Gupta, A., Schlichting, H. E., 2019. Sculpting the valley in the radius distribution of small exoplanets as a by-product of planet formation: the core-powered mass-loss mechanism. *MNRAS* 487 (1), 24–33.
- Guzmán-Mesa, A., Kitzmann, D., Mordasini, C., Heng, K., 2022. Chemical diversity of the atmospheres and interiors of sub-Neptunes: a case study of GJ 436 b. *MNRAS* 513 (3), 4015–4036.
- Haar, L., Gallagher, J. S., Kell, G. S., 1984. NBS/NRC steam tables thermodynamic and transport properties and computer programs for vapor and liquid states of water in SI units.
- Haldemann, J., Alibert, Y., Mordasini, C., Benz, W., 2020. AQUA: a collection of H₂O equations of state for planetary models. *A&A* 643, A105.
- Heidari, N., et al., 2022. HD 207897 b: A dense sub-Neptune transiting a nearby and bright K-type star. *A&A* 658, A176.

- Hellard, H., Csizmadia, S., Padovan, S., Sohl, F., Rauer, H., 2020. HST/STIS Capability for Love Number Measurement of WASP-121b. *ApJ* 889 (1), 66.
- Heller, R., Rodenbeck, K., Hippke, M., 2019. Transit least-squares survey. I. Discovery and validation of an Earth-sized planet in the four-planet system K2-32 near the 1:2:5:7 resonance. *A&A* 625, A31.
- Heng, K., Mendonça, J. M., Lee, J.-M., 2014. Analytical Models of Exoplanetary Atmospheres. II. Radiative Transfer via the Two-stream Approximation. *ApJS* 215 (1), 4.
- Hernandez, J.-A., Caracas, R., Labrosse, S., 2022. Stability of high-temperature salty ice suggests electrolyte permeability in water-rich exoplanet icy mantles. *Nature Communications* 13, 3303.
- Hinkel, N. R., Mamajek, E. E., Turnbull, M. C., Osby, E., Shkolnik, E. L., Smith, G. H., Klimasewski, A., Somers, G., Desch, S. J., 2017. A Catalog of Stellar Unified Properties (CATSUP) for 951 FGK-Stars within 30 pc. *The Astrophysical Journal* 848 (1), 34.
- Hinkel, N. R., Timmes, F. X., Young, P. A., Pagano, M. D., Turnbull, M. C., 2014. Stellar Abundances in the Solar Neighborhood: The Hypatia Catalog. *The Astrophysical Journal* 148 (3), 54.
- Hinkel, N. R., et al., 2016. A Comparison of Stellar Elemental Abundance Techniques and Measurements. *ApJS* 226 (1), 4.
- Hoyer, S., et al., 2021. TOI-220b: a warm sub-Neptune discovered by TESS. *Monthly Notices of the Royal Astronomical Society* 505 (3), 3361–3379.
- Hu, R., Brandeker, A., Damiano, M., Demory, B.-O., Dragomir, D., Ito, Y., Knutson, H. A., Miguel, Y., Zhang, M., 2021. Determining the Atmospheric Composition of the Super-Earth 55 Cancri e. JWST Proposal. Cycle 1, ID. #1952.
- Hu, R., Damiano, M., 2021. Deep Characterization of the Atmosphere of a Temperate Sub-Neptune. JWST Proposal. Cycle 1, ID. #2372.
- Hu, R., Damiano, M., Scheucher, M., Kite, E., Seager, S., Rauer, H., 2021. Unveiling Shrouded Oceans on Temperate sub-Neptunes via Transit Signatures of Solubility Equilibria versus Gas Thermochemistry. *ApJ* 921 (1), L8.
- Hu, R., Ehlmann, B. L., Seager, S., 2012. Theoretical spectra of terrestrial exoplanet surfaces. *The Astrophysical Journal* 752 (1), 7.
- Hubbard, W. B., 1973. The giant planets. In: Cameron, A. G. W. (Ed.), *Cosmochemistry*, Dordrecht, pp. 153–163. Springer Netherlands.
- Ingersoll, A. P., 1969. The Runaway Greenhouse: A History of Water on Venus. *Journal of Atmospheric Sciences* 26 (6), 1191–1198.
- Innes, H., Pierrehumbert, R. T., 2022. Atmospheric dynamics of temperate sub-neptunes. i. dry dynamics. *The Astrophysical Journal* 927 (1), 38.
- Izidoro, A., Ogihara, M., Raymond, S. N., Morbidelli, A., Pierens, A., Bitsch, B., Cossou, C., Hersant, F., 2017. Breaking the chains: hot super-Earth systems from migration and disruption of compact resonant chains. *MNRAS* 470 (2), 1750–1770.

- Jeans, J., 1925. The dynamical theory of gases. Cambridge at the University Press.
- Jin, S., Mordasini, C., Parmentier, V., van Boekel, R., Henning, T., Ji, J., 2014. Planetary Population Synthesis Coupled with Atmospheric Escape: A Statistical View of Evaporation. *ApJ* 795 (1), 65.
- Journaux, B., 2022. Salty ice and the dilemma of ocean exoplanet habitability. *Nature Communications* 13, 3304.
- Kasting, J. F., 1988. Runaway and moist greenhouse atmospheres and the evolution of Earth and Venus. *Icarus* 74 (3), 472–494.
- Katyal, N., Ortenzi, G., Lee Grenfell, J., Noack, L., Sohl, F., Godolt, M., García Muñoz, A., Schreier, F., Wunderlich, F., Rauer, H., 2020. Effect of mantle oxidation state and escape upon the evolution of Earth's magma ocean atmosphere. *A&A* 643, A81.
- Keles, E., et al., 2022. The PEPSI exoplanet transit survey (PETS) I: investigating the presence of a silicate atmosphere on the super-earth 55 Cnc e. *MNRAS* 513 (1), 1544–1556.
- Khorshid, N., Min, M., Désert, J. M., Woitke, P., Dominik, C., 2021. SimAb: A simple, fast and flexible model to assess the effects of planet formation on the atmospheric composition of gas giants. *arXiv e-prints*, arXiv:2111.00279.
- Kimura, T., Ikoma, M., 2020. Formation of aqua planets with water of nebular origin: effects of water enrichment on the structure and mass of captured atmospheres of terrestrial planets. *MNRAS* 496 (3), 3755–3766.
- Kimura, T., Ikoma, M., 2022. Predicted diversity in water content of terrestrial exoplanets orbiting M dwarfs. *Nature Astronomy* 6, 1296–1307.
- Konrad, B. S., et al., 2021. Large Interferometer For Exoplanets (LIFE): III. Spectral resolution, wavelength range and sensitivity requirements based on atmospheric retrieval analyses of an exo-Earth. *arXiv e-prints*, arXiv:2112.02054.
- Kopparapu, R. K., Ramirez, R., Kasting, J. F., Eymet, V., Robinson, T. D., Mahadevan, S., Terrien, R. C., Domagal-Goldman, S., Meadows, V., Deshpande, R., 2013. Habitable Zones around Main-sequence Stars: New Estimates. *The Astrophysical Journal* 765 (2), 131.
- Kreidberg, L., et al., 2021. Hot Take on a Cool World: Does Trappist-1c Have an Atmosphere? *JWST Proposal*. Cycle 1, ID. #2304.
- Kreidberg, L., Bean, J. L., Désert, J.-M., Benneke, B., Deming, D., Stevenson, K. B., Seager, S., Berta-Thompson, Z., Seifahrt, A., Homeier, D., 2014. Clouds in the atmosphere of the super-Earth exoplanet GJ1214b. *Nature* 505 (7481), 69–72.
- Kreidberg, L., et al., 2019. Absence of a thick atmosphere on the terrestrial exoplanet LHS 3844b. *Nature* 573 (7772), 87–90.
- Krissansen-Totton, J., Fortney, J. J., 2022. Predictions for observable atmospheres of trappist-1 planets from a fully coupled atmosphere–interior evolution model. *The Astrophysical Journal* 933 (1), 115.

- Lacis, A. A., Oinas, V., 1991. A description of the correlated-k distribution method for modelling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres. *Journal of Geophysical Research (Atmospheres)* 96, 9027–9064.
- Lammer, H., Selsis, F., Ribas, I., Guinan, E. F., Bauer, S. J., Weiss, W. W., 2003. Atmospheric Loss of Exoplanets Resulting from Stellar X-Ray and Extreme-Ultraviolet Heating. *ApJ* 598 (2), L121–L124.
- Leconte, J., 2021. Spectral binning of precomputed correlated-k coefficients. *A&A* 645, A20.
- Leconte, J., Forget, F., Charnay, B., Wordsworth, R., Pottier, A., 2013. Increased insolation threshold for runaway greenhouse processes on Earth-like planets. *Nature* 504 (7479), 268–271.
- Léger, A., et al., 2009. Transiting exoplanets from the CoRoT space mission. VIII. CoRoT-7b: the first super-Earth with measured radius. *A&A* 506 (1), 287–302.
- Léger, A., et al., 2004. A new family of planets? “Ocean-Planets”. *Icarus* 169 (2), 499–504.
- Leleu, A., et al., 2021. Six transiting planets and a chain of Laplace resonances in TOI-178. *A&A* 649, A26.
- Leleu, A., Lillo-Box, J., Sestovic, M., Robutel, P., Correia, A. C. M., Hara, N., Angerhausen, D., Grimm, S. L., Schneider, J., 2019. Co-orbital exoplanets from close-period candidates: the TOI-178 case. *A&A* 624, A46.
- Levesque, M., 2019. Internal structure of low-mass planets. aix-marseille university. Master thesis supervised by Magali Deleuil and Olivier Mousis.
- Lichtenberg, T., 2021. Redox Hysteresis of Super-Earth Exoplanets from Magma Ocean Circulation. *ApJ* 914 (1), L4.
- Lillo-Box, et al., 2020. Planetary system lhs 1140 revisited with espresso and tess. *A&A* 642, A121.
- Lillo-Box, et al., 2020. Masses for the seven planets in k2-32 and k2-233 - four diverse planets in resonant chain and the first young rocky worlds. *A&A* 640, A48.
- Lincowski, A. P., Meadows, V. S., Crisp, D., Robinson, T. D., Luger, R., Lustig-Yaeger, J., Arney, G. N., 2018. Evolved Climates and Observational Discriminants for the TRAPPIST-1 Planetary System. *ApJ* 867 (1), 76.
- Linder, E. F., Mordasini, C., Mollière, P., Marleau, G.-D., Malik, M., Quanz, S. P., Meyer, M. R., 2019. Evolutionary models of cold and low-mass planets: cooling curves, magnitudes, and detectability. *A&A* 623, A85.
- Liou, K. N., 1980. *An introduction to atmospheric radiation*.
- Lissauer, J. J., et al., 2011. A closely packed system of low-mass, low-density planets transiting Kepler-11. *Nature* 470 (7332), 53–58.
- Liu, B., Lambrechts, M., Johansen, A., Pascucci, I., Henning, T., 2020. Pebble-driven planet formation around very low-mass stars and brown dwarfs. *A&A* 638, A88.
- Liu, Beibei, Lambrechts, Michiel, Johansen, Anders, Liu, Fan, 2019. Super-earth masses sculpted by pebble isolation around stars of different masses. *A&A* 632, A7.

- Lodders, K., Palme, H., Gail, H. P., 2009. Abundances of the Elements in the Solar System. *Landolt-Börnstein* 4B, 712.
- Lopez, E. D., Fortney, J. J., 2014. Understanding the Mass-Radius Relation for Sub-neptunes: Radius as a Proxy for Composition. *ApJ* 792 (1), 1.
- Lopez-Morales, M., et al., 2019. Detecting Earth-like Biosignatures on Rocky Exoplanets around Nearby Stars with Ground-based Extremely Large Telescopes. *BAAS* 51 (3), 162.
- Luck, R. E., 2016. ABUNDANCES IN THE LOCAL REGION. II. f, g, AND k DWARFS AND SUBGIANTS. *The Astronomical Journal* 153 (1), 21.
- Lupu, R. E., Zahnle, K., Marley, M. S., Schaefer, L., Fegley, B., Morley, C., Cahoy, K., Freedman, R., Fortney, J. J., 2014. The Atmospheres of Earthlike Planets after Giant Impact Events. *The Astrophysical Journal* 784 (1), 27.
- Lustig-Yaeger, J., Meadows, V. S., Lincowski, A. P., 2019. The detectability and characterization of the TRAPPIST-1 exoplanet atmospheres with JWST. *The Astronomical Journal* 158 (1), 27.
- Lynch, C. R., Murphy, T., Lenc, E., Kaplan, D. L., 2018. The detectability of radio emission from exoplanets. *MNRAS* 478 (2), 1763–1775.
- Ma, Q., Tipping, R. H., 1992. A far wing line shape theory and its application to the water vibrational bands (II). *The Journal of Chemical Physics* 96 (12), 8655–8663.
- MacDonald, M. G., et al., 2016. A Dynamical Analysis of the Kepler-80 System of Five Transiting Planets. *AJ* 152 (4), 105.
- MacDonald, M. G., Shakespeare, C. J., Ragozzine, D., 2021. A Five-Planet Resonant Chain: Reevaluation of the Kepler-80 System. *AJ* 162 (3), 114.
- MacDonald, R. J., Marley, M. S., Fortney, J. J., Lewis, N. K., 2018. Exploring hsub2/subo prominence in reflection spectra of cool giant planets. *The Astrophysical Journal* 858 (2), 69.
- Madhusudhan, N., 2012. C/O RATIO AS a DIMENSION FOR CHARACTERIZING EXOPLANETARY ATMOSPHERES. *The Astrophysical Journal* 758 (1), 36.
- Madhusudhan, N., Constantinou, S., Moses, J. I., Piette, A., Sarkar, S., 2021. Chemical Disequilibrium in a Temperate sub-Neptune. *JWST Proposal*. Cycle 1, ID. #2722.
- Madhusudhan, N., Nixon, M. C., Welbanks, L., Piette, A. A. A., Booth, R. A., 2020. The Interior and Atmosphere of the Habitable-zone Exoplanet K2-18b. *ApJ* 891 (1), L7.
- Malik, M., et al., 2017. HELIOS: An Open-source, GPU-accelerated Radiative Transfer Code for Self-consistent Exoplanetary Atmospheres. *The Astrophysical Journal* 153 (2), 56.
- Malik, M., Kitzmann, D., Mendonça, J. M., Grimm, S. L., Marleau, G.-D., Linder, E. F., Tsai, S.-M., Heng, K., 2019. Self-luminous and Irradiated Exoplanetary Atmospheres Explored with HELIOS. *The Astrophysical Journal* 157 (5), 170.
- Mansfield, M., Bean, J. L., Kempton, E. M. R., Kite, E. S., Koll, D., Malik, M., 2021. Constraining the Atmosphere of the Terrestrial Exoplanet Gl486b. *JWST Proposal*. Cycle 1, ID. #1743.

- Marcq, E., 2012. A simple 1-d radiative-convective atmospheric model designed for integration into coupled models of magma ocean planets. *Journal of Geophysical Research: Planets* 117 (E1).
- Marcq, E., Bézard, B., Drossart, P., Piccioni, G., Reess, J. M., Henry, F., 2008. A latitudinal survey of CO, OCS, H₂O, and SO₂ in the lower atmosphere of Venus: Spectroscopic studies using VIRTIS-H. *Journal of Geophysical Research (Planets)* 113, E00B07.
- Marcq, E., Salvador, A., Massol, H., Davaille, A., 2017. Thermal radiation of magma ocean planets using a 1-D radiative-convective model of H₂O-CO₂ atmospheres. *Journal of Geophysical Research (Planets)* 122 (7), 1539–1553.
- Marcy, G. W., et al., 2014. Masses, Radii, and Orbits of Small Kepler Planets: The Transition from Gaseous to Rocky Planets. *ApJS* 210 (2), 20.
- Marleau, G.-D., Coleman, G. A. L., Leleu, A., Mordasini, C., 2019. Exploring the formation by core accretion and the luminosity evolution of directly imaged planets. The case of HIP 65426 b. *A&A* 624, A20.
- Marounina, N., Rogers, L. A., 2020. Internal Structure and CO₂ Reservoirs of Habitable Water Worlds. *ApJ* 890 (2), 107.
- Mayor, M., et al., 2003. Setting New Standards with HARPS. *The Messenger* 114, 20–24.
- Mayor, M., Queloz, D., 1995. A Jupiter-mass companion to a solar-type star. *Nature* 378 (6555), 355–359.
- Mazevet, S., Licari, A., Chabrier, G., Potekhin, A. Y., 2019. Ab initio based equation of state of dense water for planetary and exoplanetary modeling. *A&A* 621, A128.
- McBride, B. J. G., 1996. Computer program for calculation of complex chemical equilibrium compositions and applications ii. users manual and program description. Tech. rep. NASA Lewis Research Center.
- McKay, A. J., et al., 2019. The Peculiar Volatile Composition of CO-dominated Comet C/2016 R2 (PanSTARRS). *The Astrophysical Journal* 158 (3), 128.
- Miguel, Y., Cridland, A., Ormel, C. W., Fortney, J. J., Ida, S., 2020. Diverse outcomes of planet formation and composition around low-mass stars and brown dwarfs. *MNRAS* 491 (2), 1998–2009.
- Miller-Ricci, E., Fortney, J. J., 2010. The Nature of the Atmosphere of the Transiting Super-Earth GJ 1214b. *ApJ* 716 (1), L74–L79.
- Misener, W., Schlichting, H. E., 2022. The importance of silicate vapour in determining the structure, radii, and envelope mass fractions of sub-Neptunes. *MNRAS* 514 (4), 6025–6037.
- Mollière, P., et al., 2022. Interpreting the Atmospheric Composition of Exoplanets: Sensitivity to Planet Formation Assumptions. *ApJ* 934 (1), 74.
- Mollière, P., Wardenier, J. P., van Boekel, R., Henning, T., Molaverdikhani, K., Snellen, I. A. G., 2019. petitRADTRANS. A Python radiative transfer package for exoplanet characterization and retrieval. *A&A* 627, A67.

- Mollière, P. M., 2017. PhD dissertation. PhD thesis, Heidelberg University.
- Montet, B. T., Morton, T. D., Foreman-Mackey, D., Johnson, J. A., Hogg, D. W., Bowler, B. P., Latham, D. W., Bieryla, A., Mann, A. W., 2015. Stellar and Planetary Properties of K2 Campaign 1 Candidates and Validation of 17 Planets, Including a Planet Receiving Earth-like Insolation. *ApJ* 809 (1), 25.
- Mordasini, C., 2018. *Planetary Population Synthesis*, pp. 1–50. Springer International Publishing, Cham.
- Mosegaard, K., Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research (Solid Earth)* 100 (B7), 12,431–12,447.
- Mosegaard, K., Tarantola, A., 2002. 16 - probabilistic approach to inverse problems. In: Lee, W. H., Kanamori, H., Jennings, P. C., Kisslinger, C. (Eds.), *International Handbook of Earthquake and Engineering Seismology, Part A, Volume 81 of International Geophysics*. Academic Press, pp. 237–265.
- Moses, J. I., Line, M. R., Visscher, C., Richardson, M. R., Nettelmann, N., Fortney, J. J., Barman, T. S., Stevenson, K. B., Madhusudhan, N., 2013. Compositional Diversity in the Atmospheres of Hot Neptunes, with Application to GJ 436b. *ApJ* 777 (1), 34.
- Moses, J. I., Tremblin, P., Venot, O., Miguel, Y., 2022. Chemical variation with altitude and longitude on exo-Neptunes: Predictions for Ariel phase-curve observations. *Experimental Astronomy* 53 (2), 279–322.
- Mousis, O., Aguichine, A., Bouquet, A., Lunine, J. I., Danger, G., Mandt, K. E., Luspay-Kuti, A., 2021. Cold Traps of Hypervolatiles in the Protosolar Nebula at the Origin of the Peculiar Composition of Comet C/2016 R2 (PanSTARRS). *Planetary Science Journal* 2 (2), 72.
- Mousis, O., et al., 2018. Scientific rationale for Uranus and Neptune in situ explorations. *Planet. Space Sci.* 155, 12–40.
- Mousis, O., Deleuil, M., Aguichine, A., Marcq, E., Naar, J., Aguirre, L. A., Brugger, B., Gonçalves, T., 2020. Irradiated Ocean Planets Bridge Super-Earth and Sub-Neptune Populations. *The Astrophysical Journal Letters* 896 (2), L22.
- Mousis, O., Ronnet, T., Lunine, J. I., 2019. Jupiter's Formation in the Vicinity of the Amorphous Ice Snowline. *ApJ* 875 (1), 9.
- Mugnai, L. V., et al., 2021. ARES. V. No Evidence For Molecular Absorption in the HST WFC3 Spectrum of GJ 1132 b. *AJ* 161 (6), 284.
- Naar, J., 2019. Internal structure of ocean worlds: physical modelling of supercritical water. sciences sorbonne université. laboratoire d'astrophysique de marseille. Master thesis supervised by Magali Deleuil and Olivier Mousis.
- Nakajima, S., Hayashi, Y.-Y., Abe, Y., 1992. A study on the 'runaway greenhouse effect' with a one-dimensional radiative-convective equilibrium model. *Journal of Atmospheric Sciences* 49 (23), 2256–2266.
- Nettelmann, N., Fortney, J. J., Kramm, U., Redmer, R., 2011. Thermal Evolution and Structure Models of the Transiting Super-Earth GJ 1214b. *The Astrophysical Journal* 733 (1), 2.

- Nettelmann, N., Holst, B., Kietzmann, A., French, M., Redmer, R., Blaschke, D., 2008. Ab initio equation of state data for hydrogen, helium, and water and the internal structure of jupiter. *The Astrophysical Journal* 683 (2), 1217.
- Nettelmann, N., Valencia, D., 2021. Exoplanetary Interiors. In: Madhusudhan, N. (Ed.), *ExoFrontiers; Big Questions in Exoplanetary Science*. pp. 16–1.
- Noack, L., Höning, D., Rivoldini, A., Heistracher, C., Zimov, N., Journaux, B., Lammer, H., Van Hoolst, T., Bredehöft, J. H., 2016. Water-rich planets: How habitable is a water layer deeper than on Earth? *Icarus* 277, 215–236.
- Olson, S., Jansen, M. F., Abbot, D. S., Halevy, I., Goldblatt, C., 2022. The effect of ocean salinity on climate and its implications for earth's habitability. *Geophysical Research Letters* 49 (10), e2021GL095748. e2021GL095748 2021GL095748.
- Olson, S. L., Jansen, M., Abbot, D. S., 2020. Oceanographic considerations for exoplanet life detection. *The Astrophysical Journal* 895 (1), 19.
- Orell-Miquel, J., et al., 2022. A tentative detection of He I in the atmosphere of GJ 1214 b. *A&A* 659, A55.
- Ortenzi, G., Noack, L., Sohl, F., Guimond, C. M., Grenfell, J. L., Dorn, C., Schmidt, J. M., Vulpius, S., Katyal, N., Kitzmann, D., Rauer, H., 2020. Mantle redox state drives outgassing chemistry and atmospheric composition of rocky planets. *Scientific Reports* 10, 10907.
- Otegi, J. F., Dorn, C., Helled, R., Bouchy, F., Haldemann, J., Alibert, Y., 2020. Impact of the measured parameters of exoplanets on the inferred internal structure. *A&A* 640, A135.
- Owen, J. E., Jackson, A. P., 2012. Planetary evaporation by UV & X-ray radiation: basic hydrodynamics. *MNRAS* 425 (4), 2931–2947.
- Papaloizou, J. C. B., Szuszkiewicz, E., Terquem, C., 2018. The TRAPPIST-1 system: orbital evolution, tidal dissipation, formation and habitability. *MNRAS* 476 (4), 5032–5056.
- Parmentier, V., Guillot, T., 2014. A non-grey analytical model for irradiated atmospheres. I. Derivation. *A&A* 562, A133.
- Pepe, F., et al., 2021. ESPRESSO at VLT. On-sky performance and first results. *A&A* 645, A96.
- Perruchot, S., et al., 2008. The SOPHIE spectrograph: design and technical key-points for high throughput and high stability. In: McLean, I. S., Casali, M. M. (Eds.), *Ground-based and Airborne Instrumentation for Astronomy II*, Volume 7014 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 70140J.
- Perryman, M., 2018. *The Exoplanet Handbook* (2 ed.). Cambridge University Press.
- Persson, C. M., et al., 2022. TOI-2196 b: Rare planet in the hot Neptune desert transiting a G-type star. arXiv e-prints, arXiv:2208.05797.
- Petigura, E. A., et al., 2022. The California-Kepler Survey. X. The Radius Gap as a Function of Stellar Mass, Metallicity, and Age. *AJ* 163 (4), 179.

- Pierrehumbert, R. T., 2010. *Principles of Planetary Climate*.
- Piette, A. A. A., Madhusudhan, N., Mandell, A. M., 2022. HyDRo: atmospheric retrieval of rocky exoplanets in thermal emission. *MNRAS* 511 (2), 2565–2584.
- Plotnykov, M., Valencia, D., 2020. Chemical fingerprints of formation in rocky super-Earths' data. *MNRAS* 499 (1), 932–947.
- Pluriel, W., Marcq, E., Turbet, M., 2019. Modeling the albedo of earth-like magma ocean planets with h₂o-co₂ atmospheres. *Icarus* 317, 583–590.
- Polyansky, O. L., Kyuberis, A. A., Zobov, N. E., Tennyson, J., Yurchenko, S. N., Lodi, L., 2018. ExoMol molecular line lists XXX: a complete high-accuracy line list for water. *Monthly Notices of the Royal Astronomical Society* 480 (2), 2597–2608.
- Queloz, D., et al., 2009. The CoRoT-7 planetary system: two orbiting super-Earths. *A&A* 506 (1), 303–319.
- Ramos, X. S., Charalambous, C., Benítez-Llambay, P., Beaugé, C., 2017. Planetary migration and the origin of the 2:1 and 3:2 (near)-resonant population of close-in exoplanets. *A&A* 602, A101.
- Rauer, H., et al., 2014. The PLATO 2.0 mission. *Experimental Astronomy* 38 (1-2), 249–330.
- Raymond, S. N., Boulet, T., Izidoro, A., Esteves, L., Bitsch, B., 2018. Migration-driven diversity of super-Earth compositions. *MNRAS* 479 (1), L81–L85.
- Richard, C., et al., 2012. New section of the hitran database: Collision-induced absorption (cia). *Journal of Quantitative Spectroscopy and Radiative Transfer* 113 (11), 1276–1285. Three Leaders in Spectroscopy.
- Ricker, G. R., et al., 2015. Transiting Exoplanet Survey Satellite (TESS). *Journal of Astronomical Telescopes, Instruments, and Systems* 1, 014003.
- Rivera, E. J., Lissauer, J. J., Butler, R. P., Marcy, G. W., Vogt, S. S., Fischer, D. A., Brown, T. M., Laughlin, G., Henry, G. W., 2005. A $\sim 7.5 M_{\oplus}$ Planet Orbiting the Nearby Star, GJ 876. *ApJ* 634 (1), 625–640.
- Rivkin, A. S., Howell, E. S., Vilas, F., Lebofsky, L. A., 2002. *Hydrated Minerals on Asteroids: The Astronomical Record*, pp. 235–253.
- Robinson, T. D., Catling, D. C., 2012. An Analytic Radiative-Convective Model for Planetary Atmospheres. *ApJ* 757 (1), 104.
- Rodríguez Martínez, R., et al., 2022. A Reanalysis of the Composition of K2-106b: an Ultra-short Period Super-Mercury Candidate. arXiv e-prints, arXiv:2208.07883.
- Rogers, J. G., Gupta, A., Owen, J. E., Schlichting, H. E., 2021. Photoevaporation versus core-powered mass-loss: model comparison with the 3D radius gap. *MNRAS* 508 (4), 5886–5902.
- Ronnet, T., Mousis, O., Vernazza, P., 2017. Pebble Accretion at the Origin of Water in Europa. *ApJ* 845 (2), 92.

- Rothman, L. S., Gordon, I. E., Barber, R. J., Dothe, H., Gamache, R. R., Goldman, A., Perevalov, V. I., Tashkun, S. A., Tennyson, J., 2010. HITEMP, the high-temperature molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer* 111, 2139–2150.
- Sanchez-Lavega, A., 2011. An Introduction to Planetary Atmospheres. *Contemporary Physics* 52 (5), 487–487.
- Sanz-Forcada, J., Micela, G., Ribas, I., Pollock, A. M. T., Eiroa, C., Velasco, A., Solano, E., García-Álvarez, D., 2011. Estimation of the xuv radiation onto close planets and their evaporation. *A&A* 532, A6.
- Schmidt, H., Seitz, S., Hassel, E., Wolf, H., 2018. The density-salinity relation of standard seawater. *Ocean Science* 14 (1), 15–40.
- Schulze, J. G., Wang, J., Johnson, J. A., Gaudi, B. S., Unterborn, C. T., Panero, W. R., 2021. On the Probability That a Rocky Planet’s Composition Reflects Its Host Star. *Planetary Science Journal* 2 (3), 113.
- Scora, J., Valencia, D., Morbidelli, A., Jacobson, S., 2020. Chemical diversity of super-Earths as a consequence of formation. *MNRAS* 493 (4), 4910–4924.
- Seager, S., Kuchner, M., Hier-Majumder, C. A., Militzer, B., 2007. Mass-Radius Relationships for Solid Exoplanets. *ApJ* 669 (2), 1279–1297.
- Seager, S., Sasselov, D. D., 2000. Theoretical Transmission Spectra during Extrasolar Giant Planet Transits. *ApJ* 537 (2), 916–921.
- Simonelli, D. P., Veverka, J., 1988. Bolometric albedos and diurnal temperatures of the brightest regions on io. *Icarus* 74 (2), 240–261.
- Sing, D. K., et al., 2016. A continuum from clear to cloudy hot-Jupiter exoplanets without primordial water depletion. *Nature* 529 (7584), 59–62.
- Sinukoff, E., et al., 2017. K2-66b and k2-106b: Two extremely hot sub-neptune-size planets with high densities. *The Astronomical Journal* 153 (6), 271.
- Sneep, M., Ubachs, W., 2005. Direct measurement of the Rayleigh scattering cross section in various gases. *Journal of Quantitative Spectroscopy and Radiative Transfer* 92 (3), 293–310.
- Snellen, I. A. G., de Kok, R. J., le Poole, R., Brogi, M., Birkby, J., 2013. Finding Extraterrestrial Life Using Ground-based High-dispersion Spectroscopy. *ApJ* 764 (2), 182.
- Sotin, C., Grasset, O., Mocquet, A., 2007. Mass–radius curve for extrasolar earth-like planets and ocean planets. *Icarus* 191 (1), 337–351.
- Spaargaren, R. J., Ballmer, M. D., Bower, D. J., Dorn, C., Tackley, P. J., 2020. The influence of bulk composition on the long-term interior-atmosphere evolution of terrestrial exoplanets. *A&A* 643, A44.
- Stamnes, K., Tsay, S. C., Jayaweera, K., Wiscombe, W., Stamnes, S., Jin, Z., Lin, Z., 2017. DISORT: DIScrete Ordinate Radiative Transfer.

- Stevenson, D. J., Lunine, J. I., 1988. Rapid formation of Jupiter by diffusive redistribution of water vapor in the solar nebula. *Icarus* 75 (1), 146–155.
- Stevenson, K. B., et al., 2014. Thermal structure of an exoplanet atmosphere from phase-resolved emission spectroscopy. *Science* 346 (6211), 838–841.
- Stolker, Quanz, S. P., Todorov, K. O., Kühn, J., Mollière, P., Meyer, M. R., Currie, T., Daemgen, S., Lavie, B., 2020. Miracles: atmospheric characterization of directly imaged planets and substellar companions at 4-5 - i. photometric analysis of b, hip 65426 b, pz tel b, and hd 206893 b. *A&A* 635, A182.
- Terquem, C., Papaloizou, J. C. B., 2007. Migration and the Formation of Systems of Hot Super-Earths and Neptunes. *ApJ* 654 (2), 1110–1120.
- The LUVOIR Team, 2019. The LUVOIR Mission Concept Study Final Report. arXiv e-prints, arXiv:1912.06219.
- Thiabaud, A., Marboeuf, U., Alibert, Y., Leya, I., Mezger, K., 2015. Elemental ratios in stars vs planets. *A&A* 580, A30.
- Thomas, S. W., Madhusudhan, N., 2016. In hot water: effects of temperature-dependent interiors on the radii of water-rich super-Earths. *Monthly Notices of the Royal Astronomical Society* 458 (2), 1330–1344.
- Thompson, M. A., Telus, M., Schaefer, L., Fortney, J. J., Joshi, T., Lederman, D., 2021. Composition of terrestrial exoplanet atmospheres from meteorite outgassing experiments. *Nature Astronomy* 5, 575–585.
- Tinetti, G., et al., 2018. A chemical survey of exoplanets with ARIEL. *Experimental Astronomy* 46 (1), 135–209.
- Tolento, J. P., Robinson, T. D., 2019. A simple model for radiative and convective fluxes in planetary atmospheres. *Icarus* 329, 34–45.
- Tran, H., Turbet, M., Chelin, P., Landsheere, X., 2018. Measurements and modeling of absorption by CO₂+ H₂O mixtures in the spectral region beyond the CO₂ ν₃-band head. *Icarus* 306, 116–121.
- Truong, N., Lunine, J. I., 2021. Volcanically extruded phosphides as an abiotic source of Venusian phosphine. *Proceedings of the National Academy of Science* 118 (29), e2021689118.
- Tsai, S.-M., Innes, H., Lichtenberg, T., Taylor, J., Malik, M., Chubb, K., Pierrehumbert, R., 2021. Inferring Shallow Surfaces on Sub-Neptune Exoplanets with JWST. *ApJ* 922 (2), L27.
- Tsiaras, A., Rocchetto, M., Waldmann, I. P., Venot, O., Varley, R., Morello, G., Damiano, M., Tinetti, G., Barton, E. J., Yurchenko, S. N., Tennyson, J., 2016. DETECTION OF AN ATMOSPHERE AROUND THE SUPER-EARTH 55 CANCRI e. *The Astrophysical Journal* 820 (2), 99.
- Tsiaras, A., Waldmann, I. P., Tinetti, G., Tennyson, J., Yurchenko, S. N., 2019. Water vapour in the atmosphere of the habitable-zone eight-Earth-mass planet K2-18 b. *Nature Astronomy* 3, 1086–1091.

- Turbet, M., Bolmont, E., Ehrenreich, D., Gratier, P., Leconte, J., Selsis, F., Hara, N., Lovis, C., 2020. Revised mass-radius relationships for water-rich rocky planets more irradiated than the runaway greenhouse limit. *A&A* 638, A41.
- Turbet, M., Ehrenreich, David, Lovis, Christophe, Bolmont, Emeline, Fauchez, Thomas, 2019. The runaway greenhouse radius inflation effect - an observational diagnostic to probe water on earth-sized planets and test the habitable zone concept. *A&A* 628, A12.
- Turbet, M., Tran, H., Hartmann, J. M., Forget, F., 2017. Toward a more Accurate Spectroscopy of CO₂/H₂O-Rich Atmospheres: Implications for the Early Martian Atmosphere. In: Fourth International Conference on Early Mars: Geologic, Hydrologic, and Climatic Evolution and the Implications for Life, Volume 2014, pp. 3063.
- Udry, S., et al., 2014. Exoplanet Science with the European Extremely Large Telescope. The Case for Visible and Near-IR Spectroscopy at High Resolution. arXiv e-prints, arXiv:1412.1048.
- Unterborn, C. T., Desch, S. J., Hinkel, N. R., Lorenzo, A., 2018. Inward migration of the TRAPPIST-1 planets as inferred from their water-rich compositions. *Nature Astronomy* 2, 297–302.
- Unterborn, C. T., Dismukes, E. E., Panero, W. R., 2016. Scaling the Earth: A Sensitivity Analysis of Terrestrial Exoplanetary Interior Models. *ApJ* 819 (1), 32.
- Valencia, D., Guillot, T., Parmentier, V., Freedman, R. S., 2013. Bulk Composition of GJ 1214b and Other Sub-Neptune Exoplanets. *ApJ* 775 (1), 10.
- Valencia, D., Ikoma, M., Guillot, T., Nettelmann, N., 2010. Composition and fate of short-period super-Earths. The case of CoRoT-7b. *A&A* 516, A20.
- Valencia, D., O'Connell, R. J., Sasselov, D., 2006. Internal structure of massive terrestrial planets. *Icarus* 181 (2), 545–554.
- Valencia, D., Sasselov, D. D., O'Connell, R. J., 2007. Detailed Models of Super-Earths: How Well Can We Infer Bulk Properties? *ApJ* 665 (2), 1413–1420.
- Vazan, A., Sari, R., Kessel, R., 2022. A New Perspective on the Interiors of Ice-rich Planets: Ice-Rock Mixture Instead of Ice on Top of Rock. *ApJ* 926 (2), 150.
- Vernazza, P., et al., 2015. Interplanetary Dust Particles as Samples of Icy Asteroids. *ApJ* 806 (2), 204.
- Vinet, P., Rose, J. H., Ferrante, J., Smith, J. R., 1989. Universal features of the equation of state of solids. *Journal of Physics: Condensed Matter* 1 (11), 1941–1963.
- Vogt, S. S., et al., 1994. HIRES: the high-resolution echelle spectrometer on the Keck 10-m Telescope. In: Crawford, D. L., Craine, E. R. (Eds.), *Instrumentation in Astronomy VIII, Volume 2198 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 362.
- Wagner, W., Pruß, A., 2002. The iapws formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use. *Journal of Physical and Chemical Reference Data* 31 (2), 387–535.
- Wang, H. S., Lineweaver, C. H., Quanz, S. P., Mojzsis, S. J., Ireland, T. R., Sossi, P. A., Seidler, F., Morel, T., 2022. A Model Earth-sized Planet in the Habitable Zone of α Centauri A/B. *ApJ* 927 (2), 134.

- Wang, H. S., Liu, F., Ireland, T. R., Brasser, R., Yong, D., Lineweaver, C. H., 2018. Enhanced constraints on the interior composition and structure of terrestrial exoplanets. *Monthly Notices of the Royal Astronomical Society* 482 (2), 2222–2233.
- Wang, H. S., Quanz, S. P., Yong, D., Liu, F., Seidler, F., Acuña, L., Mojzsis, S. J., 2022. Detailed chemical compositions of planet-hosting stars – II. Exploration of the interiors of terrestrial-type exoplanets. *Monthly Notices of the Royal Astronomical Society* 513 (4), 5829–5846.
- Wang, S., Wu, D.-H., Barclay, T., Laughlin, G. P., 2017. Updated Masses for the TRAPPIST-1 Planets. arXiv e-prints, arXiv:1704.04290.
- Weiss, L. M., Marcy, G. W., 2014. The Mass-Radius Relation for 65 Exoplanets Smaller than 4 Earth Radii. *ApJ* 783 (1), L6.
- Werner, M. W., et al., 2004. The Spitzer Space Telescope Mission. *ApJS* 154 (1), 1–9.
- Yang, J., Ji, W., Zeng, Y., 2020. Transition from eyeball to snowball driven by sea-ice drift on tidally locked terrestrial planets. *Nature Astronomy* 4, 58–66.
- Yu, X., Moses, J. I., Fortney, J. J., Zhang, X., 2021. How to Identify Exoplanet Surfaces Using Atmospheric Trace Species in Hydrogen-dominated Atmospheres. *ApJ* 914 (1), 38.
- Zeng, L., et al., 2019. Growth model interpretation of planet size distribution. *Proceedings of the National Academy of Sciences* 116 (20), 9723–9728.
- Zeng, L., Sasselov, D., 2014. The Effect of Temperature Evolution on the Interior Structure of H₂O-rich Planets. *The Astrophysical Journal* 784 (2), 96.
- Zhang, M., Knutson, H. A., Wang, L., Dai, F., Oklopčic, A., Hu, R., 2021. No escaping helium from 55 cnc e. *The Astronomical Journal* 161 (4), 181.
- Zhang, Z., Zhou, Y., Rackham, B. V., Apai, D., 2018. The Near-infrared Transmission Spectra of TRAPPIST-1 Planets b, c, d, e, f, and g and Stellar Contamination in Multi-epoch Transit Spectra. *AJ* 156 (4), 178.
- Zieba, S., et al., 2022. K2 and Spitzer phase curves of the rocky ultra-short-period planet K2-141 b hint at a tenuous rock vapor atmosphere. arXiv e-prints, arXiv:2203.00370.