Water-rich sub-Neptunes and rocky super Earths around different Stars: Radii shaped by Volatile Partitioning, Formation, and Evolution

Remo Burn[1], Komal Bali[2], Caroline Dorn[2], Rafael Luque[3, 4], and Simon L. Grimm[5, 2]

[1] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany
    e-mail: burn@mpia.de
[2] Institute for Particle Physics and Astrophysics, ETH Zürich, HPK E32, Otto-Stern-Weg 5, 8093 Zürich,Switzerland
[3] Department of Astronomy & Astrophysics, University of Chicago, Chicago, IL 60637, USA
[4] NHFP Sagan Fellow
[5] University of Zürich, Department of Astrophysics, Winterthurerstrasse 190, 8057 Zürich, Switzerland

November 27, 2024

ABSTRACT

Context. Despite precise characterization measurements, the nature of planets with radii ranging from 2 to 4 Earth radii, the sub-Neptunes, remains unknown due to degeneracies in interior models. However, a statistical ensemble of small planets with measured masses and radii around different stars has been compiled by the field.
Aims. It can be used to test the prediction of large water reservoirs on sub-Neptunes by planet formation theory with orbital migration. We want to find out whether this water reservoir is included in photoevaporative winds and how much of it can partition into the rocky and metallic interior.
Methods. We couple the result of a planetary formation model with planetesimal and gas accretion as well as orbital migration to evolution models which assume perfect mixing of water with H/He in the envelope or complete segregation. For the mixed envelopes, we also include an analytic treatment of fractionation during photoevaporative mass-loss. Further, the effect of equilibrium dissolution of water into an assumed magma ocean and into the metallic core is studied for the first time in coupled formation-evolution models.
Results. Out of the tested scenarios, the mass-radius relation of exoplanets is best matched under the mixed assumption without water sequestration to the interior. We quantify the radius valley location and scaling with stellar mass. Fractionation is not found to significantly alter the composition of the planets for our initial conditions due to initially massive envelopes on all planets. In contrast, water sequestration has a profound effect on the radius evolution and compositional budget of the planets. The model predicts the preservation of large quantities of water even if the gaseous envelope is lost. Planets with corresponding bulk densities are not observed in comparably large numbers.
Conclusions. By combining formation and evolution model, we probe a parameter space favored by core accretion theory. We conclude that observationally searching for the imprint of fractionation on present-day exoplanet atmospheres can constrain initial gas mass fractions of exoplanets but not the occurrence of photoevaporative mass loss in general. The dissolution of different volatiles into the planetary interior and solidification of the magma ocean are natural next steps for comprehensive treatment of atmosphere-interior interaction.

Key words. Planets and satellites: formation – Planets and satellites: physical evolution – Planets and satellites: interiors – Planets and satellites: atmospheres – Planet-star interactions

1. Introduction

So far, the mass, orbital period, and radius of around two hundred sub-Saturn-mass exoplanets have been characterized with high precision (Parc et al. 2024). This success was possible thanks to various observational programs, such as K2 (Howell et al. 2014) or TESS (Ricker et al. 2015), which find nearby planets amenable to characterization via photometric and radial-velocity follow-up observations (Astudillo-Defru et al. 2020; Bonomo et al. 2023; Bonfanti et al. 2024; Brinkman et al. 2023; Damasso et al. 2023; Cloutier et al. 2024; Giacalone et al. 2022; Kossakowski et al. 2021; Leleu et al. 2024b; Luque & Pallé 2022; Luque et al. 2022, 2023; Orell-Miquel et al. 2023; Osborn et al. 2021; Passegger et al. 2024; Ulmer-Moll et al. 2023). The set of exoplanet data complements the statistical study of exoplanet occurrence from Kepler (Borucki et al. 2010; Lissauer

et al. 2023) where the masses of the planets are often unconstrained.

But even with a precisely determined mass, the interior structure of the planet is not fully characterized. This is the case for rocky planets due to an a priori unknown metallic core to rocky mantle ratio (Dorn et al. 2015) and even more so if volatile species such as water, hydrogen (H), and helium (He) are present. Without additional constraints from atmospheric spectroscopy, this degeneracy is impossible to break for an individual object. However, the statistical study of the exoplanet demographics can reveal further insights into the typical structure of an average exoplanet. One key feature which was discovered is the so-called radius valley which separates the smaller super Earths from the larger sub-Neptunes (Fulton et al. 2017; Fulton & Petigura 2018; Van Eylen et al. 2018). For planets with gaseous envelopes which are continuously stripped by photoe-

vaporation, driven by high-energy radiation from the star, such a transition was predicted to exist (Owen & Wu 2013; Lopez & Fortney 2013; Jin et al. 2014; Chen & Rogers 2016) and offered a compelling explanation. With the increasing knowledge of the mass and thus the bulk density of these planets, Luque & Pallé (2022) found that – at least for M-dwarf planets – the radius valley as seen by Kepler is fundamentally the projection of a gap in planetary bulk density, as expected from a compositional transition.

An open question remains: what is the volatile material present on sub-Neptunes but absent on super Earths? Traditionally, it was thought to be pure H/He and large quantities of water ice were excluded (Lopez 2017; Jin & Mordasini 2018; Rogers & Owen 2021). However, water accreted as ice likely changes its phase after accretion and could be present in large quantities as steam (partially) mixed with H/He (Dorn et al. 2018; Jin & Mordasini 2018; Zeng et al. 2019; Mousis et al. 2020; Venturini et al. 2020; Turbet et al. 2020; Acuña et al. 2021; Aguichine et al. 2021; Pierrehumbert 2023). Here, we will explore scenarios where water mixes with H/He, as well as a case with two distinct, stratified layers.

If water accounts for a significant portion (on the order of tens of percent) of the total mass of the planet, it likely accreted it in solid form beyond the water iceline in the protoplanetary disk. Therefore, the presence or absence of water on sub-Neptunes is a fundamental probe for large scale migration of these planets (Lin et al. 1996; Ward 1997; Paardekooper et al. 2023) and provides essential constraints for modern planet formation modeling. The orbits of sub-Neptune pairs are commonly in mean-motion resonance (Leleu et al. 2024a; Chen et al. 2024), which is likely linked to migration. However, it can be reproduced by small-scale convergent migration. Verifying the migration hypothesis through compositional analysis would prove that migration of sub-Neptunes over significant distances occurred.

Indeed, the limited masses reached by purely rocky planets forming via core accretion strongly favors that the origin of sub-Neptune mass planets lies at larger distances (Alibert et al. 2005; Emsenhuber et al. 2023b). Therefore, a water-rich composition is the focus of studies coupling formation and evolution models to explore the radius valley (Venturini et al. 2020; Chakrabarty & Mulders 2024; Burn et al. 2024).

Additionally, early evidence from JWST transmission spectroscopy suggests that some of the observed sub-Neptune atmospheres are likely water-rich. So far, the presence of water was confirmed in TOI-270 d (Holmberg & Madhusudhan 2024; Benneke et al. 2024), GJ 3470 b (Beatty et al. 2024), and GJ 9827 d (Piaulet-Ghorayeb et al. 2024), the latter begin consistent with a water-rich "steam world". Tentative detections have been obtained for other sub-Neptunes as well (MacDougall et al. 2023; Kempton et al. 2023; Wallack et al. 2024). Further observations are scheduled for this population, which have the potential to quantify the atmospheric water abundance (Kempton & Knutson 2024; Nixon et al. 2024). Combined with data on the bulk properties, these measurements help to reveal the internal composition of sub-Neptunes.

If the atmosphere consists of more than one species, it is possible that atmospheric escape preferentially leads to loss of the lighter species relative to the heavier species. If hydrogen is lost in larger abundance than the heavier species, this would change the metallicity in the envelope over time and could lead to a metal-rich late atmosphere. This process, called *fractionation*, was first studied analytically (Hunten et al. 1987; Zahnle & Kasting 1986; Zahnle et al. 1990) and is a common subject in the context of Solar System planets' atmospheres (Chasseфère

1996; Odert et al. 2018; Lammer et al. 2020). For exoplanets, the process was investigated for the fractionation of He from H (Hu et al. 2015; Malsky & Rogers 2020; Cherubim et al. 2024) and found large He fractions in some cases. Gu & Chen (2023) reported significant deuterium fractionation on sub-Neptunes occurring in a part of the parameter space. So far, no studies address the fractionation of heavier elements in sub-Neptune atmospheres. In this work, we investigate using a simple analytic approach the fractionation of oxygen (from assumed dissociated, abundant water molecules) from H, which is a required addition for water-rich sub-Neptunes.

In addition to the exact atmospheric partitioning, there is another large source of uncertainty on the water partitioning in the planet (see Ikoma et al. 2018; Lichtenberg et al. 2023, for recent reviews). Instead of mixing with the atmosphere, water can be sequestered in the deep interior (mantle and core). While solidified rocky mantle material can only hold limited amounts of volatiles (Shah et al. 2021), large quantities of volatiles can be dissolved in magma (Hirschmann et al. 2012; Dorn & Lichtenberg 2021; Sossi et al. 2023). Moreover, during core formation, volatiles can also get sequestered in the iron core (Luo et al. 2024). At temperatures above 2000 K, supercritical rock and water can mix (Vazan et al. 2022). The presence of volatiles in the deeper interior decreases their amount in the envelope. In consequence, the total radius changes which will have interesting effects on the evolution of the planet, especially on photoevaporative mass loss which is sensitive to the planetary radius. Evolution models coupled to the interior were developed for the case of (mostly) hydrogen dissolution into the magma ocean by Chachan & Stevenson (2018); Kite et al. (2019, 2020); Lichtenberg (2021); Schlichting & Young (2022) who found significant impacts on the resulting radii. In particular, Kite et al. (2019) found a truncation of the radius distribution at the upper end of the sub-Neptune radii (the radius cliff) when considering the non-ideality of the hydrogen in contact with the magma which leads to more mass being dissolved. Recently, Rogers et al. (2024) developed an evolution model for super Earths including the effect of water dissolution into magma (Kite & Schaefer 2021, see also) and iron core following Luo et al. (2024). They use it to infer allowed water mass fractions for super Earths and generally treat the initial conditions as free parameters. Here, complementary, we present a similar model but use initial conditions from formation modeling instead to predict the effect of water sequestration on the population of simulated planets.

Another aspect of exoplanet demographics is the dependency on stellar properties, in particular the stellar mass. Several studies have focused on stellar mass dependency of the radius valley (Berger et al. 2020; Cloutier & Menou 2020; Van Eylen et al. 2021; Petigura et al. 2022; Luque & Pallé 2022; Ho et al. 2024; Bonfanti et al. 2024; Parc et al. 2024) and agree in inferring the locus of the valley at lower radii around smaller stars as predicted by theoretical works (Gupta & Schlichting 2020; Rogers et al. 2023). This trend was tested including water-rich atmospheres in Venturini et al. (2024) using the prediction of a formation model where pebble accretion is the mode of solid mass delivery. They confirm a persistent density valley and report a pollution of the radius valley due to lower-mass water worlds around low-mass stars. Previously, scatter in these relations was often linked to collisions (Inamdar & Schlichting 2015; Izidoro et al. 2022). Venturini et al. (2024) include the effect of collisions but rely on estimates using single-planet formation modeling. Here, we use results from full N-body interactions between 50 growing planets per disk. However, we use a simple treat-

ment of collision outcomes, slightly modified compared to our preceding study focusing on Solar-mass stars (Burn et al. 2024).

In this paper, we include for the first time in coupled formation and evolution models 1) the effects of fractionation of oxygen from hydrogen during mass loss, and 2) water sequestration in the interior. We proceed by summarizing previously used methods and introducing in detail the new aspects in Sect. 2 before presenting our results in different dimensions in Sect 3. The results motivate some discussion (Sect. 4) on the best way to compare distributions in mass-radius space with observations (Sect. 4.1), the unexpectedly weak impact of fractionation (Sect. 4.2), and a possible dissimilarity between observational data and the model results including water sequestration (Sect. 4.3). Finally, we will summarize and conclude our findings in Sect. 5.

2. Methods

2.1. Formation

The (exo-)planetary mass and radius distribution are shaped by both the embedded stage in the protoplanetary disk as well as the long-term evolution of each planet. For the former, which we call here the *formation stage*, we use the results of the population synthesis work by Burn et al. (2021). In that work, for five different stellar masses (0.1, 0.3, 0.5, 0.7, and 1.0 $M_{\odot}$), an initial protoplanetary disk with 50 embedded, moon-mass embryos was assumed. Since the protoplanetary disk properties are expected to depend on the stellar mass, the initial disk masses following the observations of Class I disks by Tychoniec et al. (2018) were scaled linearly with the stellar mass. Similarly, the disk radius decreases for lighter disks (Andrews et al. 2010, 2018, $\propto M_{\mathrm{gas}}^{0.625}$) while the disk lifetime, solid mass fraction, and inner edge orbital period were chosen to not depend on the stellar mass. In Emsenhuber et al. (2023a), it was confirmed to a large degree that such initial conditions can lead to the observed, evolved Class II disk properties. The approach to randomize those initial conditions allows for statistical comparison of the resulting planets with observations, which is the idea of population syntheses (Mordasini & Burn 2024; Burn & Mordasini 2024). The disk subsequently evolves, which is modeled in Burn et al. (2021) assuming a turbulent viscosity parameter α of 3×10^{-3} and a simple photoevaporation prescription tuned such that the disk lifetimes are distributed around 3 Myr for all modeled stars [1]

The formation of planets as modeled by Burn et al. (2021) uses the model described in detail in Emsenhuber et al. (2021) which is based on the works of Alibert et al. (2005); Mordasini et al. (2009, 2012b). It includes the N-body forces between the embryos (using mercury, Chambers 1999) and an analytically described accretion of planetesimals by a protoplanet with a gaseous envelope (Inaba et al. 2001; Inaba & Ikoma 2003; Fortier et al. 2013), which dominate the first modeled stage of growth. At intermediate masses, the accretion of gas becomes important which is calculated by solving the one-dimensional internal structure equations (Bodenheimer & Pollack 1986) of the gaseous envelope and tracking its energy content which allows the description of its cooling and contraction. As long as the planet's growth is not limited by the mass of the surrounding gaseous disk, the amount of gas accreted is given by the contraction of the envelope. When the planet mass becomes dominated

by gas instead of solids, the larger compressibility of gas allows for increased contraction which leads to even more gas accretion. Thus, the planet enters a runaway accretion stage (Makino et al. 1998). This rapid gas accretion phase is stopped once the disk is not able to supply by viscous transport the amount of gas that could be accreted by the planet given its cooling and contraction rate. At this point, the planet detaches from the gaseous disk, contracts, and its accretion rate is modeled using an analytical estimate for the maximum gas accretion rate (Emsenhuber et al. 2021).

While the disk is present, it will be perturbed by the young planet and exert a torque on it. Therefore, the planets orbital elements are changed and in our model follows the prescriptions of Paardekooper et al. (2011) for type I migration and Dittkrist et al. (2014) for type II migration. Migration will have the effect of typically moving volatile-rich planets formed beyond the snow line toward the star. Therefore, a common outcome to models including planetary migration is the prediction of planets with volatile (mostly water) contents of several tens of percent in bulk mass at short orbital periods (Ida & Lin 2008; Alibert et al. 2005). This effect is even more pronounced for lower stellar masses due to the lack of planets reaching the slower type II regime (Alibert & Benz 2017; Burn et al. 2021). Nevertheless, in the simulations of Burn et al. (2021), which form the basis of our calculations, rocky planets emerging interior to the water iceline can also form due to an assumed high efficiency of planetesimal formation in this region. This is in contrast to other simulations assuming reduced efficiency of accretion of dry particles or different initial planetesimal distributions (Coleman et al. 2019; Miguel et al. 2020).

2.2. Evolution

After the planets have accreted solids and gas and might have migrated, the protoplanetary disk eventually dissipates. Starting from this point in time, defined by the local disk pressure around the planet dropping to low levels, we model the evolution of the planet as individual body. In this work, we modified the approach of Emsenhuber et al. (2021) in a few key aspects similar to Burn et al. (2024) in order to better model planets in the sub-Neptune regime with large volatile contents. The main aspect is the used equation of state (EOS) for volatiles for which we use an EOS for water (Haldemann et al. 2020) which covers all phases. For simplicity, we treat for the evolution phase all volatile species accreted as ice (i.e. H_2O, CO, CO_2, CH_3OH, CH_4, NH_3, N_2, and H_2S, see Marboeuf et al. 2014a,b) as water ice and will use "water" and "volatiles" as synonyms in this work.

We focus on four different evolution models which differ in their assumed underlying partitioning of water into the interior structure. The approaches are summarized in Figure 1 and are using: (1) the *Mixed* assumption, that is, the nominal model, where all volatiles treated as water are perfectly mixing with H and He, (2) a similar model but where the mixing assumption is relaxed for the lost mass due to photoevaporation, that is, *Fractionation* is allowed to occur, (3) the *Layered* model which does not mix the volatiles accreted as ice with other constituents, and finally (4) a model exploring the effect of dissolution of water to the magma ocean and metallic core, that is, the *Water Sequestration* model.

---

[1] While there is a trend of shorter disk lifetimes around stars more massive than 1.5 $M_{\odot}$ compared to those lighter than this threshold mass, this trend is not confirmed to exist below 1 $M_{\odot}$ (Richert et al. 2018, see also the discussion in Burn et al. 2021 but note the exception of some long-lived disks Silverberg et al. 2020).

2.2.1. Planetary structure and composition

General In all cases, we start at the disk dissipation time t_{disk} with initial conditions from the formation part of the models (Sect. 2.1). To improve the consistency, we recover the luminosities of the planets at t_{disk} analytically and rerun the evolution part from this time onward. A difference to Burn et al. (2024) is that we also use the hydrogen and helium (H/He) mass $M_{H/He}$ from t_{disk} as starting condition instead of the one after 20 Myr of evolution. This has an effect on simulated planets for which an impact occurred after t_{disk}. By starting with $M_{H/He}(t_{disk})$, we assume that no H/He is lost for the more massive planet in a collision while all H/He is lost for the impacting planet. The mass of other components, that is iron, rocky materials, and volatile ices, is however assumed to perfectly merge. Future works will re-introduce a more realistic impact stripping with the possibility of mass loss of all species. Therefore, in the scenarios where we assume that volatile species mix with H/He (Mixed, Fractionation, see below) we obtain a total heavy element content (by mass) in the envelope of

$$Z_{homo} = \frac{M_{vol}}{M_{vol} + M_{H/He}} \tag{1}$$

smaller or equal to that in Burn et al. (2024).

Further common elements in all models are the presence of an iron core with a silicate $MgSiO_3$ (perovskite) mantle both materials described with a modified polytropic equation of state $\rho(P) = \rho_0 + cP^n$ of Seager et al. (2007).

In all model cases, the uppermost – third or fourth – layer is what we call the envelope and is numerically resolved in one dimension. Similar to the formation stage (described in Emsenhuber et al. 2021), the following internal structure equations (e.g. Kippenhahn et al. 2012) in hydrostatic equilibrium are solved:

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

$$\frac{dP(r)}{dr} = -\frac{Gm(r)\rho(r)}{r^2} \tag{3}$$

$$\frac{dT(r)}{dr} = \frac{T}{P}\frac{dP}{dr}\min(\nabla_{ad}, \nabla_{rad}), \tag{4}$$

where $m(r)$ is the mass within a sphere of radius r, $P(r)$ is the pressure as a function of r, $\rho(r)$ is the local density, and $T(r)$ the temperature. Depending on the local conditions, the energy transport mechanism changes. We use the limiting case of convection where the adiabatic gradient, calculated from the EOS,

$$\nabla_{ad} = \left(\frac{\partial \ln T}{\partial \ln P}\right)_S \tag{5}$$

is smaller than the radiative gradient. The latter depends on the local optical depth τ. If $\tau \lesssim 10$, we use the atmosphere solution of Guillot (2010) which considers an outgoing infrared flux and a well separated in-going flux in the visible wavelength regime in a in-parts gray atmosphere under the Eddington approximation (see also Jin et al. 2014). In case of large optical depths, radiative diffusion is a well justified approximation, thus

$$\nabla_{rad} = -\frac{3}{16\pi ac}\frac{\kappa L \rho}{r^2 T^3} \tag{6}$$

where c is the speed of light in vacuum, $a = 4\sigma/c$ with the Stefan-Boltzmann constant σ. In an intermediate regime $14.4 < \tau < 144$, we linearly interpolate between the two regimes to find ∇_{rad}.

Following Mordasini et al. (2012b), we assume that the luminosity L is constant throughout the structure which reduces the number of equations and is justified as only the upper, radiative regime depends directly on the luminosity. The system of equations is closed by the EOS (see below). A shooting method is used to solve the boundary value problem which arises due to core size, mass, and temperature being given as lower boundary while external pressure and density are upper boundaries to the envelope structure. A dependency on the stellar type enters directly into the upper boundary temperature of the planetary structures in the form of the irradiation temperature $T_{irr} = T_\star \sqrt{R_\star/a_P}$, where the stellar temperature $T_\star$ and radius $R_\star$ are taken from the tracks of Baraffe et al. (2015).

Assuming that a hydrostatic equilibrium is established on short timescales, the structure of the planet is recalculated after each step forward in time. From the obtained pressure and temperature profile, the corresponding change in energy content (gravitational and internal) can then be estimated. This estimate can be used as the luminosity for the next timestep and leads to a method which was found to result in a consistent treatment of a cooling and contracting planet (Mordasini et al. 2012b). In addition to the energy budget from solid and gas accretion, we consider radiogenic heating from the decay of ^{40}K, ^{232}Th, and ^{238}U (Mordasini et al. 2012a) as an additional energy source. We further add an empirically determined source of energy causing the observed bloating of close-in planets following Sarkis et al. (2021).

Mixed As for the solid part of the planetary structure, we require an equation of state for the envelope matter. For the mixed model, the envelope is assumed to consist of a perfect mixture of water, hydrogen, and helium. Therefore, the envelope metallicity is set to $Z_{env} = Z_{homo}$. While perfect mixing at all conditions is an extreme assumption, there are clear indications that H/He and water are well miscible at the high temperatures within the runaway greenhouse limit (Vazan et al. 2022; Pierrehumbert 2023), which is a safe assumption for the observed population of small transiting exoplanets. The equation of state for water follows Haldemann et al. (2020) while that for H/He follows Chabrier et al. (2019). To combine them, we assume that Amagat's law of additive volumes holds.

Layered In the Layered model, water and H/He are assumed to not mix at all to explore the other fringe case. In this case, the planetary structure follows a pure H/He envelope on top of the solid ice or rock layers. The ice layer density and thermal coefficients are modeled following Seager et al. (2007). This is the case as long as H/He is present. After it is lost, we switch for the volatiles to a pure water envelope following the AQUA equation of state of Haldemann et al. (2020). Because AQUA also includes gaseous and supercritical phases of water, this transition implies for hot planets that we first force the ice to be solid, motivated by the pressure of the H/He on top of it, and then transform it to a typically supercritical vapor envelope. In our static model, this transition occurs instantaneously. Especially at short orbital periods, this introduces an inconsistency as the water would evaporate somewhat earlier even if the assumption of non-mixing holds. However, this transition is relatively short and we explore this layered case as a useful extreme assumption which complements the mixing one. We note that for planets which do not contain volatile ices or for planets without any H/He after formation the Layered and Mixed models are equiv-
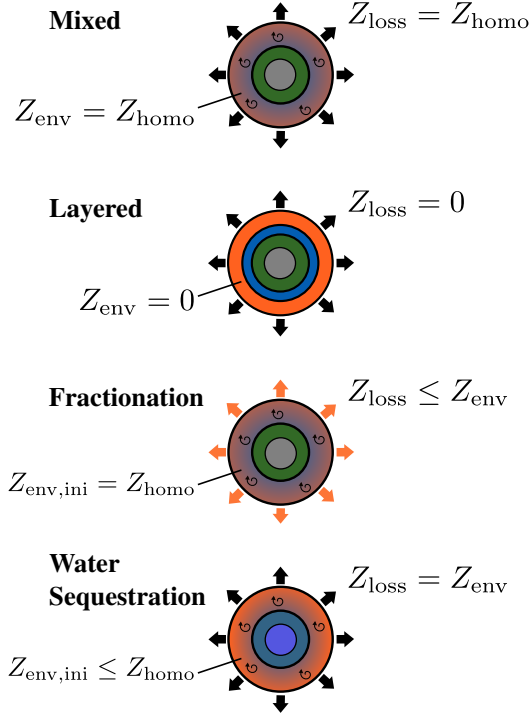
Mixed

$Z_{\text{loss}} = Z_{\text{homo}}$

$Z_{\text{env}} = Z_{\text{homo}}$

Layered

$Z_{\text{loss}} = 0$

$Z_{\text{env}} = 0$

Fractionation

$Z_{\text{loss}} \leq Z_{\text{env}}$

$Z_{\text{env,ini}} = Z_{\text{homo}}$

Water Sequestration

$Z_{\text{loss}} = Z_{\text{env}}$

$Z_{\text{env,ini}} \leq Z_{\text{homo}}$

Fig. 1. Schematic visualization of the four considered different structure models. Considered layers (inside-out) are a metallic core (gray), a silicate mantle (green) expected to be in molten form as a magma ocean, in the Layered model an ice layer, and at the top a gaseous envelope (orange). Its metallicity Z_{env} is in the Mixed and initially in the Fractionation case set to Z_{homo} – the metallicity resulting from mixing all volatiles with H/He uniformly in the gaseous envelope of the planet. For the Water Sequestration model (bottom), Z_{env} is lower since a part of the accreted volatiles are distributed to the magma ocean and metallic core (indicated by blue coloring of those two layers). The Layered model does not consider volatiles in the envelope. The arrows indicate mass loss and are colored in the Fractionation model to indicate varying metallicities in the lost gas (equal to or smaller than the envelope metallicity).

alent and that the assumption of the Layered model are the same as what is used during the early formation stage of the model.

Fractionation The structure and initialization of the Fractionation model is equal to the Mixed model explained above. The difference lies in loosening the assumption that the gas is perfectly mixed also in the photoevaporative wind. Instead, an analytic prescription (Hunten et al. 1987; Zahnle & Kasting 1986; Zahnle et al. 1990) detailed in Appendix A (optimistic estimate)[2] is used to determine the metallicity in the photoevaporative mass loss $Z_{\text{loss}} = \frac{\dot{M}_{\text{H}_2\text{O}}}{\dot{M}_{\text{H/He}} + \dot{M}_{\text{H}_2\text{O}}} \leq Z_{\text{env}}$. Consequently, this also implies that Z_{env} can evolve, that is, increase, over time.

Water Sequestration A novelty in evolutionary calculations is the here presented Water Sequestration model. Instead of mixing the accreted volatile species (treated as water) mass M_{vol} only within the envelope, we maximize the amount of water sequestered in the deep interior by assuming a molten magma ocean. In that case, some amount of volatiles $M_{\text{vol,magma}}$ can dis-
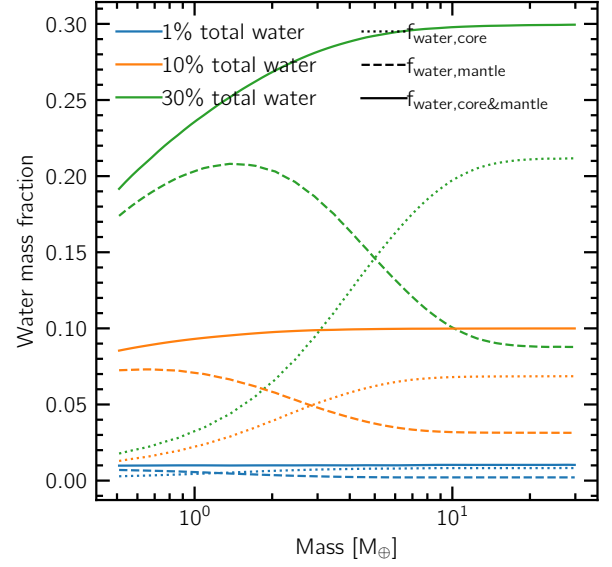
---

[2] We use the analytic estimates from Appendix A and the expressions for $\dot{M}_{\text{H}_2\text{O}}$ there instead of the output from the tables of Johnstone (2020) (Sect. 2.2.2) for its broader applicability.

Fig. 2. Water mass fractions in core and mantle relative to total mass of the planet. These values provide an upper limit and are used in the Water Sequestration model. The three different colors indicate cases with 1, 10, and 30 % of total planetary volatile mass content (including the gaseous envelope) M_{vol}/M while the different linestyles show the two interior reservoirs (dotted for core, dashed for mantle), respectively the total water mass fraction sequestered in the interior (solid).

solve into the magma and $M_{\text{vol,core}}$ can even dissolve into the metallic core based on the calculations in Luo et al. (2024) (see also Dorn & Lichtenberg 2021). The envelope volatile mass at the start of our long-term evolution phase is then only $M_{\text{vol,env}} = M_{\text{vol}} - M_{\text{vol,magma}} - M_{\text{vol,core}}$ and the envelope metallicity $Z_{\text{env,ini}} = M_{\text{vol,env}}/(M_{\text{vol,env}} + M_{\text{H/He}})$. The rationale of the model is to provide an upper limit for the mass of volatile species, treated as water, which could be removed from the envelope mass budget. This has implications for the evolution of the planetary envelope.

The partitioning of water between the mantle and surface water layers is modeled using a modified version of Henry's law which relates the water mass fraction in the magma $M_{\text{vol,magma}}/(M_{\text{vol,magma}} + M_{\text{Rock}})$ to the pressure at the bottom of the gaseous envelope P_{B} (see Dorn & Lichtenberg 2021 for details).

$$M_{\text{vol,magma}}/(M_{\text{vol,magma}} + M_{\text{Rock}}) = \alpha P_{\text{B}}^{1/\beta}. \quad (7)$$

The parameters $\alpha > 0$ and $\beta > 0$ were fitted to experimental data by Dorn & Lichtenberg (2021). Partitioning from the magma ocean to the metallic core (to determine $M_{\text{vol,core}}$) follows Luo et al. (2024) who assume iron silicate equilibration at mid-mantle pressures. Their partitioning coefficients depend on the concentration of water in the molten silicates as well as mid-mantle pressures but only marginally on the temperature at a given pressure.

For the coupling to the evolutionary cooling and photoevaporation model, tabulated results from these partitioning and solubility laws and further using the interior structure model detailed in Appendix B were used. We note that we did not include an explicit temperature dependency on the water partitioning leading to a weak effect of external temperature via the pressure dependencies. Therefore, the tables were using an exterior temperature of 3000 K for all models and varying total water mass budgets (from 1 % to 30 %) and planetary masses (0.5 to 30 $M_{\oplus}$)

were considered. For the equilibration of water between silicates and iron, we use mid-mantle pressures and the iron mass fraction of the dry core was assumed to be 0.244 consistent with the outcome of the equilibrium condensation model used in our formation model (Thiabaud et al. 2014). To construct the tables, the presence of H/He, which would increase pressures and thus solubilities, was excluded for simplicity but supported by the fact that most planets in the considered size regime contain little H/He in mass.

The resulting water mass fractions distributed at initialization to the molten mantle and the metallic core are shown in Fig. 2. We see a general trend of increasing water budget in the metallic core with increasing total mass. This is expected given the more siderophile nature of water with increasing pressure. The magma ocean water content peaks at intermediate masses. Combined, this results in a moderately increasing mass-dependency for the total fraction of water in the solid interior (solid lines).

During time evolution, the total mass of the planet and the atmospheric pressure can decrease due to photoevaporative mass loss. Therefore, the balance between sequestered water in the magma ocean and envelope (Eq. 7) is re-evaluated after each timestep and $M_{\mathrm{vol,magma}}$ is re-determined using the same tables as initially. The water partitioning between envelope and magma ocean is therefore dynamically adapted to the new conditions which leads to outgassing of water from the magma. In contrast, we assume that the volatiles trapped in the iron core remain constant over time (similar to outgassed scenarios in Rogers et al. (2024)).

For better comparison to the other structure models, the Water Sequestration model calculates the average density of the volatile-enriched metallic core and magma ocean using the equation of state of Seager et al. (2007) assuming the same separation into iron core, perovskite mantle, and water ice layer as in the Layered model. This implies that for the radius and bulk density of the interior, we do not assume that sequestration to the mantle and core occurred. Instead, the sequestered water (of mass $M_{\mathrm{vol,magma}} + M_{\mathrm{vol,core}}$) is treated as a high-pressure ice layer on top of the rocky material. A comparison to densities calculated consistently with the structure assumption of the Water Sequestration model (Fig. 1) is presented in Appendix B.

2.2.2. Photoevaporative mass loss

To model the escape of particles from the upper atmosphere due to high-energy radiation, we use a similar prescription for photoevaporation as in Affolter et al. (2023); Burn et al. (2024). The approach is equivalent between models where the envelope metallicity Z_{env} varies. We assume that the escape is in the thermally-driven hydrodynamic regime (see Gronoff et al. 2020, for a recent review). The total mass loss is a combination of the loss of H/He (Kubyshkina & Fossati 2021) and that of volatile species treated as water (Johnstone 2020). Given the Z_{env}, differing between models (see Sect. 2.2.1), it is given by

$$\dot{M}_{\mathrm{evap}} = Z_{\mathrm{env}} \dot{M}_{\mathrm{H_2O,J}} + (1 - Z_{\mathrm{env}}) \dot{M}_{\mathrm{H/He,Ku}} , \tag{8}$$

where we use the tabulated values of Kubyshkina & Fossati (2021) for $\dot{M}_{\mathrm{H/He,Ku}}$. The results are based on the chemical-hydrodynamic models of Kubyshkina et al. (2018), assume a 15% heating efficiency of the high-energy radiation, and depend on stellar mass, X-ray and extreme ultraviolet (EUV) fluxes F_{XEUV}, planetary masses and radii, as well as equilibrium temperature. To transition from these models applicable for pure H/He planets to water-dominated cases, the results of Johnstone

(2020) are used to determine $\dot{M}_{\mathrm{H_2O,J}}$ in Eq. (8) by adjusting the efficiency parameter ϵ in the energy-limited mass loss equation

$$\dot{M}_{\mathrm{esc,EL}} = \epsilon \frac{\pi F_{\mathrm{XEUV}} R_{\tau=2/3} R_{\mathrm{base}}^2}{G M_{\mathrm{tot}} K(\xi)} , \tag{9}$$

where F_{XEUV} is the received flux at the planet location in either X-ray (dominating early) or EUV wavelengths, $R_{\tau=2/3}$ is the radius of the layer where an optical depth of 2/3 is reached from the inside-out, and R_{base} is the base of the ionization layer. It is found by searching within the resolved planetary structure for the location where an optical depth of one if reached for UV photons (Murray-Clay et al. 2009). The factor $K(\xi) = 1 - \frac{3}{2\xi} - \frac{1}{2\xi^3}$, with $\xi = R_{\mathrm{Roche}}/R_{\tau=2/3}$ being the ratio of the planets Roche limit to its radius. As in Burn et al. (2024), the time evolution of F_{XEUV} follows (McDonald et al. 2019) and we linearly interpolate or extrapolate to lower stellar masses if required. In contrast to Burn et al. (2024), we reduced the rates of Johnstone (2020) to account for an additional cooling effect fitted following Yoshida et al. (2022) as $\dot{M}_{\mathrm{H_2O,J}} = \dot{M}_{\mathrm{H_2O,J,non-mod}} \times 2/(\log_{10}(Z_{\mathrm{env}} + 3))$, where $\dot{M}_{\mathrm{H_2O,J,non-mod}}$ is the unmodified rate fit to the results of Johnstone (2020).

2.3. Observational data and sampling

To compare simulations and observations, we use the collection of well characterized exoplanets from Parc et al. (2024) with the goal of comparing mass-radius relations. For our purposes, we restrict ourselves to planets within 30 days orbital periods and smaller than 7 $R_\oplus$. Furthermore we require precisely determined masses and radii with relative errors below 25% and 8% respectively. This results in a set of 201 observed planets. As a second exploration, we will relate the Solar-mass population of simulated exoplanet radii to the data from the California-Kepler survey (Fulton & Petigura 2018). For this second set of data, no precise masses are available but the application of the bias from Kepler can be done in a rigorous fashion using the KOBE package Mishra et al. (2021).

To allow for a tentative, quantitative comparison also for the sample from Parc et al. (2024), we apply an estimate for the observational bias and re-sample the synthetic population that is an outcome of the model described above such that the distribution of stellar masses and the right number of observed stars is reproduced. The observational data originates from heterogeneous sources and a human intervention bias is introduced when follow-up characterization observations are scheduled. Nevertheless, a significant fraction of the planets were discovered by TESS (Ricker et al. 2014) which is a transit survey. Therefore, we calculate the geometrical transit probability (Petigura et al. 2018)

$$p_{\mathrm{trans}} = 0.9 \frac{R_\star}{a} \tag{10}$$

and an estimate for the probability that a transiting planet is also detected (based on the Kepler mission, as detailed by Fulton et al. 2017; Petigura et al. 2018)

$$p_{\mathrm{det}} = \Phi\left(R_{\mathrm{trans}}, \log_{10}(\mu), \sigma\right) , \tag{11}$$

where Φ is the standard cumulative distribution function, $\mu = 1.387\,R_\oplus \times \sqrt{T_{\mathrm{Kepler}}/T_{\mathrm{TESS}}} \times (P/100\,\mathrm{days})^{0.19}$, and $\sigma = 0.145\,\mathrm{dex} \times 2$. Here, we use the expected scaling of signal to noise with observation time (Petigura et al. 2018), estimate the mission duration ratio of Kepler to TESS $T_{\mathrm{Kepler}}/T_{\mathrm{TESS}} = 13.33$,

and increase σ by a factor of two to account for the heterogeneous sample as well as the variability in TESS observation duration per start (depends on the sector). We visually verify the estimated detection probabilities by comparing them to the results of the detailed analysis of TESS planet detection yields by Kunimoto et al. (2022).

Most of the models (except the Layered one) rely on the assumption that water mixes with H/He. This is established for planets which are interior to the runaway greenhouse limit where water cannot condense due to too high temperatures. Therefore, we apply a cut at the runaway greenhouse limit which affects planets around different stars differently: The runaway greenhouse limit lies at close orbits similar to the detection thresholds for late M dwarfs only. For the estimate we use either the observed luminosities for the observational data or for synthetic stars, we use stellar luminosities $L_\star$ at 5 Gyr from the tracks by Baraffe et al. (2015). The incident flux of shortwave radiation on the planetary surface is on average $I_{ISR} = (1 - \alpha_{al})L_\star/(16\pi a^2)$, where α_{al} is the planets' albedo. Assuming that the planet is in thermal equilibrium and for a limiting case assuming no internal heat content, we can equate this to a critical outgoing long-wave radiation limit of $I_{OLR} = 281\,\mathrm{W\,m^{-2}}$ from Boukrouche et al. (2021). The resulting critical distance

$$a_{runaway} = \sqrt{\frac{(1 - \alpha_{al})L_\star}{16\pi I_{OLR}}}, \qquad (12)$$

is then a conservative upper limit within which our assumption that the water evaporates and mixes with other constituents holds. To remain conservative, we use a relatively high albedo of $\alpha_{al} = 0.4$ motivated by the results of Kopparapu et al. (2014) for high-pressure envelopes.

To generate a synthetic sample of planets we sample 10 000 systems from the original synthetic data weighted by stellar mass chosen such that the stellar mass distribution of the synthetic data matches the observed one (equivalent to Schlecker et al. 2022). We then proceed to draw "observed" planets by repeatedly (1000 times a set of 201) sampling planets weighted by their individual $p_{trans} \times p_{det}$ thus allowing for an individual, modeled planet to be present several times in the synthetic comparison sample. We note that this approach ignores possible correlations of the detection probability within a planetary system and we used discrete stellar mass bins (0.1, 0.3, 0.5, 0.7, 1.0, and 1.5 $M_\odot$) when the synthetic data was created (Burn et al. 2021).

Since the observed planets, to which we want to compare our models, are also characterized by radial velocity measurements, we further require for planets to be included in the synthetic comparison data that the semi-amplitude (e.g. Cumming et al. 1999, their Eq. 1) is larger than 0.3 m s^{-1}.

3. Results

3.1. Evolutionary tracks

In order to understand the dependencies and behavior of the different model assumptions, it is illustrative to visualize the time evolution of three different planetary mass cases under the four different model assumptions. This covers qualitatively the different possibilities for the final states of initially water-rich planets. Here, we initialized the planets by using a total mass – luminosity relation of Linder et al. (2019), and H/He and water fractions typical for the formation output. The parameters are tabulated in Table 1. Furthermore, the planets are placed after 3×10^6 yr at the same orbit with semi-major axis of 0.01 au around a 0.5 $M_\odot$ star.

	Light	Intermediate	Massive
Total mass [M$_\oplus$]	2.54	4.96	9.35
Envelope mass M_{env} [M$_\oplus$]	1.07	3.31	6.24
Core mass M_{Fe+Si} [M$_\oplus$]	1.69	2.26	4.68
Iron mass fraction f_{iron}	0.24	0.24	0.24
Envelope metallicity Z_{homo}	0.79	0.73	0.67
Luminosity L_{ini} [L$_{Jup}$]	2.81	12.29	50.50

Table 1. Initial conditions for the planetary evolution tracks shown in Fig. 3 and discussed in Sect. 3.1.
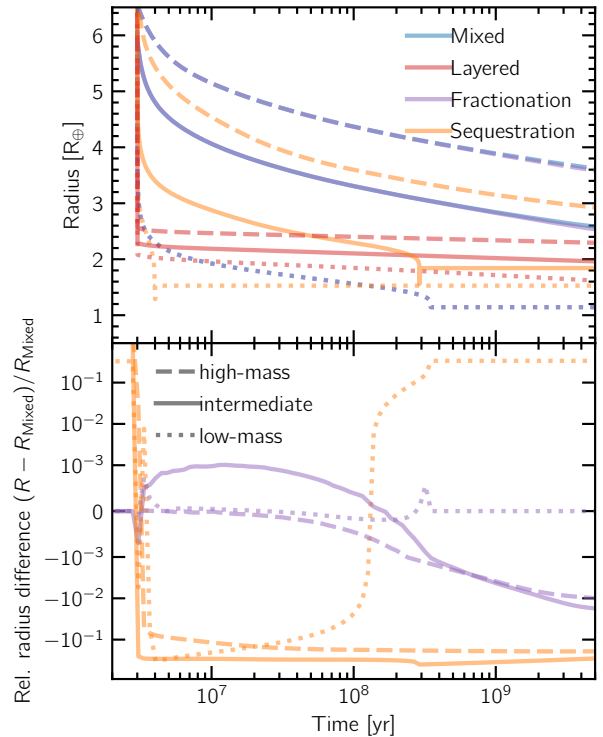

Fig. 3. Time evolution of planetary radii using the different models. The dashed lines show a more massive sub-Neptune with an initial mass of 9.35 M$_\oplus$, the solid ones that of an intermediate case (initial mass 4.96 M$_\oplus$), and the dotted lines represent the evolution of a lighter planet 2.54 M$_\oplus$. All planets are placed at 0.01 au around a 0.5 M$_\odot$ star (see text for further initial conditions). Note that the Fractionation model results overlap with the Mixed model result in the top panel. For this reason, the bottom plot shows the relative radius difference to the equal-initial mass Mixed model case. For the Layered model, the line is omitted in the bottom panel since the differences can be seen clearly from the top panel.

At time zero of the model, the Water Sequestration and Layered models distribute a fraction or all of the water contained in the envelope $Z_{homo}M_{env}$ either to the core or it is layered on top of the core. The gaseous envelope is therefore depleted in heavy particles and adjusts itself to a larger radius. As soon as the simulation starts evolving in time, the larger radius leads to a higher photoevaporative mass loss rate. Thus, the planet looses mass and consequently shrinks in size more rapidly than the cases with a Mixed or Fractionation assumption. This can be observed for the intermediate and massive planet and the effect is more pronounced in the Layered scenario compared to the Water Sequestration scenario as all water instead of just a part, corresponding to the possible uptake of the core and magma ocean, is removed from the envelope.

Depending on the initial mass and structure assumption, the gaseous envelope can be lost completely. For the Layered model, the $Z_{env} = 0$ envelopes are removed in all cases. For the Water Sequestration model, the most massive planet retains its gaseous envelope while the intermediate and light planets transition to a secondary envelope stage. This can be seen from the yellow curve dipping to the size of a bare core before the numerical scheme establishes a secondary atmosphere due to the reservoir of water contained in the magma ocean being allowed to re-equilibrate with the atmosphere. As mentioned in the model description and will be discussed in Sect. 4.3, this approach is relatively crude and could be improved in the future. Generally, the outgassed atmosphere is almost entirely made of heavy elements and of relatively low-mass. Although it could be removed on timescale comparable to the lifetime of the system, the replenishment from the much larger magma ocean reservoir typically sustains it for planets with $M > 2\,M_\oplus$ (Kite & Schaefer 2021).

Due to the faster loss of H/He in the initial stage, the Layered and Water Sequestration model result in planets with lower radii up to the point where also the Mixed model planet looses its atmosphere completely. In the cases shown in Fig. 3, this is the case after 300 kyr for the light planet. Since this planet in the Mixed model will have lost not only H/He but also its complete water budget, the planet ends as a smaller rocky core compared to the Water Sequestration model where an outgassed atmosphere remains and even more so compared to the Layered model result where a thicker steam atmosphere is present.

The Fractionation model differs less than a percent in radius from the Mixed model. There are hints of weakly fractionating mass-loss at the end of the lifetime of the atmosphere on the light planet (300 kyr, dotted purple line) where the atmosphere is removed slightly faster in the Mixed model case. If an observation at that point would have been made, the case where fractionation is included would appear slightly larger. For the more massive planets, the effect of fractionation is to marginally increase Z_{env} over time which leads to percent-level smaller radii due to the slightly increased mean molecular weight. A numerical artifact is the initially increased radius for the Fractionation model output for the intermediate-mass planet. It originates due to differing timesteps just after initialization where mass loss rates are large and slight overshooting occurred for the Mixed model, revealing the level of precision obtainable with our approach. Generally, the model is not suited for radius comparisons below the 1% level.

3.2. The mass-radius-period distribution in the steam and stone interpretation

3.2.1. Mixed model

We start by comparing the mass-radius relation of the synthetic planets using the nominal, mixed envelope assumption against the sample of observed planets in the top, left panel of Fig. 4. The synthetic distribution consists of volatile-rich planets with significant scatter in mass-radius relation distinct from a population of higher-density, rocky planets with a tightly correlated mass-radius relation. [3]

---

[3] Not discussed in the rest of this study, we find a few planets below the main rocky composition relation, made of pure iron in the model, which are sometimes observed (Goffo et al. 2023). In our calculations, they tidally migrated towards the star and lost their silicate mantle due to photoevaporation, using a simple approximation of hydrodynamic or jeans escape of an atmosphere with an assumed mean molecular weight of 49 u.

The rocky population seems to be in agreement with the observational data although there might be too many massive rocky bodies. The upper end of the rocky body distribution will be discussed quantitatively in Sect. 4.1.

The volatile-rich population of the mixed model has a decreasing number density with increasing radius which is more clearly seen if estimates of the occurrence probability density are added (Appendix C). The contours in Fig. C.1 show this information and make apparent that both the observed and synthetic planets are more likely to occur at around $2.7\,R_\oplus$ in contrast to larger radii. The synthetic planets however populate a parameter space at low masses and large radii where few observations are found. This regime is sensitive to the observational bias since the radial velocity semi-amplitudes are low. With increasing radius, there is no clear water mass fraction correlation visible in Fig. 4 but we further explored envelope metallicities, which do show a decreasing trend with increasing radius.

In addition to masses and radii, it is insightful to also compare orbital periods of the planets. Here, observations give very precise measurements which is why we omitted their errors from Fig. 5. For the Mixed model shown in the upper left panel, we can see an under-density of planets corresponding to the observed radius valley. It separates in our model high-density planets from low-density planets. The Mixed model output agrees well with the observational data in period-radius space. The observation that the innermost planets above the radius valley, the sub-Neptunes, are located at longer orbital periods than the innermost rocky planets is recovered.

To summarize, we qualitatively compared the Mixed model output against stellar-mass independent observational data in terms of masses, radii, and orbital periods. From this, we cannot discard the Mixed model and reasonably reproduce the observations.

3.2.2. Layered model

The results can be contrasted with those from the Layered model. In this case, two sharp transitions can be made out in the top right panel in Fig. 4. H/He-rich planets differ fundamentally from the water worlds as mixing between the two components is not allowed. Therefore, H/He can also be more easily lost to photoevaporation as H/He-rich planets are initially larger in radius if it is not mixed with heavier species as seen in Fig. 3 and discussed in Sect. 3.1.

This leaves a large population of planets with pure water envelopes which is then protected from photoevaporative loss due to the higher atmospheric mean molecular weight compared to mixtures. Therefore, less planets populate the low-density parameter space at larger masses $M > 5\,M_\oplus$. Instead, many planets are located in the pure water world regime at around two $R_\oplus$.

The lowest-mass volatile-rich planet of each population will be discussed in detail Sect. 4.1, but already from qualitative comparison, the Layered model can be considered a worse match to observations in terms of the volatile-rich planet distribution with too many low-mass planets and too few higher-mass, lower density planets. A similar conclusion can be drawn from a comparison in period-radius space (Fig. 5).

A noteworthy feature is that, in the Layered model, less planets populate the massive super-Earth regime. The pathway to form those planets requires stripping of the volatile elements by photoevaporation which is not efficient at those high masses if no light elements are present to decrease the overall density – as that is the case in the second stage of the Layered model evolution.
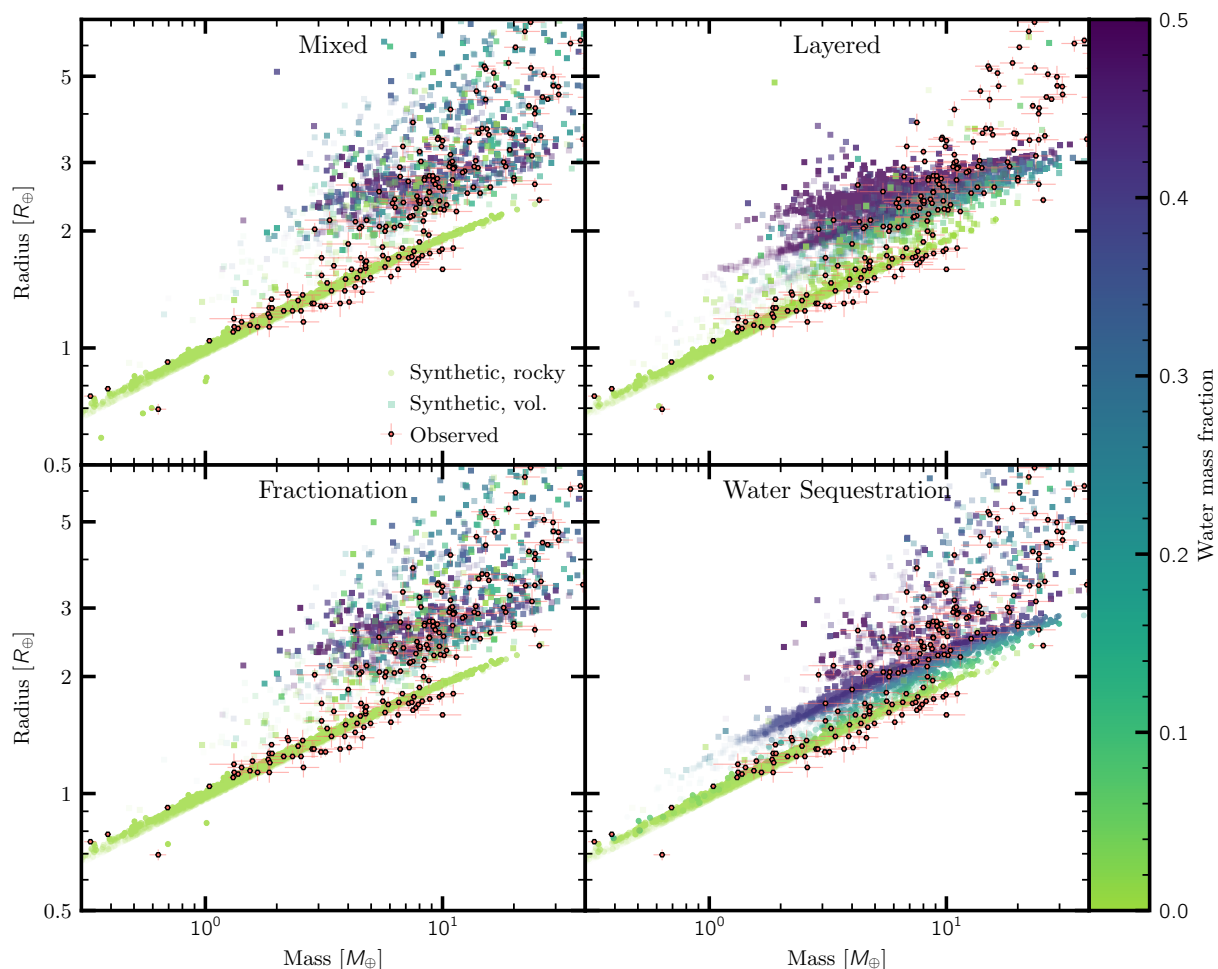
Fig. 4. Mass versus radius diagram of observed and synthetic planets. The four panels show the resulting planets using different model variations contrasted against the observational data from Parc et al. (2024). The synthetic data for all models is sampled according to the stellar mass distribution of the observations and an estimate of the detection and transit probability is applied. More saturated dots imply that the planet was sampled multiple times due to being more likely to be detected and their color is given by the planets' water mass fraction.

3.2.3. Fractionation model

The results of the Fractionation model can be summarized briefly: the effect of fractionating mass-loss has negligible consequences for the bulk planet properties using the considered initial conditions. Therefore, the results of the Fractionation model are similar to that of the Mixed model and we will in the following omit a separate discussion for the Fractionation model results. This result is interpreted in Sect. 4.2.

3.2.4. Water Sequestration model

The results of the Water Sequestration model can be broadly placed between the Layered and the Fractionation model. A part of the water content separates from the gaseous envelope and is therefore not lost to photoevaporation. From Fig. 4, we see that down to a mass of $2\,M_\oplus$, large water reservoirs can remain on the planet. At the end of the evolution (see also Sect. 3.1), this water is present in the form of oxygen and hydrogen in the metallic core, in the magma ocean, as well as outgassed as steam atmosphere. Planets in this state are tightly correlated in mass-radius while planets with primordial H/He content are distributed to larger radii. The shape of the mass-radius distribution is similar to the output of the Layered model, however, the planets with

steam atmospheres are significantly smaller than in that model. This is due to only a fraction of the initial water content remaining instead of the total amount as in the Layered model.

Compared to observations, the extension of water-rich planets to masses blow $3\,M_\oplus$ is not observed (see also Sect. 4.1). Furthermore, the model also predicts a population of H/He free planets with large water contents (i.e. water worlds) on a narrow mass-radius relation. While some observed planets populate this region of the parameter space, they are not as numerous as predicted by the model (see also the analysis of Parviainen et al. (2024) on observational evidence for water worlds). Furthermore, from Fig. 5, we find a population of these water wolds (with $R \sim 2\,R_\oplus$) at short orbital periods (~ 1 day) which is not matched by observations. The Mixed model reproduced this innermost distribution of planets better. However, the mass-radius distribution of larger sub-Neptunes is well-matched. In particular, there is less scatter than in the Mixed model and no planets with too low densities exist in the sub-Neptune regime. This indicates that lowering the water fraction in the gaseous envelope led to a better match with observations for planets keeping their primordial envelopes.
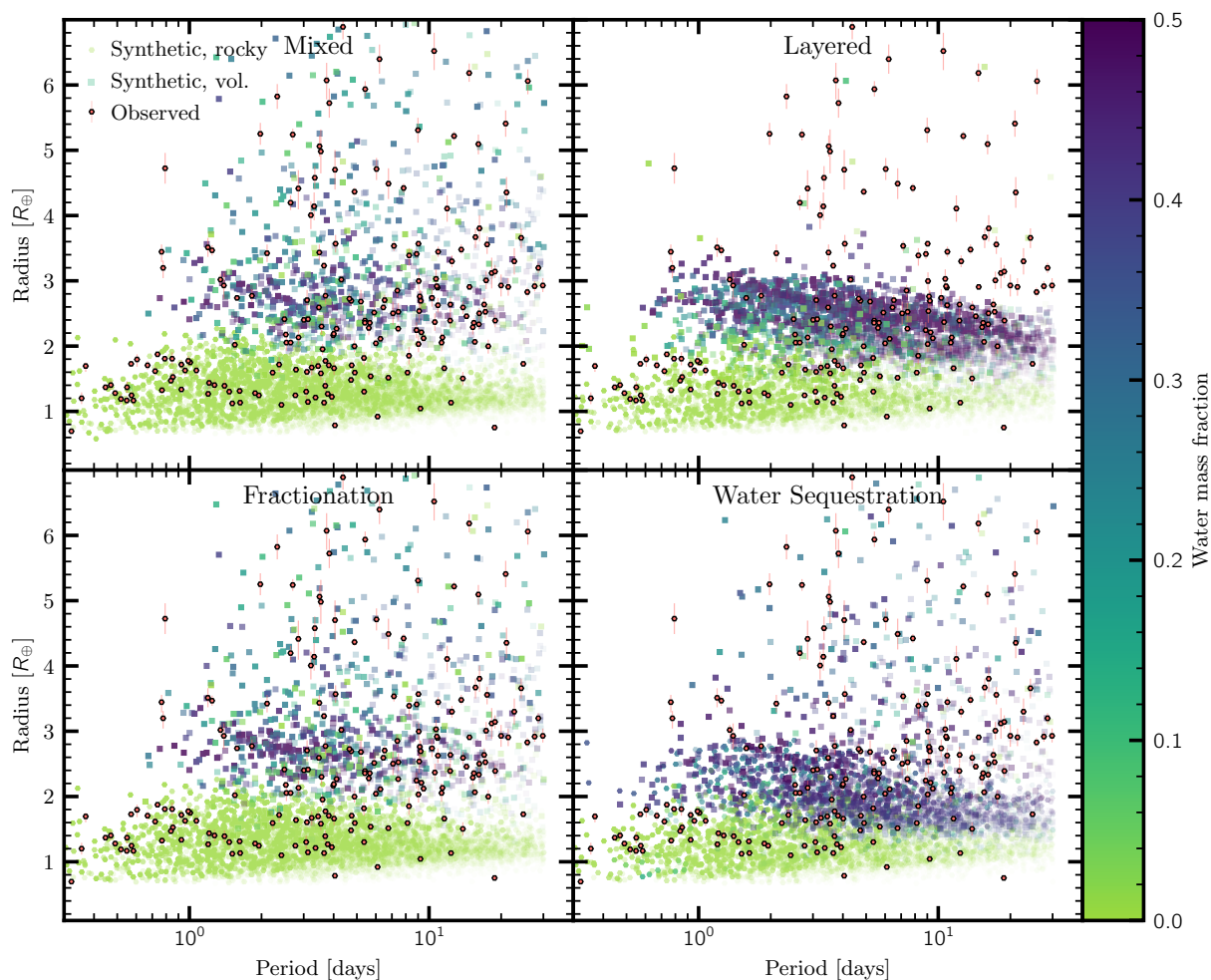
Fig. 5. Orbital period against planetary radii of observed and synthetic planets. The samples and coloring are the same as in Fig. 4.

3.3. The radius valley as a function of stellar mass

Here, we investigate the dependency of the emerging radius valley on stellar mass. In Fig. 6, we show for the Mixed model the full synthetic data without an observational bias applied for different discrete stellar masses. Since the previously introduced bias is applicable to planets characterized in mass and radius, we opt to not use an estimate of the bias in this radius comparison (irrespective of planetary masses) to investigate the actual trends predicted by the model. Complementarily, we will discuss the radius valley trends using the populations around Solar mass stars with applied Kepler observational bias in Sect. 3.4. To support the discussion here, we also show a summary of fits done for observed or synthetic data in Fig. 7 and mass-radius relations resulting from the Mixed model for the different stellar mass bins in Fig. 8. The results of the other models are presented in the same fashion in Appendix C. In each panel of Figs. 6, C.4, and C.5, `gapfit` (Loyd et al. 2020) was used with bootstrapping to infer a gap in the region of the observed gap, that is, using planets from 1 day to 200 days orbital period for stellar masses larger than $0.1\,M_\odot$ and from 0.6 to 20 days for the $0.1\,M_\odot$ case. The gap is fit using a power-law

$$R_{\mathrm{gap}}(P) = R_{\mathrm{gap}}(P/1\,\mathrm{day})^{u_{\mathrm{gap}}}\,R_\oplus.\qquad(13)$$

The search for a gap is performed within 0.15 dex of $\log_{10} R_{\mathrm{gap}} = 0.3$, with an initial slope guess of $u_{\mathrm{gap}} = -0.1$, and a Kernel width of 0.15.

The synthetic data at low stellar mass for the Mixed model shows a similar gap location at $\sim 1.9\,R_\oplus$ in agreement with Luque & Pallé (2022). We can see a trend with stellar mass: the locus of the valley at 1 d orbital period shifts upward for late M dwarfs before converging to $R_{\mathrm{gap}} = 2.38$ for stellar masses larger than $0.5\,M_\odot$. This kind of threshold stellar mass is similar to the mass at which the small planet population no longer increases significantly in mass, as discussed in Burn et al. (2021). In that work, we attributed this to the fact that growth of small planets is no longer limited by the amount of solids in the disk at stellar masses larger than $0.5\,M_\odot$ which we might recover here since growth determines the initial mass and thus indirectly the radius of the Super-Earth population.

The observed radius valley indicates a dependency on stellar mass around FGK-type stars based on Kepler and K2 data (Berger et al. 2020; Petigura et al. 2022; Ho & Van Eylen 2023). The reported logarithmic slopes of the radius valley location with stellar mass $d \log_{10} R_{\mathrm{gap}}/d \log_{10} M_\star$ range from 0.1 to 0.4 with likely values around 0.2. However, for M stars, Luque & Pallé (2022) report a slope consistent with no stellar mass dependency (0.08 ± 0.12). Although it is likely not the relevant functional form, we fit a power-law to the inferred radius valley locations for the five presented stellar mass bins at 1 day orbital period and obtain $d \log_{10} R_{\mathrm{gap}}/d \log_{10} M_\star = 0.08$, which agrees with the observed trend for M dwarfs (Luque & Pallé 2022) but is flatter than found for FGK stars where values around 0.2 are reported
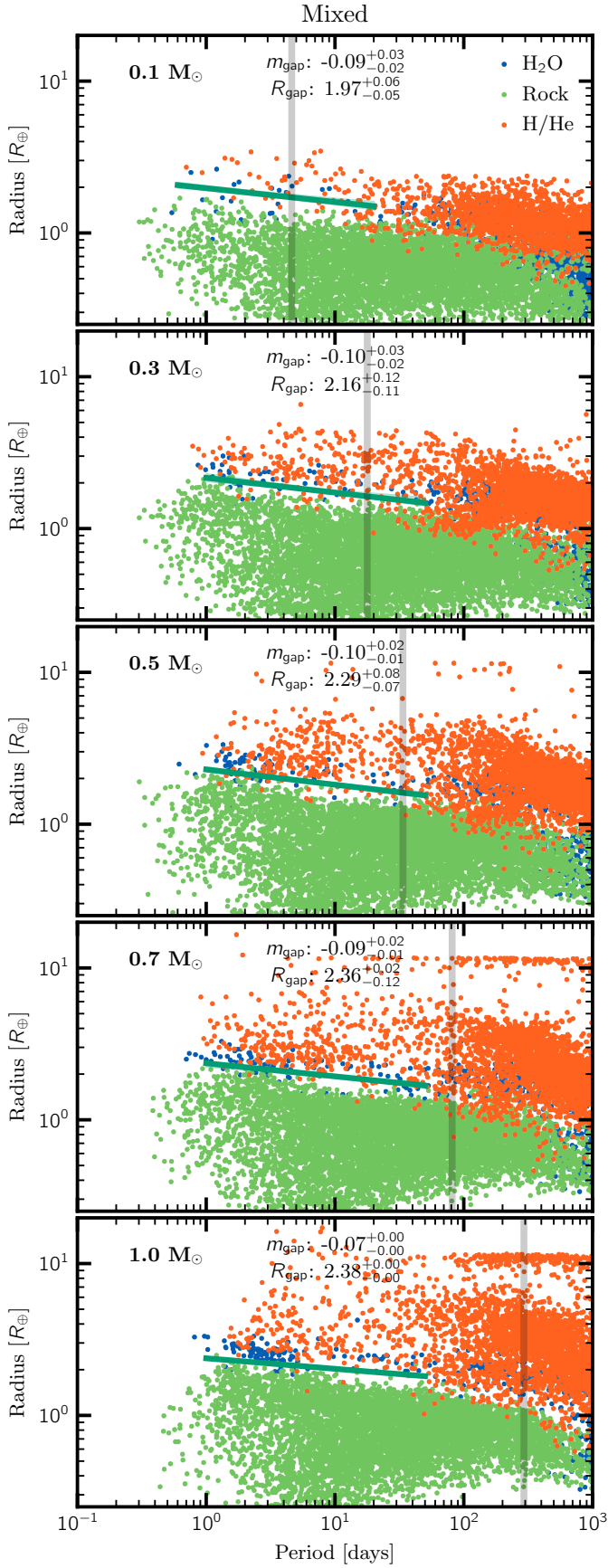
Fig. 6. Period versus radius of unbiased, synthetic planets using the mixed model around stars of different masses. For each stellar mass (top left), a gap in the data was fitted (`gapfit`, Loyd et al. 2020) and the resulting slope ($u_{\rm gap}$) and locus ($R_{\rm gap}$) in logarithmic space are listed. The gray line marks $a_{\rm runaway}$ (Eq. 12) within which runaway heating occurs.
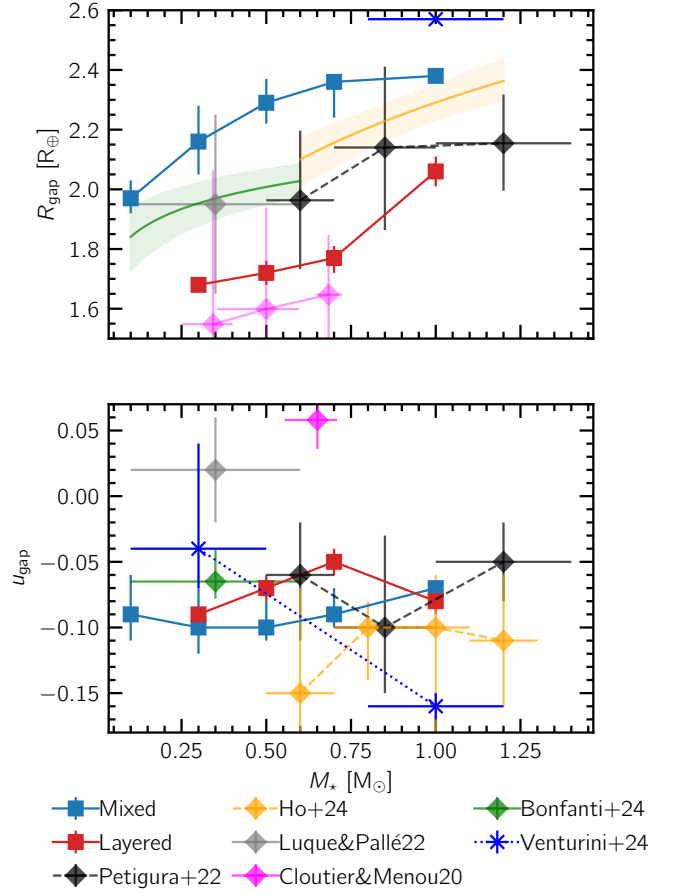

Fig. 7. Comparison of derived loci $R_{\rm gap}$ and slopes with orbital period $u_{\rm gap}$ for the radius valley. Observational data from Petigura et al. (2022); Luque & Pallé (2022); Ho et al. (2024); Bonfanti et al. (2024) is shown with diamond shaped markers and errors in x-direction showing the range of stellar masses included. From this work, only the Mixed (light blue squares) and Layered (red squares) model output is included here since the Fractionation model gives similar results to the Mixed case and for the Sequestration model no trustworthy fit can be obtained. For the same reason, Venturini et al. (2024) (dark blue) did not report $R_{\rm gap}$ for their sample of modeled planets around low stellar mass. The values for $R_{\rm gap}$ by Luque & Pallé (2022) is read-off from their figure. The results of Cloutier & Menou (2020) and Petigura et al. (2022) are based on occurrences instead of detections. For Petigura et al. (2022) the values for $R_{\rm gap}$ were shifted to 1 day orbital period (incl. Gaussian error propagation). For the results from Cloutier & Menou (2020), we report their values which were obtained from marginalizing over orbital periods. The values are only comparable to the other derivations because their derived $u_{\rm gap}$ is close to zero. For Bonfanti et al. (2024, green line and region, $M_{\rm star} < 0.6\,M_\odot$) and Ho et al. (2024, orange, $0.6\,M_\odot \lesssim M_{\rm star} \lesssim 1.2\,M_\odot$), we show their fitted continuous function for $R_{\rm gap}(M_\star)$ with shaded regions marking 16 to 84% confidence intervals.

(Petigura et al. 2022; Ho & Van Eylen 2023). From Fig. 7, we see that also the data by Petigura et al. (2022) shows a flattening of $d \log_{10} R_{\rm gap}/d \log_{10} M_\star$ for the two higher stellar mass bins.

While the locus of the valley increases, our inferred slopes with orbital period hints at a flatter dependency with increasing stellar mass. The values range from $d \log_{10} R/d \log_{10} P = u_{\rm gap} = -0.10$ at $0.3\,M_\odot$ to -0.07 for Solar mass stars. This is comparable to other theoretical works relying on photoevaporation predictions from hydrodynamic models instead of simple energy- or recombination-limited escape models (the latter resulting in steeper slopes Mordasini 2020; Affolter et al. 2023). Although

the error from the bootstrap method is small, we expect that the flattening is not based on enough statistical data from the model to confirm it as a significant trend. The flattening could be influenced from below, by growing more massive super-Earths at larger orbital periods around more massive stars, or from the top by less efficient photoevaporation and thus smaller radii of volatile-rich planets at fixed orbital period around lower-mass stars. Our results remain for all stellar masses in the regime of negative slopes indicating that photoevaporation is the dominant effect shaping the valley. The compilation of observational data (Fig. 7) does not show a significant trend with the exception of the inferred slope by Luque & Pallé (2022). However, the theoretical work of Venturini et al. (2024) predicts a steeper slope around more massive stars which was not recovered here.

Ho et al. (2024) and Bonfanti et al. (2024) gave indications for a more populated, radius valley around low-mass stars. Similarly, Cloutier & Menou (2020) already noted a narrowing of the radius valley towards lower stellar masses. Apart from the $0.1\,M_\odot$ case where the statistics becomes poor, we find a similar trend in Fig. 6. This trend was also recovered and discussed in Venturini et al. (2024). From the detailed discussion of mass-radius relations (Sect. 4.1), we interpret part of it being linked to the low-mass tail of the volatile distribution extending into the radius valley (but not the density valley, see Luque & Pallé 2022; Venturini et al. 2024). We note that the M star sample probes the planetary population at lower irradiations where condensation could occur, but we excluded this region when we plot biased populations. We leave a more detailed analysis of the pollution of the valley to future works.

In addition to the data, we show an estimate of the runaway greenhouse limit (Sect. 2.3) as gray line in Fig. 6. The fact that for low-mass stars, this limit is not necessarily outside the regime of interest further motivates the investigation of a model which does not assume that water and H/He mix, that is, the Layered model. Thus, in Fig. C.4, we show the resulting unbiased planetary population with the previously introduced compositional color code using the Layered model. The most striking effect is that the population of H/He-free sub-Neptunes is more numerous and extends to longer orbital periods than in the Mixed model. If the assumption that no water enters the low-density envelope is made, also planets at larger orbital periods can lose their H/He content. This also leads to a more abrupt transition from H/He free to H/He-rich planets at a threshold radius for a given orbital period. Therefore, a second radius valley emerges between H/He-rich and H/He-free planets at $3\,R_\oplus$–$4\,R_\oplus$. So far and to our knowledge, no reports on such a valley have been published. However, a visual indication of such an under-density can be made out in the data of Ho & Van Eylen (2023) for Solar mass stars, their Fig. 1, although it does not seem to be of high enough significance. We further note that in the Solar mass data, the observable region is well within the runaway greenhouse limit where we expect mixing.

The algorithm could determine a radius valley for all but the $0.1\,M_\odot$ case for the Layered model. It lies between rocky planets and water worlds and at too low radii compared to observations (Fig. 7) because pure water envelopes instead of mixed-composition ones are less prone to being removed by photoevaporation (see e.g. Mordasini 2020, for a parameter study). However, the distance to the observed valley is comparable to the Mixed model. This seems to naively suggest that reality lies between the modeled assumptions.

For the Water Sequestration model, no radius valley can be fitted conclusively around the three lowest stellar mass bins. Although there is a transition in density space, the existence of
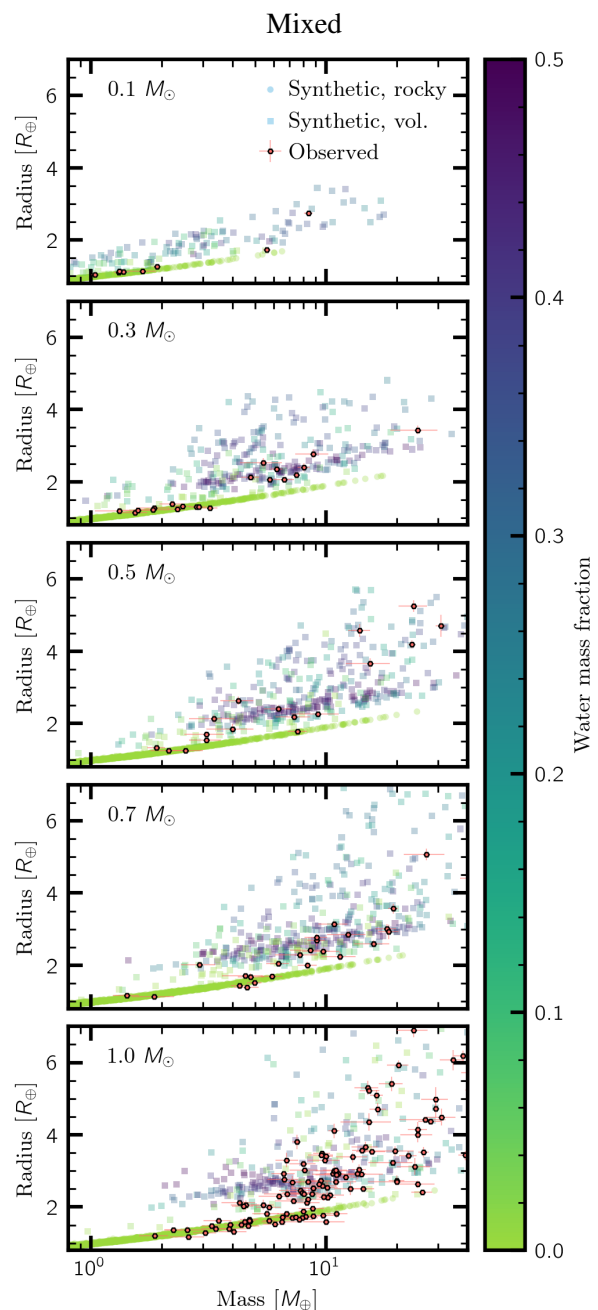
Fig. 8. Masses against radii for observed and synthetic planets using the Mixed model for various stellar masses. In contrast to Fig. 4, no observational bias is applied. The selection is only restricted to planets within 30 days, respectively the runaway greenhouse limit (for 0.1 and $0.3\,M_\odot$). The synthetic data is calculated at the stellar mass indicated at the top left; observational data is grouped by restricting stellar masses (as tabulated by Parc et al. 2024, see references therein) to the nearest corresponding reference stellar mass.

steam worlds with low mass ($\sim 2\,M_\oplus$) makes this density valley undetectable in radius space. This can also be seen when comparing the mass-radius relations shown in Fig. 8 with Fig. C.3.

3.4. The radius valley as seen by Kepler

The ability of the different models to reproduce the observed radius valley can also be tested against the Kepler dataset. As men-
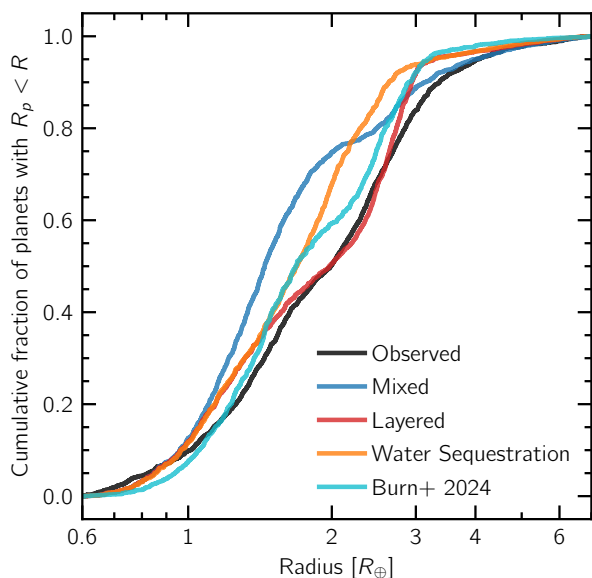
Fig. 9. Cumulative distribution of synthetic planetary radii compared to Kepler data. The synthetic data is obtained after modeling the observability with Kepler using KOBE (Mishra et al. 2021). In addition to the datasets obtained here, we show the results of Burn et al. (2024). Only planets with radii ranging from 0.6 to $7\,R_\oplus$ are included. The observed distribution is taken from Fulton & Petigura (2018).

tioned in Sect. 2.3, we constructed a set of synthetic systems as the Kepler spacecraft would observe it. Given by the sensitivity of Kepler, we restrict the selection of simulations to those around Solar mass stars. In Fig. 9, the resulting cumulative distributions of the synthetic planetary radii can be compared to the observed ones. We further added the distribution by Burn et al. (2024) since several assumptions on the initial conditions were changed although the Mixed model is physically similar to the model used in that work.

We find a reduced number of sub-Neptunes in the Mixed model compared to Burn et al. (2024) in disagreement with observations. The high-radius tail of sub-Neptunes, that is, the radius cliff, is best matched in the Mixed model. The radius valley location, also discussed in Sect. 3.3 where a radius valley location at $2.4\,R_\oplus$ was found, is shifted to too large radii in the Mixed model. This implies that the sub-Neptunes in the Mixed model loose their envelopes too rapidly. However, we caution that the sub-Neptune to rocky planet ratio is sensitive to uncertain formation model assumptions (production of rocky planetesimals and planets), as well as to the application of the observational bias. The locus of the radius valley is less sensitive to these considerations.

From Fig. 9, we find that Layered model reproduces better the observed radius valley location, although shifted to too low radii for lower stellar masses (see Fig. 7). Furthermore, it has a more favorable sub-Neptune to rocky planet ratio compared to the Mixed model. From this comparison on planetary radii only, one might conclude that it should be favored over the Mixed model. In contrast, the Water Sequestration model lacks the distinct radius valley feature.

4. Discussion

4.1. Key quantities for comparison of mass-radius relations

To better compare the models and observational data, we introduce here four quantities which are not expected to be influenced by observational biases and should relate to different, stellar-mass-dependent processes shaping the period-mass-radius distribution of small planets.

- **The upper-mass end of the rocky planet distribution** $M_{\mathrm{rock,max}}$. To not rely on outliers, this is measured as the 90th quantile of the rocky planet mass distribution. Rocky planets are identified by having a normalized density of at least 0.65 times that of an equal-mass Earth composition planet (using the relation by Zeng et al. 2016).
  It should be either determined by the most massive planet which can still lose its envelope or the most massive planet which can form in the inner system. If the former is the case, it should be sensitive to the shortest orbital period planets where photoevaporation is strongest. Since transit and radial-velocity methods are most sensitive at short orbital periods, we expect no influence from observational biases.
- **The lower-mass end of the volatile-rich planet distribution** $M_{\mathrm{vol,min}}$. Analogous to $M_{\mathrm{rock,max}}$, the 10th quantile of all other planets but the rocky ones is measured.
  Since all planets are expected to form with gaseous envelopes, this quantity is determined by the lightest planet which can retain its envelope. Again, this should be sensitive to orbital periods and is therefore influenced to some degree by the observational bias (the outermost planets which can be detected). However, it is nevertheless a useful property for inter-comparison of models and stellar-mass dependent trends. Alternatively, if stripping of (water-rich) envelopes is not efficient, this point could also be set by the lowest-mass planet which is volatile-rich and observable. Practically, since planets beyond the water iceline are not observable, this would give direct access to the lowest-mass planet which migrated from beyond the water iceline.
- **The volatile-planet mass-radius relation slope** $m_{\mathrm{vol,\,M-R}}$. This can be fitted using least-squares fitting as implemented in the `scipy` python package (Virtanen et al. 2020). By using a single slope, we assume that the mass-radius relation is of the form $R(M) = R_0/R_\oplus (M/M_\oplus)^{m_{\mathrm{vol,\,M-R}}}$. To allow for an estimation of the error, we use a bootstrapping approach for the observed data where for 1000 repetitions, an equal-sized distribution of planets is drawn with replacing from the observed data. In addition, we randomize the values of the observed data by sampling from a skewed normal distribution based on the errors in mass and radius space. For synthetic data, we obtain comparable error estimates by drawing 1000 times a sample of planets with the same number of points as the observed data (using a Jackknife approach, i.e. without replacing).
  Physically, the slope $m_{\mathrm{vol,\,M-R}}$ results from a combination of atmospheric density scaling from envelope fractions and mean-molecular weight effects as well as being shaped by mass-loss. For a mixed envelope, a more steep slope is expected for lower metallicity envelopes and vice-versa.
- Finally, the **radius valley** characteristics are further important quantities to reproduce. We introduced our approach to quantifying this aspect in Sects. 3.3 and 3.4.

Fig. 10 provides a means for comparison of mass-radius distributions with corresponding power-law fits to the volatile distribution as well as visualizing the quantities $M_{\mathrm{rock,max}}$ and
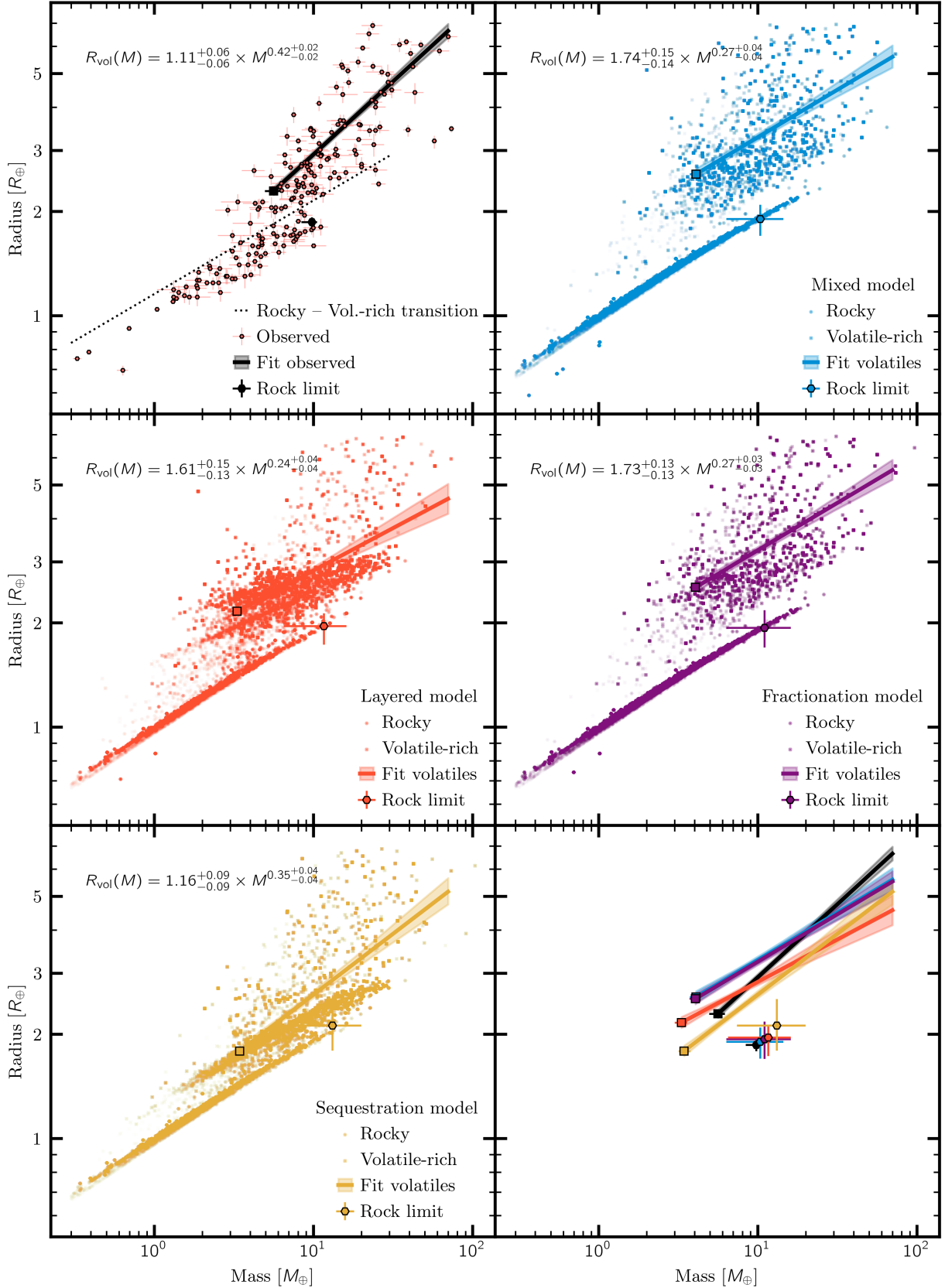
Fig. 10. Comparison of observed exoplanets and different model results in the mass radius plane. The synthetic data includes the estimate of the observational bias as detailed in Sect. 2.3. A power-law was fitted to the volatile-rich populations for each panel and indicated in the top left using units of Earth radii and masses (see text for details). The lower end is drawn at the 10th quantile of the volatile distribution ($M_{vol,min}$). Furthermore, the 90th quantile of the rocky planet population ($M_{rock,max}$) is indicated with errors obtained from the same resampling technique. The bottom right panel shows repeated fits for the different populations for better comparison to each other.

$M_{\mathrm{vol,min}}$. We find that all models predict a smaller value than the observed $M_{\mathrm{vol,min}}$, match within errors $M_{\mathrm{rock,max}}$, and predict a significantly shallower slope of the volatile planet mass-radius relation $m_{\mathrm{vol, M-R}}$ than the observed $m_{\mathrm{vol, M-R}}$. We find that the quantity $M_{\mathrm{vol,min}}$ is in our model shaped by mass-loss and not imprinted from formation. The population of water-rich planets after formation extends to lower masses (i.e. to approximately $1\,M_\oplus$ smaller than $M_{\mathrm{vol,min}}$). Since the values lie below the observed ones, this hints at too inefficient photoevaporation caused by the prescription itself or by too high mean molecular weight atmospheres in the modeled planets.

While it can be a useful probe for photoevaporation, the values for $M_{\mathrm{vol,min}}$ could be influenced by the observational bias. In contrast, the slope of the mass-radius relation is a less bias-dependent indication that too many planets with high metallicity envelopes are present in the models. This can also be seen from a bifurcation in the volatile-rich population with a group clustered at relatively low radii and another with larger radii. The lower population is made from almost pure steam worlds which is at odds with observation. Although, some planets populate this region (e.g. GJ 9827 d, Passegger et al. 2024), they are not as commonly observed as predicted by the model. This group is even more abundant in the Water Sequestration and Layered models. We note that the analysis for the Water Sequestration model is made more difficult as many water-rich planets with small outgassed atmospheres are classified to be part of the rocky population.

The qualitative observation that $M_{\mathrm{vol,min}} < M_{\mathrm{rock,max}}$ is captured by all models. If all rocky planets originate from being stripped from an equal-composition massive envelope, there should be an irradiation trend visible with rocky planets close to $M_{\mathrm{rock,max}}$ orbiting on short orbital periods and volatile-rich planets at the low-end $M_{\mathrm{vol,min}}$ of the volatile-rich mass-radius distribution an wide orbits. We visually inspect the data for such a trend and do not find any correlation in the observed data. In the synthetic data from the Mixed model, however, the most massive rocky planets indeed orbit on short orbital periods and there are more long-orbit volatile-rich planets around $M_{\mathrm{vol,min}}$.

We can further try to identify trends with stellar masses for the mass-radius characteristics. In Fig. 11, the observational data and Mixed and Water Sequestration model output is split into two stellar mass bins with a separation at $0.6\,M_\odot$. One apparent effect is that the two characteristic masses $M_{\mathrm{vol,min}}$ and $M_{\mathrm{rock,max}}$ lie closer to each other around lower stellar masses for the synthetic and observed data. Instead of an overlap in mass space of the volatile-rich and rocky population, they show a clear transition at $\sim 4\,M_\oplus$. In the Mixed model, this is mainly because the rocky planet population decreases in mass. It is to a large degree sourced by planets which formed interior to the iceline and this population shrinks with decreasing stellar mass (Burn et al. 2021). The findings here seem to contradict the theoretical findings of the work by Venturini et al. (2024). They found for their pebble-accretion based planets (with structures analogous to the Mixed model) that the overlap of the volatile-rich planets decreases with increasing stellar mass – opposite to what we find. However, this difference is because we applied a cut at the runaway greenhouse limit to our synthetic planets (Sect. 2.3). The planets with radii larger than that expected from condensed water are also not present around low mass stars in the work of Venturini et al. (2024, their Fig. 1). Indeed, when excluding planets cold enough to allow for condensed water, their data reveals a similar transition from rocky to volatile-rich planets at 3-4 $M_\oplus$ around low-mass stars.

The comparison suggests that our Mixed model could be improved if planets have lower initial envelope metallicities: the slope of the volatile-rich planet mass-radius relation would steepen and the $M_{\mathrm{vol,min}}$ would increase towards a generally better match with observations. The solution is however not to dissolve large quantities of water in the magma ocean and metallic core as this water will be later outgassed and remain on too low-mass planets in disagreement with observations. Instead, it is more likely that less water or more H/He should be accreted during formation. The latter option is unlikely as the formation model already includes optimistic estimates on the envelope cooling efficiency (via reduced assumed opacities, Mordasini 2014). Potential mechanisms for a reduction of water accretion could be radioactive heating and subsequent loss of water from planetesimals (Lichtenberg et al. 2019) or recycling of water back to the ambient gas (Moldenhauer et al. 2022; Wang et al. 2023). However, with a reduction of the atmospheric metallicity the radius valley location would shift to even larger radii (Mordasini 2020) in disagreement with Kepler data (see Sects. 3.3 and 3.4). Thus, the process which reduces the water content might require a planet mass-dependency which does not affect low-mass planets but more higher mass ones, such as expected for recycling (Wang et al. 2023).

4.2. Negligible effect of fractionation

In this work, we included the effect of fractionating outflows of oxygen relative to hydrogen for the evolution of synthetic exoplanets (see Appendix A for the implementation). This effect was studied in the context of exoplanets and coupled to evolutionary calculations by Hu et al. (2015), Malsky & Rogers (2020); Malsky et al. (2023), Gu & Chen (2023), and Cherubim et al. (2024) although for He or deuterium. Since oxygen is heavier, the effect should be more pronounced. However, we did not find a significant effect on the radius.

Since the initial envelope masses are considerably massive, following the output of our formation modeling, a large photoevaporative mass loss is required to alter their radius significantly. Such a flow is however not in the regime where fractionation occurs as can be seen from Eq. (A.1), where the mass loss Φ_{H} is in the denominator of the subtrahend. Thus, for $\Phi_{\mathrm{H}} \gg (m_{\mathrm{O}} - m_{\mathrm{H}}) b_{\mathrm{H,O}} \frac{GM}{R^2 k_{\mathrm{B}} T_{\mathrm{iso}}}$, which evaluates to a relative mass loss of $\dot{M}_{\mathrm{H}}/M \gg 5 \times 10^{-13}$ 1/yr, fractionation does not occur $x_{\mathrm{O}} \approx 1$ and mass of oxygen is lost proportionally to its fraction (Hunten et al. 1987; Zahnle & Kasting 1986; Zahnle et al. 1990). While this estimate is based on analytical theory, the numerical investigation of Johnstone (2020) found a similar estimate of this so-called critical flux. This is satisfied for most planets in our simulation until the last stage of the evolution (see Sect. 3.1) in general agreement with the findings of Cherubim et al. (2024). However, in that study, He-rich atmospheres remained in some cases. Although this difference should be investigated in more details, it could be due to their study also sampling small initial envelope fractions on the percent level. This speculation is supported by the study of Gu & Chen (2023) who found a dependency on the initial gas envelope mass for deuterium fractionation. Such low envelope fractions are not obtained from our formation modeling and are therefore not included in this study but could be candidates for fractionation to occur to a more significant degree.

We do not find a significant influence of the estimate of the upper atmosphere number density (see Appendix A). Further extension of the fractionation model to different species, such as
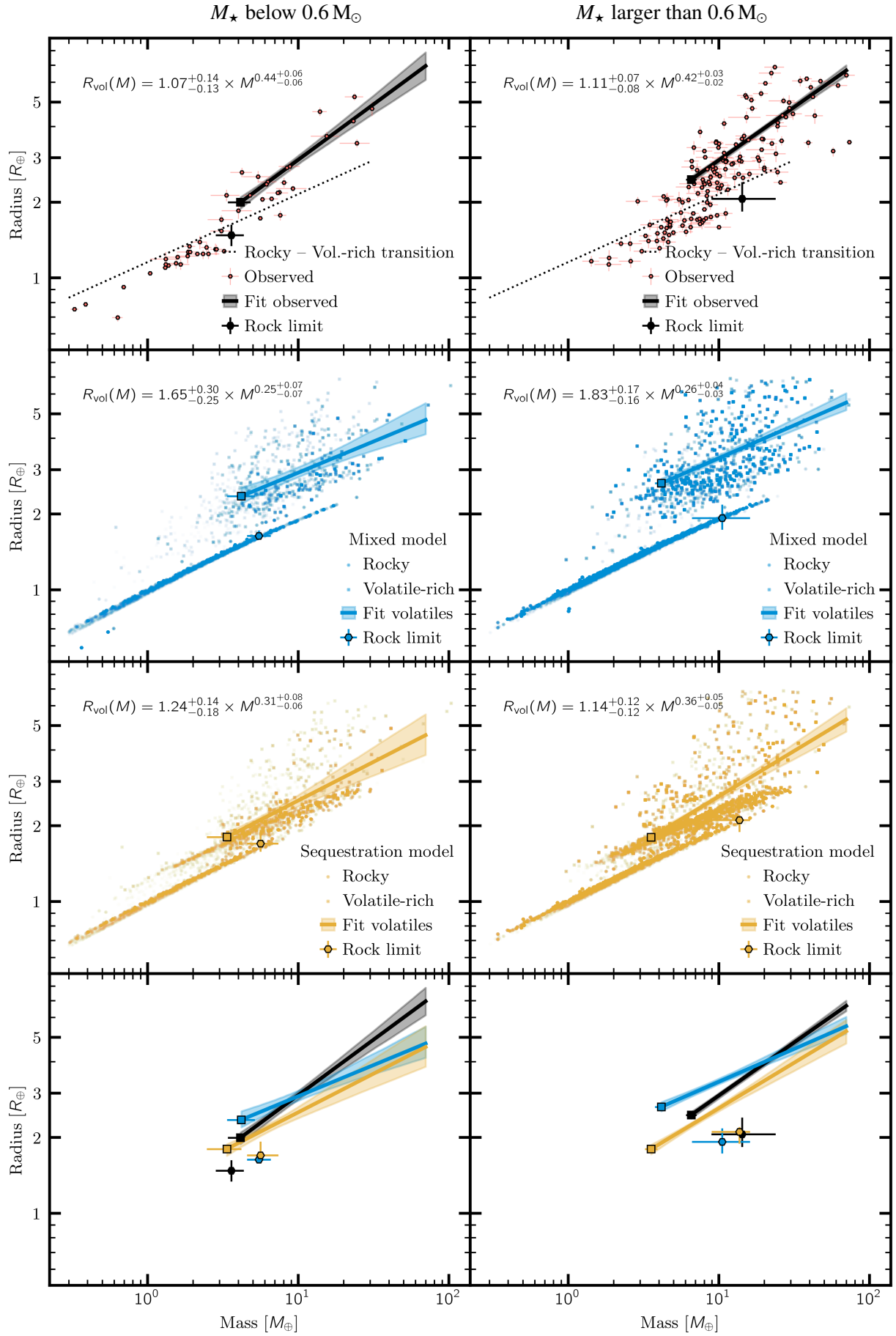
$M_\star$ below $0.6\,M_\odot$

$M_\star$ larger than $0.6\,M_\odot$

$R_{\rm vol}(M) = 1.07^{+0.14}_{-0.13} \times M^{0.44^{+0.06}_{-0.06}}$

Rocky – Vol.-rich transition
Observed
Fit observed
Rock limit

$R_{\rm vol}(M) = 1.11^{+0.07}_{-0.08} \times M^{0.42^{+0.03}_{-0.02}}$

Rocky – Vol.-rich transition
Observed
Fit observed
Rock limit

$R_{\rm vol}(M) = 1.65^{+0.30}_{-0.25} \times M^{0.25^{+0.07}_{-0.07}}$

Mixed model
Rocky
Volatile-rich
Fit volatiles
Rock limit

$R_{\rm vol}(M) = 1.83^{+0.17}_{-0.16} \times M^{0.26^{+0.04}_{-0.03}}$

Mixed model
Rocky
Volatile-rich
Fit volatiles
Rock limit

$R_{\rm vol}(M) = 1.24^{+0.14}_{-0.18} \times M^{0.31^{+0.08}_{-0.06}}$

Sequestration model
Rocky
Volatile-rich
Fit volatiles
Rock limit

$R_{\rm vol}(M) = 1.14^{+0.12}_{-0.12} \times M^{0.36^{+0.05}_{-0.05}}$

Sequestration model
Rocky
Volatile-rich
Fit volatiles
Rock limit

Radius [$R_\oplus$]

Mass [$M_\oplus$]

Fig. 11. Comparison of mass-radius relations with observational data split by stellar mass. Data is presented and fitted equivalently to Fig. 10. The left column shows simulations and observations with $M_\star < 0.6\,\mathrm{M_\odot}$ while the panels on the right show the data for more massive stars.

He, and including a realistic temperature and flow profile should be investigated in the future (see also the discussions of deviations from analytical estimates in Johnstone 2020; Schulik & Booth 2023).

There are observational prospects to probe the atmospheres of planets at the upper edge of the radius valley (Malsky et al. 2023; Cherubim et al. 2024). Since He and deuterium are unlikely to be outgassed from the interior, they provide a probe for the process of fractionation occurring. However, a future observational confirmation of the absence of He-dominated atmospheres should not be over-interpreted as being caused by the complete absence of photoevaporation. Instead, the planet could have had a more massive initial gaseous envelope and photoevaporation could be efficiently acting and also remove the He content of the planet. Rather, the presence of He or deuterium rich atmospheres could be used as a probe for initial gaseous envelope masses.

4.3. Outlook on models including sequestration

Here, we report comparisons to observational data which do not favor the most physically advanced Water Sequestration model. However, this finding should be put into relation to the developments required for a comprehensive model with atmosphere-interior coupling.

First, the Water Sequestration model includes only the dissolution of water but excludes that of the other modeled constituents, namely He and hydrogen. For He, this could be justified, but hydrogen dissolution into the interior is expected to occur (Hirschmann et al. 2012; Kite et al. 2019). Moreover, the dissolved material is bound to undergo chemical reactions. Instead of outgassing water, the secondary envelope can be of a different composition. However, if large quantities of oxygen were accreted via water ice and correspondingly, the mantle is in a high redox state, water is indeed expected to also be the main constituent of the outgassed atmosphere (Bower et al. 2022; Lichtenberg et al. 2023). Nevertheless, given how sensitive photoevaporative mass-loss is on the radius and therefore the molecular weight of the atmosphere, more outgassed species should be included. In particular, gas accretion predicts that the hydrogen to oxygen ratio increases with increasing planetary mass (Mordasini et al. 2016), therefore, the evolution of the high-mass planets in our simulation would change more significantly and outgassed hydrogen and other species could improve the match with observations in this regime.

Another neglected effect is the temperature dependency of the solubility. For simplicity, we assumed a constant exterior temperature of $3000\,\mathrm{K}$ for the computation of the water sequestration. Instead, the planet would cool and the retention of water in the magma ocean would therefore decrease (Chachan & Stevenson 2018). This would imply earlier outgassing than found here and potentially more efficient loss of water from the planet. Qualitatively, this effect is not expected to change the mass-radius relation of the Water Sequestration model, it would only reduce the average radius of the closely-correlated, water-rich, (almost) envelope-free planets.

A potentially more important related effect is however the solidification of the magma ocean (Hamano et al. 2013; Bower et al. 2018, 2022). Depending on the details of crystallization pattern and viscosity (e.g. whether or not a solid lid forms at the top of the mantle), most of the volatile budget could be removed from the mantle. If mantle solidification occurred on hot planets up to four Earth masses, this effect might reconcile with observa-

tions the low-mass end of the Water Sequestration model results presented here.

An exploration in this direction is also motivated by the findings of Gaidos et al. (2024) who explore the age dependency of the small planet population. They suggest that the atmosphere-interior coupling could produce effects on Gyr timescales as observed and further motivate an inclusion of water condensation. Qualitatively, water condensation could have a similar effect as switching from the Mixed to the Layered scenarios and could be quantified using these approaches as a basis (see also the discussion in Venturini et al. 2024).

5. Summary and Conclusions

We investigated the masses and radii of planets resulting from coupled formation and evolution modeling under the assumption that water can be abundant on sub-Neptune and super Earth mass exoplanets. The underlying formation model relies on planetesimal accretion and includes both initially rocky and water-rich planets in the observable regime. The large water fraction on planets with masses ranging from Earth mass to tens of Earth masses is a prediction of planet formation models including orbital migration and was in a prior work found to match reasonably well the radius distribution of observed exoplanets around Solar-type stars (Burn et al. 2024). Here, we followed-up on this work and explore the mass-radius relations of these planets. Furthermore, we slightly revised the initial hydrogen and helium content and extend the simulations to lower stellar masses and discussed trends with stellar mass. In Burn et al. (2024), the water was mixed with H/He in the gaseous envelope or alternatively segregated from H/He and layered on top of the rocky core. Here, we additionally explore the effect of preferential removal of hydrogen over oxygen (fractionation) and sequestration of water into the interior of the planet.

The following conclusions can be drawn:

– We found a qualitatively good match of the observational data if water is mixed with H/He in the envelope and is not allowed to dissolve in the interior. In contrast, planets with layered structures do not match observed mass-radius relations.

– For the planets with mixed envelope composition, the location of the radius valley is increasing for stellar masses ranging from 0.1 to $0.5\,\mathrm{M_\odot}$ before remaining constant. The observed slope with orbital period is reproduced within errors. The locus of the radius valley is found to be at slightly too large radii which did not occur for the initial conditions with less H/He used in Burn et al. (2024). However, we recover the trend of longer minimal orbital periods for the sub-Neptunes compared to super Earths, which was not found in Burn et al. (2024), thanks to enforcing the presence of H/He on all planets at the end of the disk lifetime. Another iteration of modeling is required to match both aspects.

– The obtained mass-radius relation is in general agreement with the coupled formation-evolution modeling by Venturini et al. (2024). The differences are due to different orbital period distributions and applied limits as well as due to their pebble-accretion-informed initial conditions segregating more distinctly water-rich from rocky planets compared to the prediction of planetesimal accretion used here. As in that work, the lowest-mass volatile-rich planets are predicted at lower irradiation levels, while a cut at fixed irradiation predicts a stellar-mass independent emergence of the lowest-mass sub-Neptune.

– The quantification of mass-radius characteristics suggests that a mechanism should act during formation which moderately lowers the initial water content compared to standard formation modeling (e.g. atmospheric recycling with the gaseous disk or ^{26}Al heating and water-loss on planetesimals). Further, a reduction of the number of rocky planet building blocks is motivated. Combined, this would allow for a better match with observations by steepening the mass-radius relation of volatile-rich planets, and increasing the ratio of sub-Neptune to super Earths, which was found to be lower than found by Kepler in the nominal model. However, this conclusion relies on a heterogeneous sample of exoplanetary masses and radii and the comparison of the same model to Kepler statistics would rather indicate less mass loss (or higher metallicities). Overall, this strongly motivates the construction of a homogeneous sample of well characterized planets or occurrence rates in mass-radius-orbital period space, which is being pursued by the THIRSTEE project (Lacedelli et al. 2024).

– Mass loss driven by photoevaporation is found to be mostly equally removing hydrogen and oxygen for the modeled planets because all planets with surviving volatile budgets initially contain large envelopes which require large mass loss rates. This implies, that fractionation is not necessarily occurring on sub-Neptunes and the absence of enriched (e.g. He-rich) planets cannot be used as a probe to answer the nature versus nurture question on whether evolution or formation determines the present-day exoplanet properties. Instead, occurrence or absence of fractionation in nature constrains the initial volatile budget of the planet.

– Sequestration of water into the interior has a large effect on the resulting mass-radius relation and should therefore be considered in future studies. The resulting mass-radius relation includes less low-density sub-Neptunes in better agreement with observations compared to the model without any interior-atmosphere coupling. However, the model with sequestration predicts a population of planets without significant primary envelopes but large water abundances. These planets would have densities similar to water-worlds which do not seem to be present in significant numbers in nature (Parviainen et al. 2024). This might indicate that less than maximum sequestration or earlier outgassing occurs. However, the model used here is likely too simplistic for a conclusive assessment and we suggest to include time-dependent solidification of the assumed magma ocean and comprehensive tracking of other chemical species and elements.

References

Acuña, L., Deleuil, M., Mousis, O., et al. 2021, A&A, 647, A53
Affolter, L., Mordasini, C., Oza, A. V., Kubyshkina, D., & Fossati, L. 2023, A&A, 676, A119
Aguichine, A., Mousis, O., Deleuil, M., & Marcq, E. 2021, ApJ, 914, 84
Alibert, Y. & Benz, W. 2017, A&A, 598, L5
Alibert, Y., Mordasini, C., Benz, W., & Winisdoerffer, C. 2005, A&A, 434, 343
Andrews, S. M., Terrell, M., Tripathi, A., et al. 2018, ApJ, 865, 157
Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., & Dullemond, C. P. 2010, ApJ, 723, 1241
Astudillo-Defru, N., Cloutier, R., Wang, S. X., et al. 2020, A&A, 636, A58
Bajgain, S., Ghosh, D. B., & Karki, B. B. 2015, Nature Communications, 6, 8578
Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, A&A, 577, A42
Beatty, T. G., Welbanks, L., Schlawin, E., et al. 2024, ApJ, 970, L10
Benneke, B., Roy, P.-A., Coulombe, L.-P., et al. 2024ArXiv e-prints
Berger, T. A., Huber, D., Gaidos, E., van Saders, J. L., & Weiss, L. M. 2020, AJ, 160, 108
Bodenheimer, P. & Pollack, J. B. 1986, Icarus, 67, 391
Bonfanti, A., Brady, M., Wilson, T. G., et al. 2024, A&A, 682, A66
Bonomo, A. S., Dumusque, X., Massa, A., et al. 2023, A&A, 677, A33
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977
Boukrouche, R., Lichtenberg, T., & Pierrehumbert, R. T. 2021, ApJ, 919, 130
Bower, D. J., Hakim, K., Sossi, P. A., & Sanan, P. 2022, The Planetary Science Journal, 3, 93
Bower, D. J., Sanan, P., & Wolf, A. S. 2018, Physics of the Earth and Planetary Interiors, 274, 49
Brinkman, C. L., Weiss, L. M., Dai, F., et al. 2023, AJ, 165, 88
Burn, R. & Mordasini, C. 2024, in Handbook of Exoplanets, 2nd edn., ed. H. Deeg, J. A. Belmonte, & R. E. Pudritz (Cham: Springer International Publishing)
Burn, R., Mordasini, C., Mishra, L., et al. 2024, Nature Astronomy, 8, 463
Burn, R., Schlecker, M., Mordasini, C., et al. 2021, A&A, 656, A72
Chabrier, G., Mazevet, S., & Soubiran, F. 2019, ApJ, 872, 51
Chachan, Y. & Stevenson, D. J. 2018, ApJ, 854, 21
Chakrabarty, A. & Mulders, G. D. 2024, ApJ, 966, 185
Chambers, J. E. 1999, MNRAS, 304, 793
Charnay, B., Meadows, V., & Leconte, J. 2015, ApJ, 813, 15
Chassefière, E. 1996, Icarus, 124, 537
Chen, D.-C., Mordasini, C., Xie, J.-W., Zhou, J.-L., & Emsenhuber, A. 2024, A&A, 687, A25
Chen, H. & Rogers, L. A. 2016, ApJ, 831, 180
Cherubim, C., Wordsworth, R., Hu, R., & Shkolnik, E. 2024, ApJ, 967, 139
Cloutier, R., Greklek-McKeon, M., Wurmser, S., et al. 2024, MNRAS, 527, 5464
Cloutier, R. & Menou, K. 2020, AJ, 159, 211
Coleman, G. A. L., Leleu, A., Alibert, Y., & Benz, W. 2019, A&A, 631, A7
Connolly, J. A. D. 2009, Geochemistry, Geophysics, Geosystems, 10, Q10014
Coppari, F., Smith, R. F., Wang, J., et al. 2021, Nature Geoscience, 14, 121
Cumming, A., Marcy, G. W., & Butler, R. P. 1999, ApJ, 526, 890
Damasso, M., Rodrigues, J., Castro-González, A., et al. 2023, A&A, 679, A33
Dittkrist, K.-M., Mordasini, C., Klahr, H., Alibert, Y., & Henning, T. 2014, A&A, 567, A121
Dorn, C., Khan, A., Heng, K., et al. 2015, A&A, 577, A83
Dorn, C. & Lichtenberg, T. 2021, ApJ, 922, L4
Dorn, C., Mosegaard, K., Grimm, S. L., & Alibert, Y. 2018, ApJ, 865, 20
Dorn, C., Venturini, J., Khan, A., et al. 2017, A&A, 597, A37
Emsenhuber, A., Burn, R., Weder, J., et al. 2023a, A&A, 673, A78
Emsenhuber, A., Mordasini, C., & Burn, R. 2023b, The European Physical Journal Plus, 138, 181
Emsenhuber, A., Mordasini, C., Burn, R., et al. 2021, A&A, 656, A69
Faik, S., Tauschwitz, A., & Iosilevskiy, I. 2018, Computer Physics Communications, 227, 117
Fischer, R. A., Campbell, A. J., Shofner, G. A., et al. 2011, Earth and Planetary Science Letters, 304, 496
Fortier, A., Alibert, Y., Carron, F., Benz, W., & Dittkrist, K.-M. 2013, A&A, 549, A44
Fulton, B. J. & Petigura, E. A. 2018, AJ, 156, 264
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
Gaidos, E., Ali, A., Kraus, A. L., & Rowe, J. F. 2024, MNRAS, 534, 3277
Giacalone, S., Dressing, C. D., Hedges, C., et al. 2022, AJ, 163, 99
Goffo, E., Gandolfi, D., Egger, J. A., et al. 2023, ApJ, 955, L3
Gronoff, G., Arras, P., Baraka, S., et al. 2020, Journal of Geophysical Research (Space Physics), 125, e27639
Gu, P.-G. & Chen, H. 2023, The Astrophysical Journal Letters, 953, L27
Guillot, T. 2010, A&A, 520, A27
Guo, J. H. 2019, ApJ, 872, 99
Gupta, A. & Schlichting, H. E. 2020, MNRAS, 493, 792
Hakim, K., Rivoldini, A., Van Hoolst, T., et al. 2018, Icarus, 313, 61
Haldemann, J., Alibert, Y., Mordasini, C., & Benz, W. 2020, A&A, 643, A105
Hamano, K., Abe, Y., & Genda, H. 2013, Nature, 497, 607

Hemley, R. J., Stixrude, L., Fei, Y., & Mao, H. K. 1992, Geophysical Monograph Series, 67, 183

Hirschmann, M. M., Withers, A. C., Ardia, P., & Foley, N. T. 2012, Earth and Planetary Science Letters, 345, 38

Ho, C. S. K., Rogers, J. G., Van Eylen, V., Owen, J. E., & Schlichting, H. E. 2024, MNRAS, 531, 3698

Ho, C. S. K. & Van Eylen, V. 2023, MNRAS, 519, 4056

Hollenbach, D., Johnstone, D., Lizano, S., & Shu, F. 1994, ApJ, 428, 654

Holmberg, M. & Madhusudhan, N. 2024, A&A, 683, L2

Howell, S. B., Sobeck, C., Haas, M., et al. 2014, Publications of the Astronomical Society of the Pacific, 126, 398

Hu, R., Seager, S., & Yung, Y. L. 2015, ApJ, 807, 8

Hunten, D. M., Pepin, R. O., & Walker, J. C. G. 1987, Icarus, 69, 532

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90

Ida, S. & Lin, D. N. C. 2008, ApJ, 673, 487

Ikoma, M., Elkins-Tanton, L., Hamano, K., & Suckale, J. 2018, Space Science Reviews, 214, 76

Inaba, S. & Ikoma, M. 2003, A&A, 410, 711

Inaba, S., Tanaka, H., Nakazawa, K., Wetherill, G. W., & Kokubo, E. 2001, Icarus, 149, 235

Inamdar, N. K. & Schlichting, H. E. 2015, MNRAS, 448, 1751

Izidoro, A., Schlichting, H. E., Isella, A., et al. 2022, ApJ, 939, L19

Jin, S. & Mordasini, C. 2018, ApJ, 853, 163

Jin, S., Mordasini, C., Parmentier, V., et al. 2014, ApJ, 795, 65

Johnstone, C. P. 2020, ApJ, 890, 79

Katz, R. F., Spiegelman, M., & Langmuir, C. H. 2003, Geochemistry, Geophysics, Geosystems, 4, 1073

Kempton, E. M. R. & Knutson, H. A. 2024, Reviews in Mineralogy and Geochemistry, 90, 411

Kempton, E. M. R., Zhang, M., Bean, J. L., et al. 2023, Nature, 620, 67

Kippenhahn, R., Weigert, A., & Weiss, A. 2012, Stellar Structure and Evolution (Berlin, Heidelberg: Springer)

Kite, E. S., Fegley, Jr., B., Schaefer, L., & Ford, E. B. 2019, ApJ, 887, L33

Kite, E. S., Fegley, Jr., B., Schaefer, L., & Ford, E. B. 2020, ApJ, 891, 111

Kite, E. S. & Schaefer, L. 2021, ApJ, 909, L22

Komacek, T. D., Showman, A. P., & Parmentier, V. 2019, ApJ, 881, 152

Kopparapu, R. K., Ramirez, R. M., SchottelKotte, J., et al. 2014, ApJ, 787, L29

Kossakowski, D., Kemmer, J., Bluhm, P., et al. 2021, A&A, 656, A124

Kubyshkina, D., Fossati, L., Erkaev, N. V., et al. 2018, A&A, 619, A151

Kubyshkina, D. I. & Fossati, L. 2021, Research Notes of the American Astronomical Society, 5, 74

Kunimoto, M., Winn, J., Ricker, G. R., & Vanderspek, R. K. 2022, AJ, 163, 290

Lacedelli, G., Pallè, E., Luque, R., et al. 2024, ArXiv e-prints

Lammer, H., Scherf, M., Kurokawa, H., et al. 2020, Space Science Reviews, 216, 74

Leleu, A., Delisle, J.-B., Burn, R., et al. 2024a, A&A, 687, L1

Leleu, A., Delisle, J. B., Delrez, L., et al. 2024b, A&A, 688, A211

Lichtenberg, T. 2021, ApJ, 914, L4

Lichtenberg, T., Golabek, G. J., Burn, R., et al. 2019, Nature Astronomy, 3, 307

Lichtenberg, T., Schaefer, L. K., Nakajima, M., & Fischer, R. A. 2023, in Protostars and Planets VII, Astronomical Society of the Pacific Conference Series, ed. Inutsuka, S. and Aikawa, Y. and Muto, T. and Tomida, K. and Tamura, M., Vol. 534, Kyoto, 907

Lin, D. N. C., Bodenheimer, P., & Richardson, D. C. 1996, Nature, 380, 606

Linder, E. F., Mordasini, C., Mollière, P., et al. 2019, A&A, 623, A85

Lissauer, J. J., Batalha, N. M., & Borucki, W. J. 2023, in Protostars and Planets VII, Astronomical Society of the Pacific Conference Series, ed. Inutsuka, S. and Aikawa, Y. and Muto, T. and Tomida, K. and Tamura, M., Vol. 534, Kyoto, 839

Lopez, E. D. 2017, MNRAS, 472, 245

Lopez, E. D. & Fortney, J. J. 2013, ApJ, 776, 2

Loyd, R. O. P., Shkolnik, E. L., Schneider, A. C., et al. 2020, ApJ, 890, 23

Luo, H., Dorn, C., & Deng, J. 2024, Majority of Water Hides Deep in the Interiors of Exoplanets

Luque, R., Fulton, B. J., Kunimoto, M., et al. 2022, A&A, 664, A199

Luque, R., Osborn, H. P., Leleu, A., et al. 2023, Nature, 623, 932

Luque, R. & Pallé, E. 2022, Science, 377, 1211

MacDougall, M. G., Gilbert, G. J., & Petigura, E. A. 2023, AJ, 166, 61

Makino, J., Fukushige, T., Funato, Y., & Kokubo, E. 1998, New Astronomy, 3, 411

Malsky, I., Rogers, L., Kempton, E. M.-R., & Marounina, N. 2023, Nature Astronomy, 7, 57

Malsky, I. & Rogers, L. A. 2020, ApJ, 896, 48

Marboeuf, U., Thiabaud, A., Alibert, Y., Cabral, N., & Benz, W. 2014a, A&A, 570, A36

Marboeuf, U., Thiabaud, A., Alibert, Y., Cabral, N., & Benz, W. 2014b, A&A, 570, A35

Mason, E. A. & Marrero, T. R. 1970, in Advances in Atomic and Molecular Physics, ed. D. R. Bates & I. Esternnan, Vol. 6 (Academic Press), 155–232

McDonald, G. D., Kreidberg, L., & Lopez, E. 2019, ApJ, 876, 22

Miguel, Y., Cridland, A., Ormel, C., Fortney, J., & Ida, S. 2020, MNRAS, 491, 1998

Mishra, L., Alibert, Y., Leleu, A., et al. 2021, A&A, 656, A74

Moldenhauer, T. W., Kuiper, R., Kley, W., & Ormel, C. W. 2022, A&A, 661, A142

Mordasini, C. 2014, A&A, 572, A118

Mordasini, C. 2020, A&A, 638, A52

Mordasini, C., Alibert, Y., & Benz, W. 2009, A&A, 501, 1139

Mordasini, C., Alibert, Y., Georgy, C., et al. 2012a, A&A, 547, A112

Mordasini, C., Alibert, Y., Klahr, H., & Henning, T. 2012b, A&A, 547, A111

Mordasini, C. & Burn, R. 2024, Reviews in Mineralogy and Geochemistry, 90, 55

Mordasini, C., van Boekel, R., Mollière, P., Henning, Th., & Benneke, B. 2016, ApJ, 832, 41

Mousis, O., Deleuil, M., Aguichine, A., et al. 2020, ApJ, 896, L22

Murray-Clay, R. A., Chiang, E. I., & Murray, N. 2009, ApJ, 693, 23

Musella, R., Mazevet, S., & Guyot, F. 2019, Physical Review B, 99, 064110

Nixon, M. C., Piette, A. A. A., Kempton, E. M. R., et al. 2024, ApJ, 970, L28

Odert, P., Lammer, H., Erkaev, N. V., et al. 2018, Icarus, 307, 327

Orell-Miquel, J., Nowak, G., Murgas, F., et al. 2023, A&A, 669, A40

Osborn, A., Armstrong, D. J., Cale, B., et al. 2021, MNRAS, 507, 2782

Owen, J. E. & Wu, Y. 2013, ApJ, 775, 105

Paardekooper, S., Dong, R., Duffell, P., et al. 2023, in ASP Conference Series, Vol. 534, Protostars and Planets VII, Astronomical Society of the Pacific Conference Series, ed. Inutsuka, S. and Aikawa, Y. and Muto, T. and Tomida, K. and Tamura, M., Kyoto, 685

Paardekooper, S.-J., Baruteau, C., & Kley, W. 2011, MNRAS, 410, 293

Parc, L., Bouchy, F., Venturini, J., Dorn, C., & Helled, R. 2024, A&A, 688, A59

Parviainen, H., Luque, R., & Palle, E. 2024, MNRAS, 527, 5693

Passegger, V. M., Suárez Mascareño, A., Allart, R., et al. 2024, A&A, 684, A22

Pavlov, A. V. 2019, Surveys in Geophysics, 40, 247

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011Journal of Machine Learning Research, 12, 2825

Petigura, E. A., Marcy, G. W., Winn, J. N., et al. 2018, AJ, 155, 89

Petigura, E. A., Rogers, J. G., Isaacson, H., et al. 2022, AJ, 163, 179

Piaulet-Ghorayeb, C., Benneke, B., Radica, M., et al. 2024, JWST/NIRISS Reveals the Water-Rich "Steam World" Atmosphere of GJ 9827 d

Pierrehumbert, R. T. 2023, ApJ, 944, 20

Richert, A. J. W., Getman, K. V., Feigelson, E. D., et al. 2018, MNRAS, 477, 5191

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Rogers, J. G., Dorn, C., Raj, V. A., Schlichting, H. E., & Young, E. D. 2024, Most Super-Earths Have Less Than 3% Water

Rogers, J. G. & Owen, J. E. 2021, MNRAS, 503, 1526

Rogers, J. G., Schlichting, H. E., & Owen, J. E. 2023, ApJ, 947, L19

Sarkis, P., Mordasini, C., Henning, T., Marleau, G. D., & Mollière, P. 2021, A&A, 645, A79

Schlecker, M., Burn, R., Sabotta, S., et al. 2022, A&A, 664, A180

Schlichting, H. E. & Young, E. D. 2022, The Planetary Science Journal, 3, 127

Schulik, M. & Booth, R. A. 2023, MNRAS, 523, 286

Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007, ApJ, 669, 1279

Shah, O., Alibert, Y., Helled, R., & Mezger, K. 2021, A&A, 646, A162

Silverberg, S. M., Wisniewski, J. P., Kuchner, M. J., et al. 2020, ApJ, 890, 106

Sossi, P. A., Tollan, P. M. E., Badro, J., & Bower, D. J. 2023, Earth and Planetary Science Letters, 601, 117894

Stixrude, L. 2014, Philosophical Transactions of the Royal Society of London Series A, 372, 20130076

Stixrude, L. & Lithgow-Bertelloni, C. 2022, Geophysical Journal International, 228, 1119

Thiabaud, A., Marboeuf, U., Alibert, Y., et al. 2014, A&A, 562, A27

Turbet, M., Bolmont, E., Ehrenreich, D., et al. 2020, A&A, 638, A41

Tychoniec, Ł., Tobin, J. J., Karska, A., et al. 2018, The Astrophysical Journal Supplement Series, 238, 19

Ulmer-Moll, S., Osborn, H. P., Tuson, A., et al. 2023, A&A, 674, A43

Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2018, MNRAS, 479, 4786

Van Eylen, V., Astudillo-Defru, N., Bonfils, X., et al. 2021, MNRAS, 507, 2154

Vazan, A., Sari, R., & Kessel, R. 2022, ApJ, 926, 150

Venturini, J., Guilera, O. M., Haldemann, J., Ronco, M. P., & Mordasini, C. 2020, A&A, 643, L1

Venturini, J., Ronco, M. P., Guilera, O. M., et al. 2024, A&A, 686, L9

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Wallack, N. L., Batalha, N. E., Alderson, L., et al. 2024, JWST COMPASS: A NIRSpec/G395H Transmission Spectrum of the Sub-Neptune TOI-836c

Wang, Y., Ormel, C. W., Huang, P., & Kuiper, R. 2023, MNRAS, 523, 6186

Ward, W. R. 1997, Icarus, 126, 261

Yoshida, T., Terada, N., Ikoma, M., & Kuramoto, K. 2022, ApJ, 934, 137

Zahnle, K., Kasting, J. F., & Pollack, J. B. 1990, Icarus, 84, 502

Zahnle, K. J. & Kasting, J. F. 1986, Icarus, 68, 462

Zeng, L., Jacobsen, S. B., Sasselov, D. D., et al. 2019, Proceedings of the National Academy of Sciences, 116, 9723

Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, ApJ, 819, 127

Appendix A: An analytic model for fractionation

For large escape rates, the analytic estimate by Hunten et al. (1987) should apply (Zahnle et al. 1990). They derived a fractionation factor

$$x_O = 1 - \frac{(m_O - m_H)\, b_{H,O}\, GM}{\Phi_H R^2 k_B T_{iso}\left(1 + \frac{n_O(R)}{n_H(R)}\right)}, \tag{A.1}$$

where M is the mass of the planet, T_{iso} is the temperature in the atmosphere assumed in this model to be isothermal, $b_{H,O} = 4.8 \times 10^{17}\,\mathrm{cm^{-1}\,s^{-1}}\,(T_{iso}/1\,\mathrm{K})^{0.75}$ is the binary diffusion coefficient (Mason & Marrero 1970; Zahnle & Kasting 1986; Pavlov 2019), R is the effective planetary radius, and Φ_H (respectively Φ_O) is the escaping flux of hydrogen (oxygen) particles per surface such that the total particle loss $\dot{N}_H = -4\pi R^2 \Phi_H$ ($\dot{N}_O = -4\pi R^2 \Phi_O$).

The oxygen fractionation factor is defined to relate to the fluxes as

$$x_O \equiv \frac{\Phi_O}{\Phi_H}\frac{N_H}{N_O}, \tag{A.2}$$

where the total number of particles of both species can be obtained from the total masses used in the rest of the evolution model $N_O = M_O/m_O$ and $N_H = M_H/m_H$. The number density ratio n_O/n_H at radius R is more challenging to obtain as discussed below. This factor is relevant for major abundances of the heavier oxygen and whenever the mass loss rate Φ_H is relatively low.

After some algebra and relating these to the total mass loss $\dot{M}_{evap} = \dot{M}_H + \dot{M}_O$, the expression can be written as

$$\dot{M}_O = \frac{M_O}{M}\left(\dot{M}_{evap} - \frac{4\pi m_H(m_O - m_H)b_{H,O}GM}{k_B T_{iso}\left(1 + \frac{n_O(R)}{n_H(R)}\right)}\right) \tag{A.3}$$

implying (given that the amount of water is determined by the available oxygen and never limited by hydrogen availability) a water loss of

$$\dot{M}_{H_2O} = \dot{M}_O \frac{m_{H_2O}}{m_O}. \tag{A.4}$$

To investigate the effect of different estimates of n_O/n_H, we assume uniform mixing $n_O(R)/n_H(R) = N_O/N_H$ and $T_{iso} = 3000\,\mathrm{K}$ in one case (lower limit for fractionation) and an estimate assuming isothermal atmospheres and no mixing as detailed in the following to give a more optimistic estimate for fractionation.

For the optimistic fractionation model, we determine the base of the flow analogous to literature on disk photoevaporation (Hollenbach et al. 1994) by requiring that the mass flux launched as a decoupled wind at the absorption layer of high-energy radiation (Murray-Clay et al. 2009) is sustained from a deeper 'base' radius R_{base} of the flux. This can be written as

$$\dot{M}_{evap} = c_{s,ions} 4\pi R_{base}^2 \rho_{base}, \tag{A.5}$$

where $c_{s,ions} = \frac{k_B 3000\,\mathrm{K}}{\bar{m}}$, with $\bar{m} = m_H\,(X/0.5 + Y/2 + Z/3)^{-1}$ for mass fractions of hydrogen X, helium Y and heavy elements assumed to be water Z. This averaging assumes that particles are singly ionized. For example H_2O is present as $2\times H^+$, O^+, and 3 electrons with a total weight of $18\,\mathrm{u}$ distributed to 6 particles, thus an average weight of $3\,\mathrm{u}$ per particle results. The sound speed is an order-of-magnitude estimate of average velocity with which the ionized particles in the wind are launched. The base

radius can be located outside of the resolved structure of the envelope which sometimes requires extrapolation. Then, the number density ratio $\frac{n_O(R_{base})}{n_H(R_{base})}$ is estimated assuming individual hydrostatic distributions for O and H with scale heights $H_{O,H} = \frac{k_B T}{m_{O,H}}$. This implies no interaction between O and H, thus a lower limit for $\frac{n_O}{n_H}$, respectively an upper limit for the fractionation effect. More detailed considerations should include different species (Odert et al. 2018), the effect of turbulent mixing (e.g. Charnay et al. 2015; Komacek et al. 2019), ionization and the resulting interactions with ions (Hu et al. 2015; Guo 2019), and the deviations of detailed numerical studies from the analytical approach (Johnstone 2020; Schulik & Booth 2023).

Appendix B: Iron core and mantle densities including volatiles

The Water Sequestration model presented here simplifies the treatment of the interior structure to determine the radius of the iron core and magma ocean by assuming a water layer on top of the rocky material instead of consistently mixing water with mantle and core. In this section, we aim at comparing the radii of the planets obtained in this way with a more consistent model based on Dorn et al. (2017).

This model focuses on Earth-like rocky interiors with steam atmospheres. Water can exist in the core, mantle, or surface, depending on the thermal state of the planet. The core is composed of iron (Fe) with light elements such as hydrogen (H) and oxygen (O). For solid Fe, we use the equations of state for hexagonal close-packed iron from Hakim et al. (2018), and for liquid iron and its alloys, we follow Luo et al. (2024) who use a Mie-Grüneisen equation of state with parameters derived from ab-initio calculations. Luo et al. (2024) state that their approach is valid in the super-Earth to sub-Neptune regime, that is for high pressures above $50\,\mathrm{GPa}$. The core is assumed to be adiabatic, but there is a temperature jump at the core-mantle boundary (CMB) due to residual heat from core formation, based on Stixrude (2014).

The mantle consists of three primary components: MgO, SiO_2, and FeO, assuming an Earth-like composition. For solid mantle properties, we employ the Perple_X thermodynamic model Connolly (2009) and the database of Stixrude & Lithgow-Bertelloni (2022). For pressures above $125\,\mathrm{GPa}$, stable minerals are defined a priori using equations of state from various sources (Fischer et al. 2011; Faik et al. 2018; Hemley et al. 1992; Musella et al. 2019; Coppari et al. 2021). The liquid mantle's density is computed as an ideal mixture of Mg_2SiO_4, SiO_2, and FeO, using Mg_2SiO_4 instead of MgO due to the recent data for forsterite being updated for high-pressure regimes.

The mantle is considered fully adiabatic, with water present only in the melts, while solid mantle is assumed to be dry. Water decreases the mantle's density following Bajgain et al. (2015), and this reduction is nearly pressure and temperature independent for small water fractions. The mantle melting curve is calculated for pure $MgSiO_3$, and the addition of water (Katz et al. 2003) and iron (Dorn et al. 2018) lowers the melting temperature. Water in the core also reduces the melting temperature, as modeled by Luo et al. (2024). The water which is not dissolved in the mantle (see Sect. 2.2.1) is modeled as being in steam or supercritical phases using the AQUA equation of state (Haldemann et al. 2020), with an isothermal profile below $0.1\,\mathrm{bar}$ that transitions to an adiabatic profile.

Hydrated silicates are not considered motivated by a minor impact on the density (Shah et al. 2021). Instead, water is treated

using the EOS of Haldemann et al. (2020) and the additive volume law to combine it with rock. In this way, interior structures and planetary core and mantle radii can be obtained comparable to the simplified model adopted in this work.

The results of this comparison for the planets which contain some water but without H/He envelopes from our simulations can be seen in Fig. B.1. We see a difference of up to 10% at low mass and low water fraction mainly influenced by the differing assumed rocky mantle material. For substantial water fractions, the radii obtained with the consistent interior model are lower by up to 5% compared to the same planets with a layered structure. We find a compensation of the effect of the varying rocky material density and the treatment of water. Isolated, the assumption that water is layered on top of the mantle instead of mixed leads to a decrease in radius of 15%. We conclude that for our main purpose, the evolution of the planetary atmosphere, this difference is likely not substantial to change the outcome significantly also thanks to the compensation effects. A change in core radius on the order of 5% (15%) leads to a 15% (50%) change in atmospheric mass loss for planets whose radius is dominated by the core. However, for detailed comparison to observations and retrieval of planetary parameters, our radii should not be used.

Appendix C: Supporting figures

This appendix contains figures used for more detailed and higher-dimensional analysis of the synthetic populations presented in the main part of this work.

Appendix D: Supporting Tables

Here, we provide tables of simulated planets in the regime where our assumptions apply, which is considered to be within the runaway greenhouse limit (see Sect. 2.3). The tables provided here do not use an estimate of the observational bias and therefore include every simulated planet once irrespective of detection likelihood. The full tables are available in electronic form.
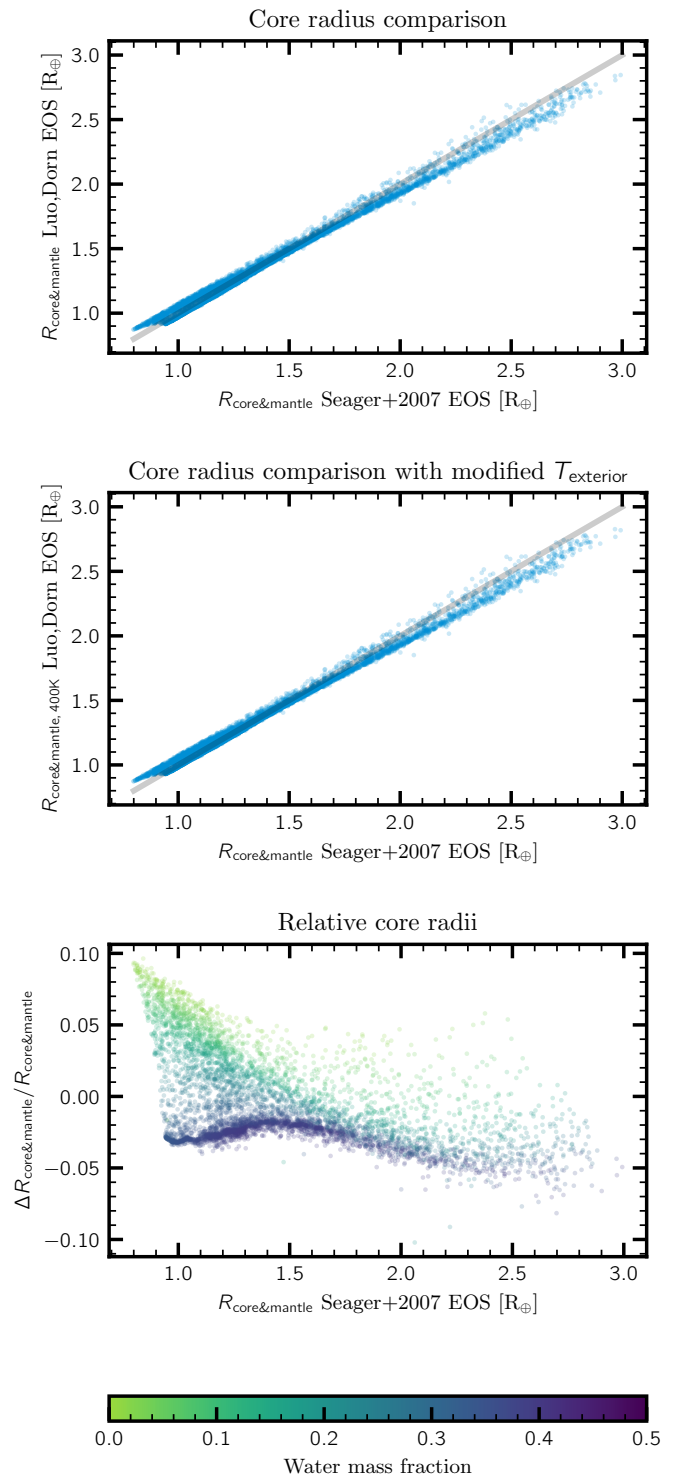

Fig. B.1. Comparison of radii resulting from two different interior structure calculations in the Water Sequestration scenario without outgassing. The adopted interior structure model uses the equation of states (EOS) based on Seager et al. (2007) (see Sect. 2.2.1) and assumes internally that water forms a layer of ice on top of the rocky material. The resulting radii of the solid planetary core and mantle are shown on the horizontal axis. Those from a model based on Luo et al. (2024); Dorn & Lichtenberg (2021), which calculates more consistently the density of materials including water (see text), is shown on the vertical axis (top and middle panel). It assumes an exterior temperature of 3000 K (top panel), respectively 400 K (middle). The bottom panel shows the relative difference between the two modeled radii, with positive values indicating larger radii and total water mass fraction color-coded.
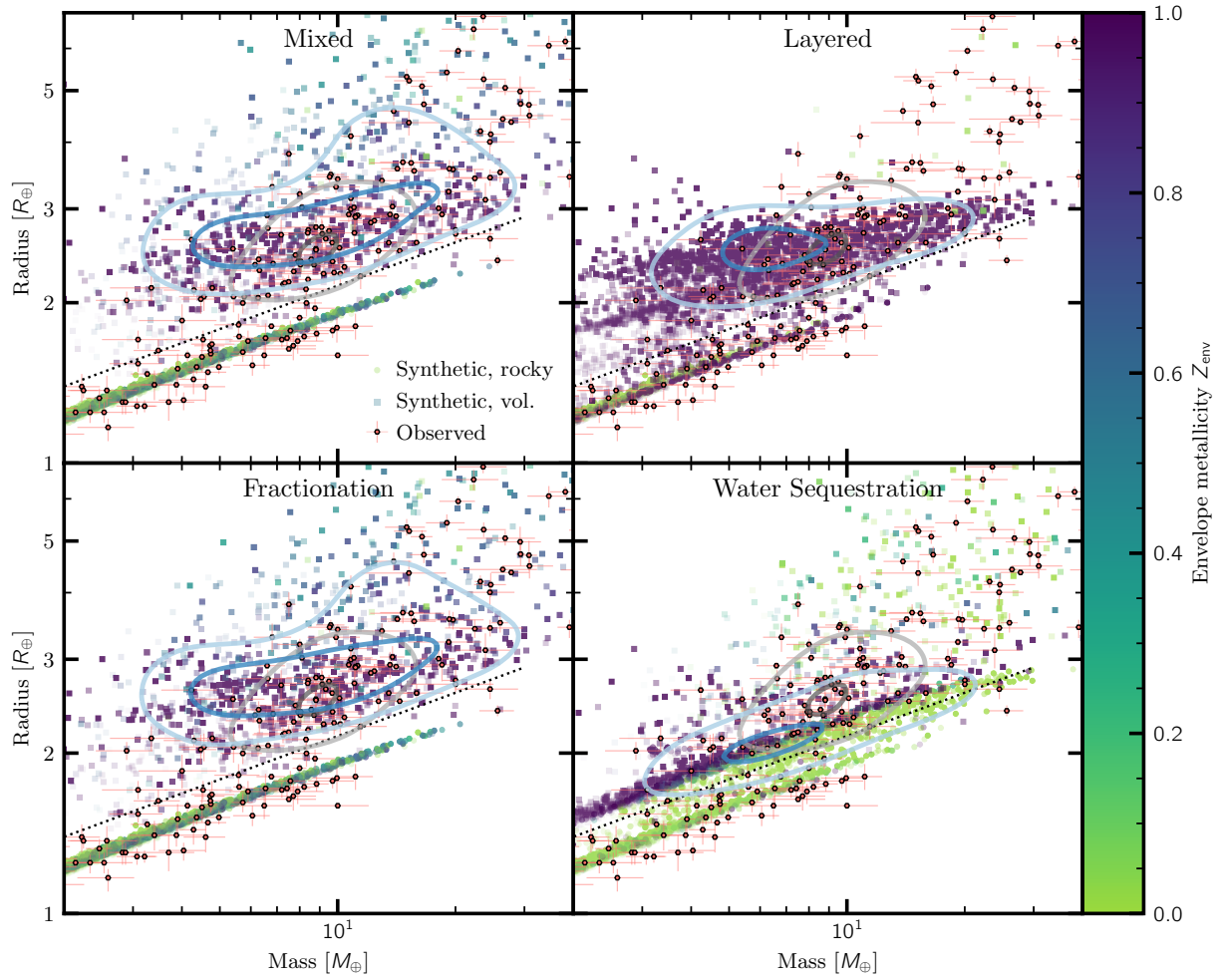
Fig. C.1. Zoom into sub-Neptune region of the mass radius relation. Similar to Fig. 4, but here the envelope metallicity Z_{env} is coded in color after 5 Gyr of evolution. The coloring of rocky planets corresponds to the metallicity of the last gaseous envelope before removal. In addition, the plot contains the contours of a two-dimensional kernel density estimate (KDE Pedregosa et al. 2011) of the probability density of volatile rich planets (bulk densities below 0.65 the expected Earth density, Zeng et al. 2016) for both observed (gray) and synthetic (blue). To account for uneven observation errors, the KDE uses a bandwidth of 0.1 measured in logarithms base 10 and e for masses and radii, respectively.
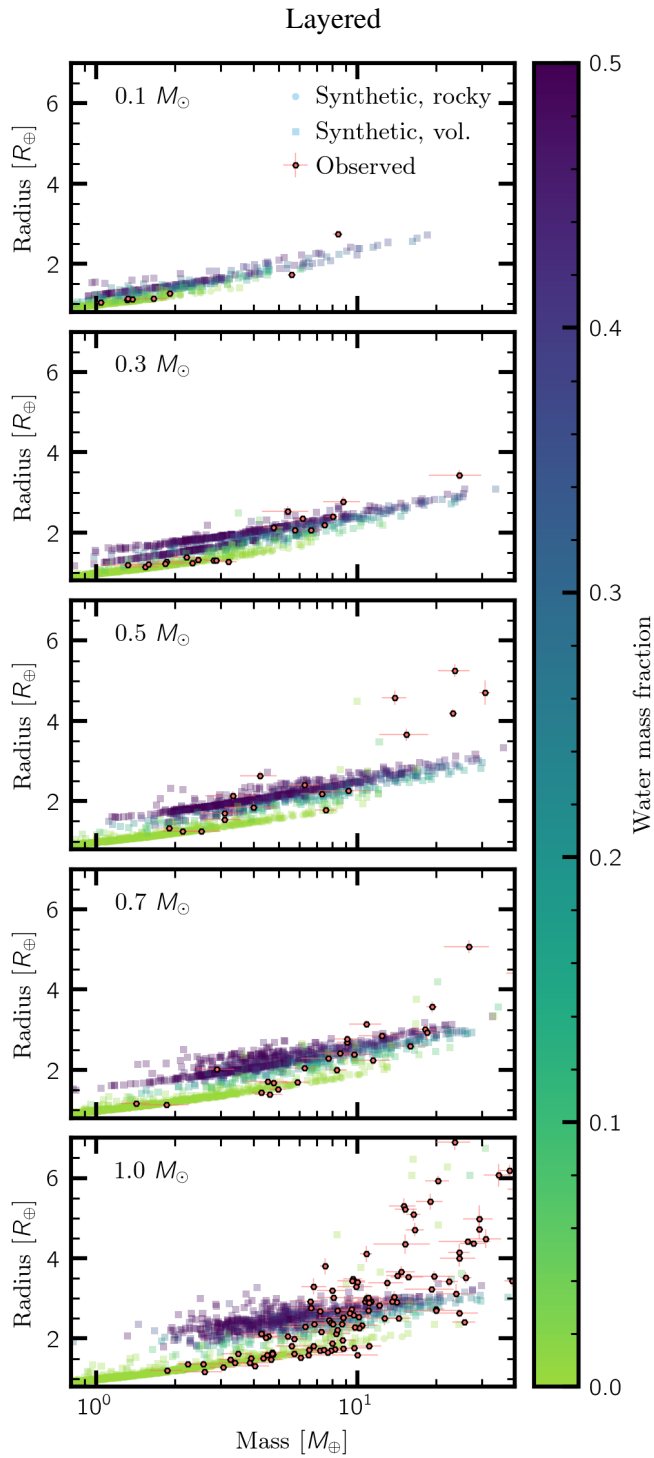
Fig. C.2. As Fig. 8 but showing the output of the Layered model. The bifurcation of water-rich (blue colored) planet densities at low masses, best seen in the 0.3 $M_\odot$ panel, originates from the phase transition of water. The panels here include for the 0.1 and 0.3 $M_\odot$ cases planets exterior to the runaway greenhouse limit where water condenses. With increasing stellar mass, as temperatures increases, so does the scatter in water envelope radii at fixed planetary mass.
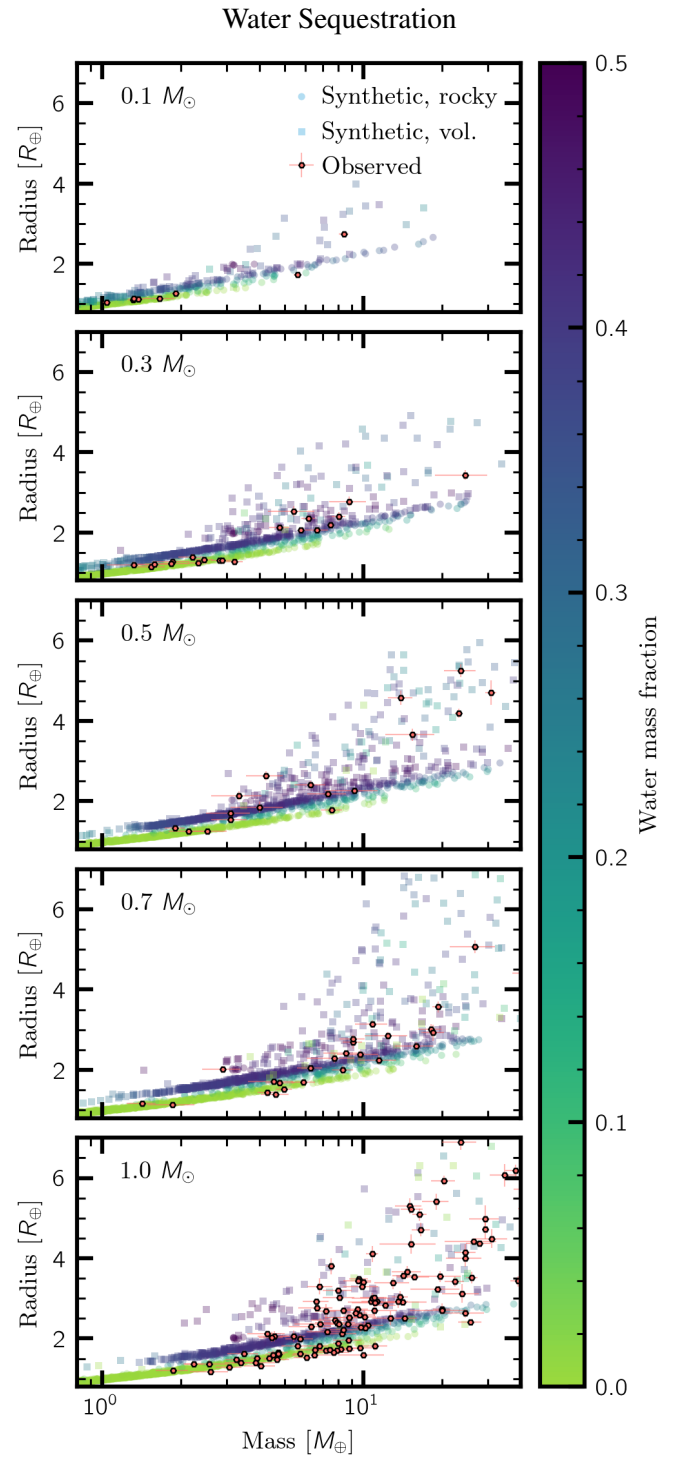
Fig. C.3. As Fig. 8 but showing the output of the Water Sequestration model.

Layered

Water Sequestration
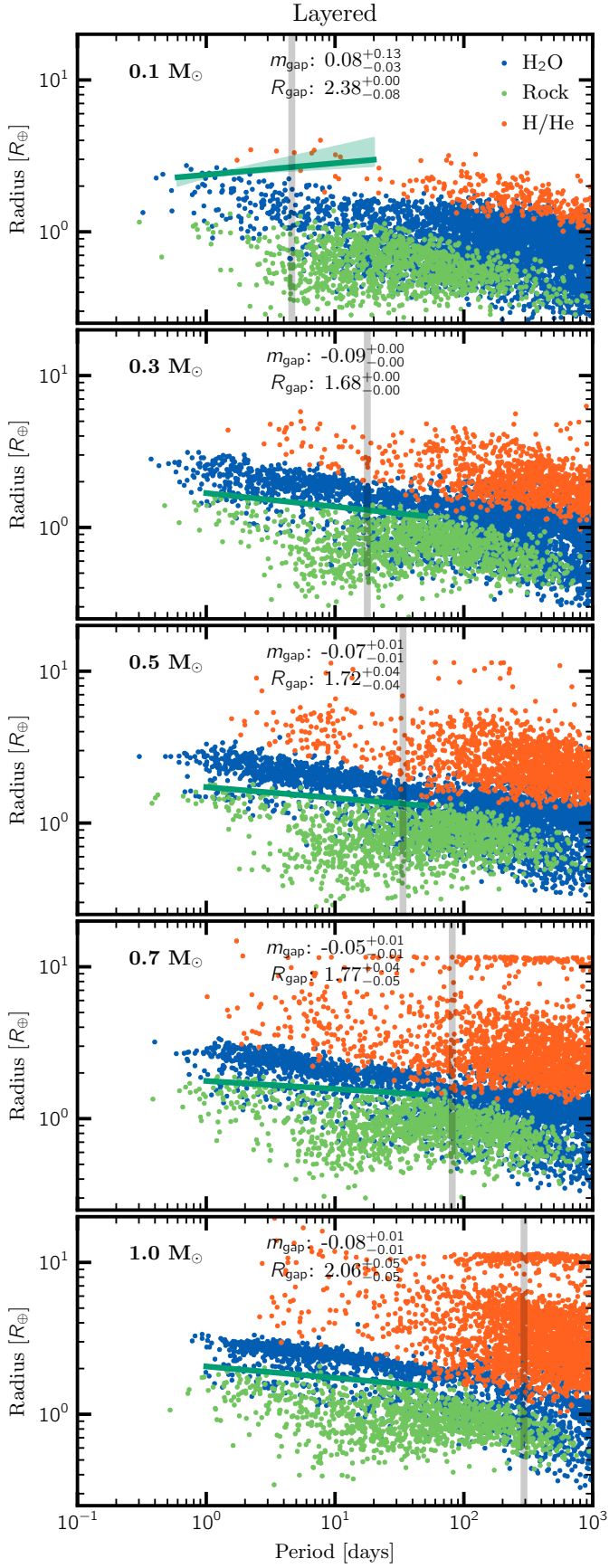
Fig. C.4. Period versus radius of unbiased, synthetic planets using the layered model around stars of different masses. The gap was fitted analogous to the mixed model results shown in Fig. 6. We suggest to not use the 0.1 M$_\odot$ case due to low-number statistics.

Fig. C.5. Period versus radius of unbiased, synthetic planets using the Water Sequestration model around stars of different masses. In this model, some of the rocky planets (green) contain water in their magma ocean or metallic core. The gap was fitted analogous to the mixed model results shown in Fig. 6. We report here but suggest to not use the fitted values since no conclusive radius valley is found.

Table D.1. Simulated planet parameters for the Mixed model

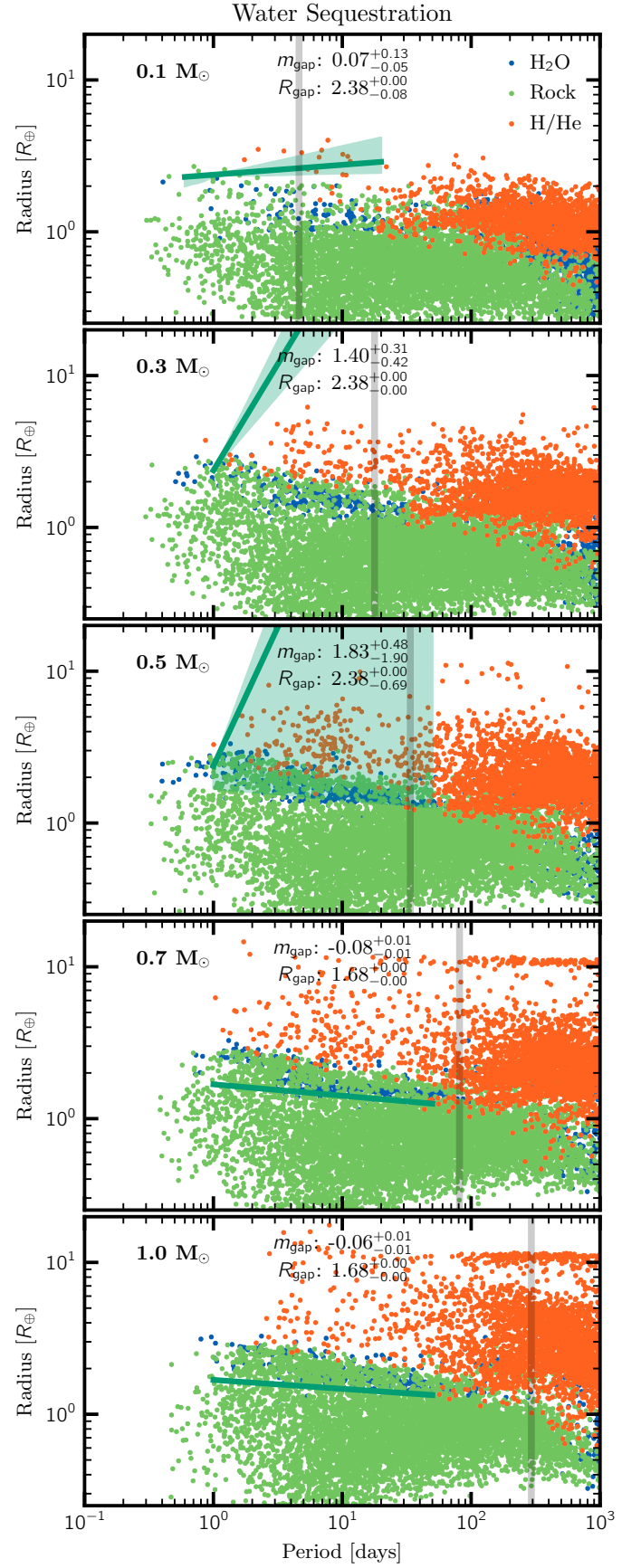| ID	M [M$_\oplus$]	R [R$_\oplus$]	a [au]	P [days]	F [F$_\oplus$]	i	e	M_{env} [M$_\oplus$]	M_{core} [M$_\oplus$]	Z_{env}	$f_{\mathrm{H_2O}}$	sysID	pltID	$M_\star$ [M$_\odot$]	T_{eff} [K]	P_{rgh} [days]
0	0.9870	0.9744	0.0177	2.7170	0.2200	5.24e-7	0.028	0.0	0.9870	0.772	0.0	4	11	0.1	2810	4.600
1	1.7341	1.3626	0.0061	0.5461	1.8685	5.24e-7	0.019	0.0987	1.6354	1.0	0.0569	4	21	0.1	2810	4.600
2	1.3143	1.0589	0.0215	3.6345	0.1493	5.24e-7	0.021	0.0	1.3143	0.698	0.0	4	38	0.1	2810	4.600

Note. The full table with column descriptions is available in electronic form.

Table D.2. Simulated planet parameters for the Layered model

ID	M [M$_\oplus$]	R [R$_\oplus$]	a [au]	P [days]	F [F$_\oplus$]	i	e	M_{env} [M$_\oplus$]	M_{core} [M$_\oplus$]	Z_{env}	$f_{\mathrm{H_2O}}$	sysID	pltID	$M_\star$ [M$_\odot$]	T_{eff} [K]	P_{rgh} [days]
0	1.3249	1.4416	0.0177	2.7170	0.220	5.24e-7	0.028	0.3379	0.9870	0.987	0.2550	4	11	0.1	2810	4.600
1	1.7425	1.3708	0.0061	0.5458	1.870	5.24e-7	0.019	0.1072	1.6354	1.0	0.0	4	21	0.1	2810	4.600
2	2.1690	1.4828	0.0215	3.6344	0.149	5.24e-7	0.021	0.8547	1.3143	1.0	0.3940	4	38	0.1	2810	4.600

Note. The full table with column descriptions is available in electronic form.

Table D.3. Simulated planet parameters for the Fractionation model

ID	M [M$_\oplus$]	R [R$_\oplus$]	a [au]	P [days]	F [F$_\oplus$]	i	e	M_{env} [M$_\oplus$]	M_{core} [M$_\oplus$]	Z_{env}	$f_{\mathrm{H_2O}}$	sysID	pltID	$M_\star$ [M$_\odot$]	T_{eff} [K]	P_{rgh} [days]
0	0.9879	0.9741	0.0177	2.7170	0.2200	5.24e-7	0.028	0.0	0.9870	0.772	0.0	4	11	0.1	2810	4.600
1	1.7423	1.3706	0.0061	0.5460	1.8692	5.24e-7	0.019	0.1070	1.6354	1.0	0.0614	4	21	0.1	2810	4.600
2	1.3143	1.0589	0.0215	3.6345	0.1493	5.24e-7	0.021	0.0	1.3143	0.698	0.0	4	38	0.1	2810	4.600

Note. The full table with column descriptions is available in electronic form.

Table D.4. Simulated planet parameters for the Sequestration model

ID	M [M$_\oplus$]	R [R$_\oplus$]	a [au]	P [days]	F [F$_\oplus$]	i	e	M_{env} [M$_\oplus$]	M_{core} [M$_\oplus$]	Z_{env}	$f_{\mathrm{H_2O}}$	sysID	pltID	$M_\star$ [M$_\odot$]	T_{eff} [K]	P_{rgh} [days]
0	1.2930	1.2621	0.0177	2.7170	0.220	5.24e-7	0.028	0.0001	1.2929	1.0	0.2366	4	11	0.1	2810	4.600
1	1.7425	1.2143	0.0061	0.5474	1.863	5.24e-7	0.019	0.0	1.7425	1.0	0.0615	4	21	0.1	2810	4.600
2	2.0400	1.4427	0.0215	3.6344	0.149	5.24e-7	0.021	0.0	2.0400	1.0	0.3556	4	38	0.1	2810	4.600

Note. The full table with column descriptions is available in electronic form.

Burn et al.: Water-rich sub-Neptunes and rocky super Earths around different stars