



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
Main Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2018

Threshold Radii of Volatile-rich Planets

Lozovsky, M ; Helled, R ; Dorn, C ; Venturini, J

Abstract: Constraining the planetary composition is essential for exoplanetary characterization. In this paper, we use a statistical analysis to determine the characteristic maximum (threshold) radii for various compositions for exoplanets with masses up to 25 Earth masses (M_{\oplus}). We confirm that most planets with radii larger than 1.6 Earth radii (R_{\oplus}) are not rocky, and must consist of lighter elements, as found by previous studies. We find that planets with radii above $2.6 R_{\oplus}$ cannot be pure-water worlds, and must contain significant amounts of hydrogen and helium (H–He). We find that planets with radii larger than about $3 R_{\oplus}$, $3.6 R_{\oplus}$, and $4.3 R_{\oplus}$ are expected to consist of 2%, 5%, and 10% of H–He, respectively. We investigate the sensitivity of the results to the assumed internal structure, the planetary temperature and albedo, and the accuracy of the determination of mass and radius. We show that the envelope’s metallicity, the percentage of H–He, and the distribution of the elements play a significant role in the determination of the threshold radius. Finally, we conclude that, despite the degenerate nature of the problem, it is possible to put limits on the possible range of compositions for planets with well-measured mass and radius.

DOI: <https://doi.org/10.3847/1538-4357/aadd09>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-166681>

Journal Article

Published Version

Originally published at:

Lozovsky, M; Helled, R; Dorn, C; Venturini, J (2018). Threshold Radii of Volatile-rich Planets. *The Astrophysical Journal*, 866(1):13.

DOI: <https://doi.org/10.3847/1538-4357/aadd09>



Threshold Radii of Volatile-rich Planets

M. Lozovsky , R. Helled , C. Dorn, and J. Venturini

Center for Theoretical Astrophysics & Cosmology, Institute for Computational Science, University of Zurich, Zurich, Switzerland; michloz@mail.com

Received 2018 June 25; revised 2018 August 20; accepted 2018 August 23; published 2018 October 10

Abstract

Constraining the planetary composition is essential for exoplanetary characterization. In this paper, we use a statistical analysis to determine the characteristic maximum (threshold) radii for various compositions for exoplanets with masses up to 25 Earth masses (M_{\oplus}). We confirm that most planets with radii larger than 1.6 Earth radii (R_{\oplus}) are not rocky, and must consist of lighter elements, as found by previous studies. We find that planets with radii above $2.6 R_{\oplus}$ cannot be pure-water worlds, and must contain significant amounts of hydrogen and helium (H–He). We find that planets with radii larger than about $3 R_{\oplus}$, $3.6 R_{\oplus}$, and $4.3 R_{\oplus}$ are expected to consist of 2%, 5%, and 10% of H–He, respectively. We investigate the sensitivity of the results to the assumed internal structure, the planetary temperature and albedo, and the accuracy of the determination of mass and radius. We show that the envelope’s metallicity, the percentage of H–He, and the distribution of the elements play a significant role in the determination of the threshold radius. Finally, we conclude that, despite the degenerate nature of the problem, it is possible to put limits on the possible range of compositions for planets with well-measured mass and radius.

Key words: planets and satellites: composition – planets and satellites: detection – planets and satellites: fundamental parameters – planets and satellites: interiors

1. Introduction

The ongoing efforts to detect and characterize exoplanets from Earth and space have led to the detection of thousands of exoplanets, and allow us to study planets as a class of astrophysical objects. Measured radii of planets from the *Kepler* mission combined with radial velocity (RV) follow-ups and transit timing variations (TTV) provide information on the planetary radii and masses, and therefore on their mean densities. The measured masses and radii can be compared to theoretical mass–radius (M–R) relation of planetary objects, which is used to infer the possible bulk composition (e.g., Batygin & Stevenson 2013; Weiss & Marcy 2014; Wolfgang & Lopez 2015; Zeng et al. 2016).

Since the discovery of exoplanets with radii between those of Earth ($1 R_{\oplus}$) and Neptune ($\sim 4 R_{\oplus}$), it has been unclear whether they represent large-scale terrestrial planets (super-Earths) or small versions of Neptune (mini-Neptunes). Characterizing these planets is particularly challenging because there are no similar objects in the solar system, and they lie in a mass regime where uncertainties in the equation of state (EOS) and the material’s distribution are the largest (e.g., Baraffe et al. 2008; Vazan et al. 2016).

Determining the exact planetary structure and composition is challenging due to the intrinsic degeneracy, i.e., exoplanets with very different interiors can have identical masses and radii (e.g., Rogers & Seager 2010; Lopez & Fortney 2014; Dorn et al. 2015, 2017). Despite this inherent degeneracy, the least dense possible interiors for a given bulk composition can be derived. These represent end-member interiors that can be compared to observed exoplanets. For example, the lowest density among all rocky (silicate) interiors is that of MgSiO_3 . Based on the density of the idealized composition of MgSiO_3 , previous studies suggest that most planets with radii larger than $1.6 R_{\oplus}$ have densities too low to be consistent with purely rocky interiors (Weiss & Marcy 2014; Rogers 2015), and therefore they are expected to contain volatiles. More

specifically, Rogers (2015) employed a hierarchical Bayesian statistical approach to determine threshold radii of various rocky compositions. The threshold radius of a given composition represents the radius above which a planet has a very low probability of being of that specific composition. Rogers (2015) used a sample of 22 short-period (up to 50 days) *Kepler* planets with RV follow-ups. A threshold radius of $1.6 R_{\oplus}$ was found for purely rocky exoplanets.

Interestingly, the distribution of observed radii of small exoplanets suggests a bimodal shape of planetary sizes (Fulton et al. 2017; Fulton & Petigura 2018). A gap found at radii of $1.5\text{--}2.0 R_{\oplus}$ splits the population of close-in planets (orbital period shorter than 100 days) into two regimes: planets with $R_p < 1.5 R_{\oplus}$ and planets with $R_p = 2.0\text{--}3.0 R_{\oplus}$. This paucity in the distribution might be explained by photoevaporation of their volatile atmospheres (Lopez & Fortney 2014; Owen & Wu 2017; Van Eylen et al. 2018).

Generally, in volatile-rich planets, the thickness of the gaseous envelope depends on the mass fraction of light elements, the envelope’s metallicity, and the temperature profile of the planet. These parameters and the characteristics of the underlying deeper layers determine the planet’s density. Similarly to purely rocky planets, models of end-member interiors for volatile-rich compositions exist. For example, a planet with a mass fraction of 2% of H–He is expected to have the lowest density when the envelope’s metallicity is low and the temperatures are high. In this paper, we build on the statistical methodology of Rogers (2015) and determine different threshold radii for small and intermediate-size planets, accounting for the possibility of gaseous envelopes with different metallicities and internal structures.

2. Methods

2.1. Exoplanet Data

To date (2018 August) there are more than 3700 confirmed exoplanets. Using the *exoplanet.eu* database, we select

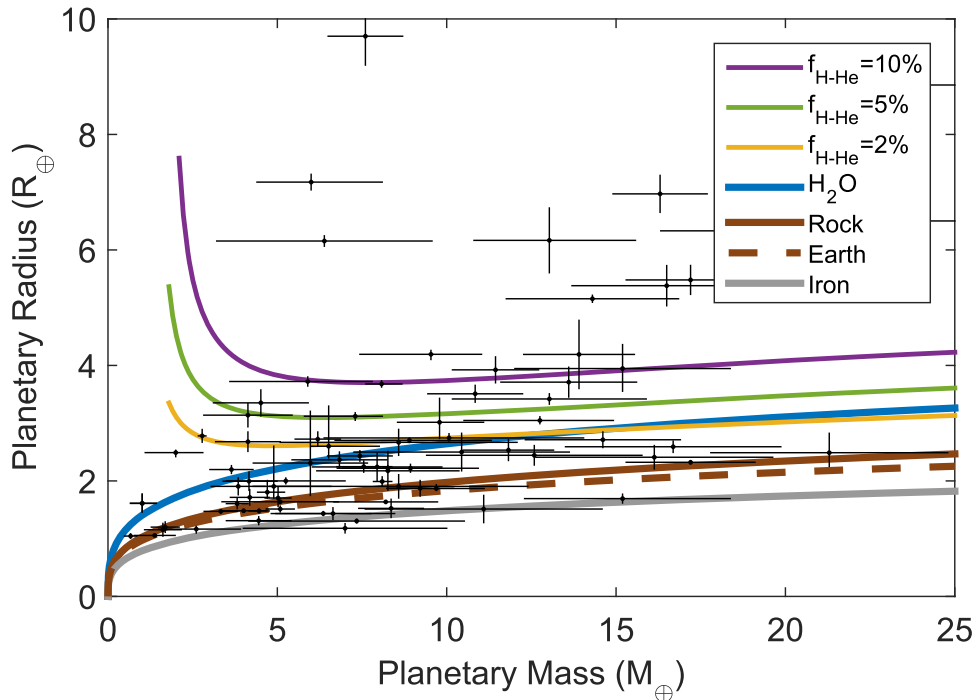


Figure 1. The M–R diagram for the planetary sample. The black dots with error bars are the planets we used in the analysis. The colored curves are examples of M–R relations for various theoretical compositions (see Section 2.2 for details). The three theoretical M–R curves for planets with H–He atmospheres correspond to a semimajor axis of 0.1 au and a homogeneously mixed planetary envelope (*scenario-1*) with $Z = 0.2$, as defined in Equation (2).

transiting planets with RV or TTV follow-ups, with radii up to $10 R_{\oplus}$ and masses up to $25 M_{\oplus}$. Planets with large uncertainties in the measured mass/radius (larger than 50%), as well as planets with a measured uncertainty larger than $1 R_{\oplus}$ and/or $4 M_{\oplus}$ are excluded. That leaves us with a sample of 83 planets. Figure 1 shows the planetary sample, and the theoretical M–R curves (see Section 2.2 for details). Corrections for data completeness are not considered since this study does not rely on the absolute frequency of planets; instead it is the measurement uncertainties that are relevant.

2.2. The Mass–Radius Relation

Several theoretical M–R relations for various compositions have been derived by several groups (e.g., Seager et al. 2007; Marcus et al. 2010; Zeng & Sasselov 2013; Lopez & Fortney 2014). We use various theoretical compositions, such as pure H_2O , pure $MgSiO_3$, Earth-like composition (32% Fe, 68% silicate), and pure Fe, based on Seager et al. (2007).

In addition to these compositions, we construct a series of planetary models with rocky cores and volatile envelopes. The volatiles assumed in the envelope include hydrogen, helium, and water. For simplicity, we consider two end-member scenarios for the planetary structure. These two structures bracket the expected radii for a given planetary composition.

1. In *scenario-1* a rocky core is surrounded by an envelope consisting of H–He and water. The hydrogen, helium, and water are assumed to be homogeneously mixed.
2. In *scenario-2* we assume a completely differentiated structure in which the rocky core is surrounded by an inner pure H_2O layer and an outer layer composed of pure H–He. For that case the envelope corresponds to the two-layer structure of water and H–He.

In both scenarios, we use the EOS of Saumon et al. (1995) for H–He, with the H–He ratio being 72% H to 28% He. For the rocky core we use the EOS of $MgSiO_3$ (Seager et al. 2007). In *scenario-1* the water EOS is based on ANEOS by Thompson (1990) (see Venturini et al. 2016 for details) while in *scenario-2* the EOS for water is based on QEOS presented by More et al. (1988) (see Vazan et al. 2013 for details). The reason for using different water EOSs is linked to the fact that we use two different codes for the two structures. This, however, does not impact the inferred M–R relation, as presented in Figure 2. The figure shows the M–R relation of a pure-water planet using ANEOS with a surface temperature $T = 500$ K, QEOS with $T = 300$ K, and QEOS with $T = 1200$ K, alongside the polytropic EOS of Seager et al. (2007) and the EOS of Wagner et al. (2011). The three upper curves, corresponding to QEOS at 300 K and 1200 K and ANEOS at 500 K, are very similar (less than 1% difference), suggesting that the inferred M–R relations should not be affected by using different water EOSs (see Section 4.1 for discussion). The M–R relations accounting for different mass fractions of H–He are derived by solving the standard internal structure equations; we use the irradiation model of Guillot (2010) for the atmosphere. More details on the structure model can be found in the Appendix and references therein.

2.2.1. Key Parameters

In both scenarios, the planetary composition is defined by two parameters: the mass fraction of H–He (f_{H-He}) and the mass fraction of water in the envelope (Z). These mass fractions are given by

$$f_{H-He} = \frac{M_{H-He}}{M_{H-He} + M_{H_2O} + M_{rock}}, \quad (1)$$

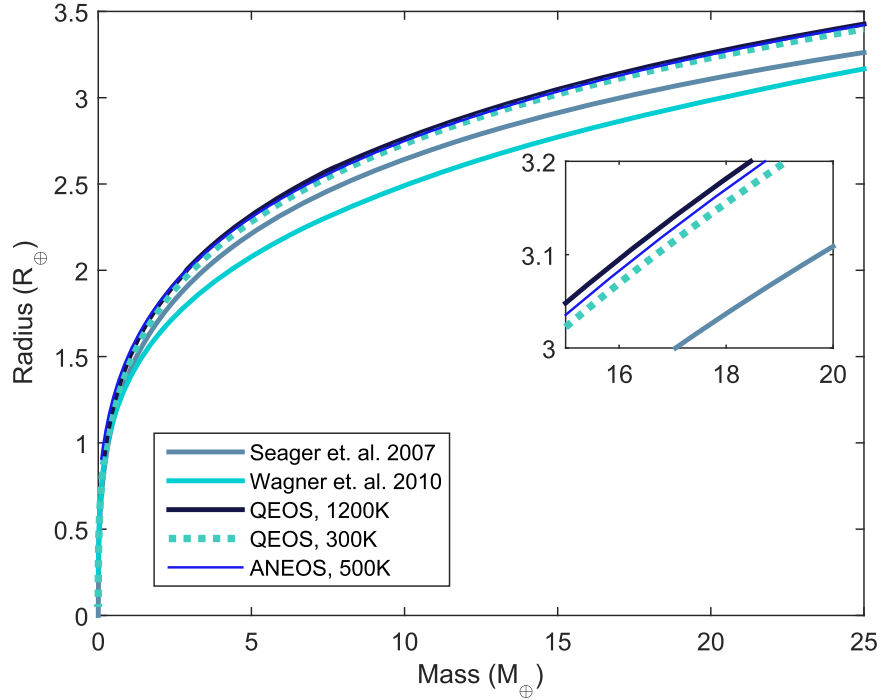


Figure 2. M–R relations for pure-water planets, using different EOSs for water. In this study we use the polytropic EOS of Seager et al. (2007) for a pure-water planet. Other EOSs presented here are that of Wagner et al. (2011), QEOS assuming surface temperatures of $T = 300$ K and $T = 1200$ K (More et al. 1988), and ANEOS (Thompson 1990) with a surface temperature of $T = 500$ K. In this work, ANEOS was used in *scenario-1* and QEOS was used in *scenario-2* (see the text for details).

$$Z = \frac{M_{\text{H}_2\text{O}}}{M_{\text{H}_2\text{O}} + M_{\text{H-He}}}, \quad (2)$$

where $M_{\text{H-He}}$, $M_{\text{H}_2\text{O}}$, and M_{rock} are the masses of H–He, water, and rock, respectively.

The envelope’s mass fraction (homogeneously mixed/differentiated) is then given by

$$f_{\text{env}} = \frac{M_{\text{H-He}} + M_{\text{H}_2\text{O}}}{M_{\text{H-He}} + M_{\text{H}_2\text{O}} + M_{\text{rock}}}. \quad (3)$$

Similarly, we can define the mass fraction of water by

$$f_{\text{H}_2\text{O}} = \frac{M_{\text{H}_2\text{O}}}{M_{\text{H-He}} + M_{\text{H}_2\text{O}} + M_{\text{rock}}}, \quad (4)$$

and that of rock by

$$f_{\text{rock}} = \frac{M_{\text{rock}}}{M_{\text{H-He}} + M_{\text{H}_2\text{O}} + M_{\text{rock}}}. \quad (5)$$

Since $f_{\text{H-He}} + f_{\text{H}_2\text{O}} + f_{\text{rock}} = 1$, the following relations can be derived:

$$f_{\text{H}_2\text{O}} = \left(\frac{Z}{1-Z} \right) f_{\text{H-He}} \quad (6)$$

$$f_{\text{rock}} = 1 - \frac{f_{\text{H-He}}}{1-Z} \quad (7)$$

$$\begin{aligned} f_{\text{W/R}} &= \frac{f_{\text{H}_2\text{O}}}{f_{\text{rock}}} \\ &= \frac{Z f_{\text{H-He}}}{1-Z-f_{\text{H-He}}} \end{aligned} \quad (8)$$

where the latter stands for the planetary mass ratio of water to rock.

It should be noted that, in principle, one can choose different key parameters to define the planetary composition, such as the $f_{\text{W/R}}$. Other possible definitions for a planetary structure model with alternative key parameters but the same composition (rock, water, H–He) are not physically different from our models, but differ in the way the mass fractions are defined.

Since we focus on intermediate- and low-mass planets, we consider $f_{\text{H-He}}$ values between 2% and 10% and Z values between 0 and 0.7. Such H–He mass fractions are the minimum expected for mini-Neptunes (e.g., Venturini & Helled 2017) and the envelope’s metallicities expected for Neptune-like planets (e.g., Helled et al. 2011). The mean density of planets with larger $f_{\text{H-He}}$ is significantly lower than the typical density of the observed exoplanets with masses up to $25 M_{\oplus}$. As a result, we do not consider $f_{\text{H-He}}$ larger than 10%.

2.2.2. M–R Relation of Volatile-rich Planets

The range of possible M–R relations for planets with volatile envelopes is presented in Figure 3. The row and the color correspond to $f_{\text{H-He}}$ (2%, 5%, 10%), and the column corresponds to various Z (0.1, 0.2, 0.4, 0.7). Each subplot shows the range of possible M–R relations for a given $f_{\text{H-He}}$ and Z . The upper limit is determined by the inferred M–R relation of a fully differentiated structure (*scenario-2*), while the lower one is given by the fully mixed models (*scenario-1*).

It is found that the range of possible models increases with increasing $f_{\text{H-He}}$ and Z values, but decreases with increasing planetary mass, as presented in Figure 4 (various colors represent different $f_{\text{H-He}}$ and the line style represents various Z). Figure 4 shows the absolute difference between *scenario-2* and *scenario-1* as a function of planetary mass. We confirm that the distribution of elements within the interior of intermediate-mass planets has a large effect on the radius (e.g., Baraffe et al. 2008; Vazan

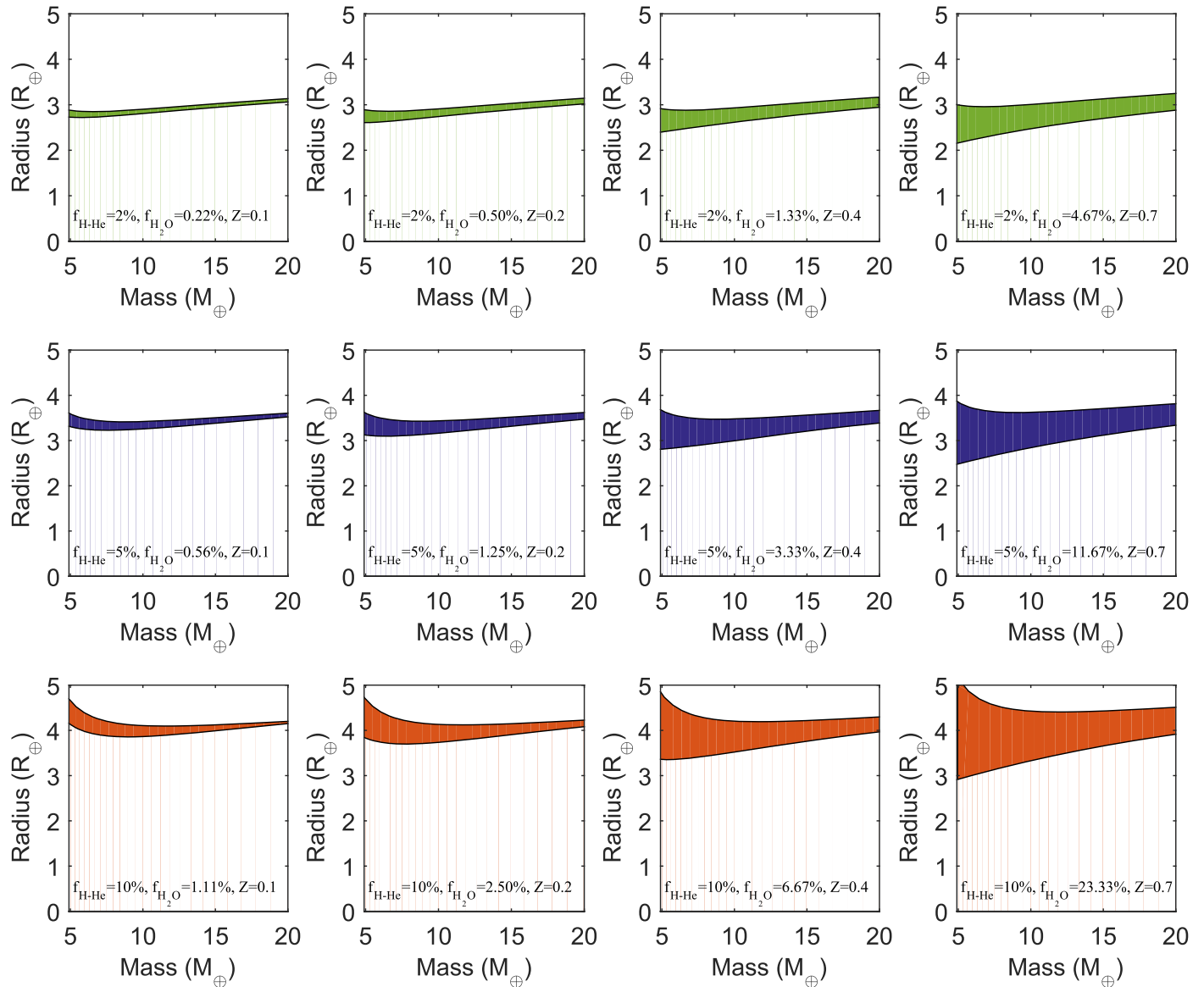


Figure 3. Ranges of M–R relation for structure models with a rocky core and various fractions of hydrogen and helium ($f_{\text{H-He}}$) and atmospheric metallicities (Z) (see the text for details). The lower limit on the radius corresponds to models with the water being mixed in the H–He envelope (*scenario-1*), while the upper one corresponds to the fully differentiated structure (*scenario-2*). The mass fractions of H–He ($f_{\text{H-He}}$) and water ($f_{\text{H}_2\text{O}}$) and the assumed envelope’s metallicity (Z) are indicated in each panel.

et al. 2016). This should be accounted for when characterizing planets in this mass/size regime.

As can be seen in the figures, $f_{\text{H}_2\text{O}}$ increases with Z (see Equations (2) and (4)), as well as with increasing $f_{\text{H-He}}$. While this might be counterintuitive, this behavior is a result of our definition of composition: increasing $f_{\text{H-He}}$ while keeping a constant Z leads to an increase in $f_{\text{H}_2\text{O}}$ (Equation (6)), and therefore to a decrease in f_{rock} (Equation (7)).

The planetary temperature must be included when studying the M–R relation of planets consisting of volatile materials (e.g., Swift et al. 2012; Lopez & Fortney 2014). The radii of planets with H–He atmospheres are larger for higher temperatures. It should be noted, however, that most of the planets in the sample have equilibrium temperatures of ~ 500 K and semimajor axes of ~ 0.1 au. Therefore, we use a semimajor axis of 0.1 au as the default when deriving the M–R relations. A more delicate analysis where the temperatures are derived for

each planet individually, accounting for semimajor axes and different albedos, was also performed and is presented in Section 4.2.

2.3. The Statistical Analysis

We aim to determine the probability of a given planet being denser than a given composition, based on the M–R relations. The theoretical M–R relations are used to define the transition between different possible compositional regimes. For example, the solid brown curve in Figure 1 corresponds to MgSiO_3 (the least dense silicate composition). If a given planet is above the curve, it indicates that the planet is less dense than pure rock, and thus has some volatiles (e.g., water and/or H–He). The gray line corresponds to pure iron and is used as the highest density possible for terrestrial planets. Planets between the lines for iron and rock are likely to consist of a mixture of silicates and iron and can therefore be referred as “potentially

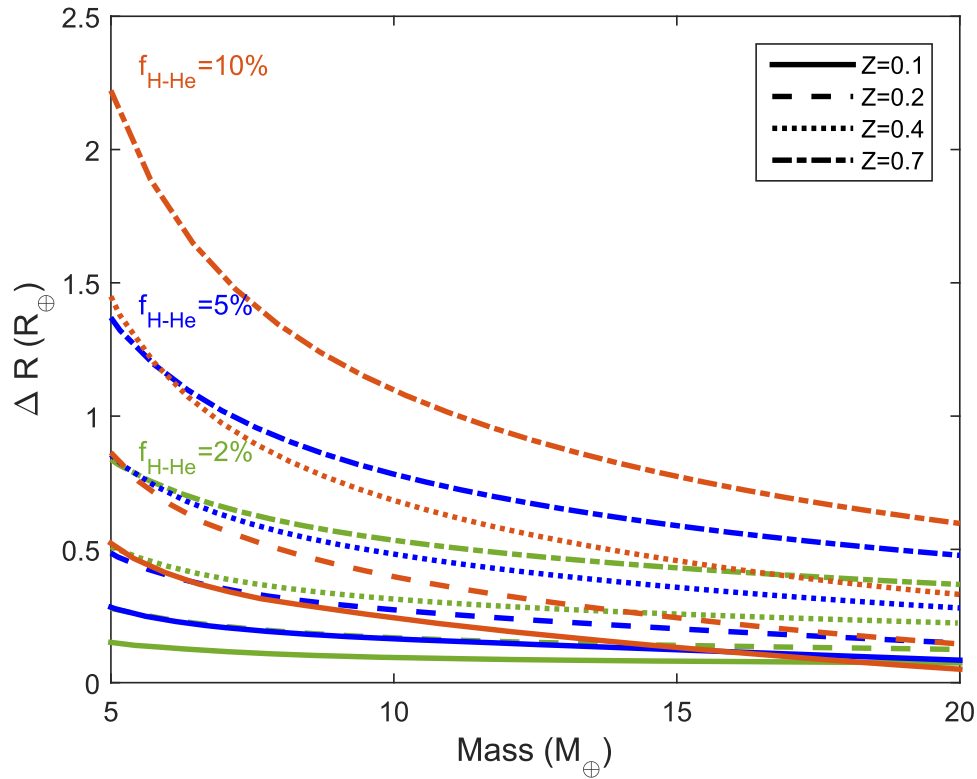


Figure 4. The differences in the inferred threshold radius between *scenario-1* and *scenario-2*. The colors represent the percentage of H–He ($f_{\text{H-He}}$), and the line style indicates the assumed atmospheric metallicity Z .

rocky” (Rogers 2015). A similar reasoning is applied for planets with H–He atmospheres: for example, a planet above our curve for $f_{\text{H-He}} = 5\%$ is likely to have an atmosphere with $f_{\text{H-He}} > 5\%$.

2.3.1. Calculation of Probability

The measured values of mass and radius and their uncertainties play a key role in the analysis; the larger the uncertainties, the larger the range of possible compositions. The planetary composition is constrained as follows: if a given planet is located below a given M–R curve, it suggests that the planet is denser than a specific composition (i.e., consists of heavier elements). Due to measurement uncertainties, we define the probability p of a planet being below a given M–R curve. Planets with $p \sim 1$ are very likely to be composed of elements that are denser than a particular structure, while planets with $p \sim 0$ are likely to be composed of lighter (i.e., more volatile) materials.

The measured values of mass and radius are assumed to have an asymmetric normal-like distribution:

$$M_{\text{pl}} \sim \mathcal{N}(M, M_{\text{err}\pm}) = \begin{cases} \mathcal{N}(M, M_{\text{err}-}) & M_{\text{pl}} < M \\ \mathcal{N}(M, M_{\text{err}+}) & M_{\text{pl}} > M \end{cases}, \quad (9)$$

$$R_{\text{pl}} \sim \mathcal{N}(R, R_{\text{err}\pm}) = \begin{cases} \mathcal{N}(R, R_{\text{err}-}) & R_{\text{pl}} < R \\ \mathcal{N}(R, R_{\text{err}+}) & R_{\text{pl}} > R \end{cases}, \quad (10)$$

where M and R are the measured radius and mass, respectively, and $M_{\text{err}\pm}$ and $R_{\text{err}\pm}$ are the corresponding measurement uncertainties.

For each measured mass–radius pair, we randomly sample 10,000 physically plausible synthetic planets. The simulated values are asymmetrically normally distributed using the measured data, $M_p \sim \mathcal{N}(M, M_{\text{err}\pm})$, $R_p \sim \mathcal{N}(R, R_{\text{err}\pm})$, as defined in Equations (9) and (10).

This simulated sample is used to determine p , the probability of a given planet being in a desired M–R region. The probability of a planet being denser than a given composition is given by the fraction of simulated points that are found below a given M–R curve:

$$p = \frac{\#\text{points below the curve}}{\#\text{simulated points}}. \quad (11)$$

Simulated points that fall below the iron curve (gray line in Figure 1) and points with negative radius and/or mass are unphysical, and are therefore excluded. We infer the probability p for each individual planet in our sample. This procedure is then repeated for every M–R curve separately. In order to ensure that the inferred value of p does not depend on the size of the simulated sample, we have run cases with smaller samples (1000 and 500) and got similar p values, suggesting that our derived value is robust.

2.4. Threshold Radius R_{th}

We investigate whether there is a sharp threshold on the distribution of p in terms of the planetary radius for different assumed compositions, which should result in a step function. Similarly to Rogers (2015), we represent the distribution of p as a function of R with a step function (upper panels in Figures 5 and 6):

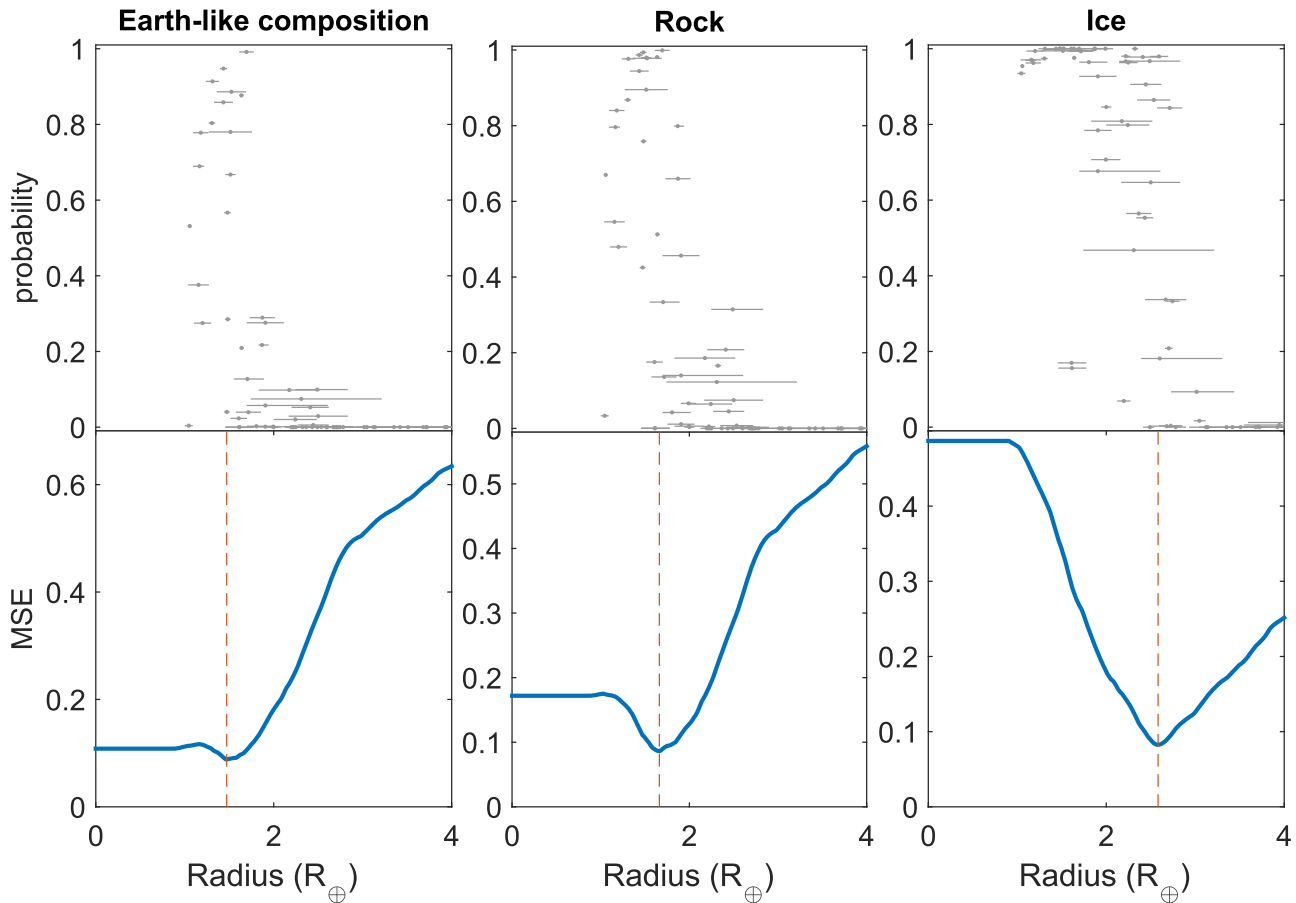


Figure 5. The probability of a planet being denser than a given composition, and the corresponding threshold radius R_{th} . Top: the probability of being denser than a given theoretical composition (Earth-like, pure rock, pure water) as a function of planetary radius. Bottom: mean squared error (MSE) of the fitted threshold step function. The red dashed line corresponds to the best fit (see Section 2.4).

$$\Theta(R_p, R_{\text{th}}) = \begin{cases} 1, & R_p < R_{\text{th}}, \\ 0, & R_p \geq R_{\text{th}}, \end{cases} \quad (12)$$

where R_p is the planetary radius and R_{th} is the threshold radius, to be found. Slightly more complex functions with gradual transitions have been explored by Rogers (2015), and it was found that the simple step function essentially coincides with the best fit of a gradual transition. As discussed above, R_{th} represents the lower limit on planetary radius that allows the planet to be denser than a given composition.

For every composition, we derive an ensemble of different possible threshold radii r_{th} in order to find the best fit for R_{th} . We search for a radius R_{th} that minimizes the mean squared error (MSE) between the p points and the curve $\Theta(R_p, r_{\text{th}})$. Formally, we minimize the term

$$\text{MSE}(r_{\text{th}}) = \frac{1}{j} \sum_j |\Theta(R_j, r_{\text{th}}) - p_j|^2, \quad (13)$$

where j runs over a fixed number of planets (83), R_j are the measured planetary radii, and p_j are the corresponding probabilities of being below a certain M–R curve (see upper panels of Figures 5 and 6). The quality of the fit is inversely proportional to the value of MSE, and therefore the threshold radius is $r_{\text{th}} = R_{\text{th}}$ that minimizes the function $\text{MSE}(r_{\text{th}})$. MSE is shown in the lower panels of Figures 5 and 6. The location of R_{th} , corresponding to the minimal value of MSE, is represented

by the red dashed line. The value of MSE starts to drop in a region of radius where p transits from ~ 1 to ~ 0 (Figure 6). The quality of the fit to R_{th} is better for lower values of MSE. The spread of planets in a transition region with $0 < p < 1$ correlates with the width of the trough, and therefore with the uncertainty on estimating R_{th} as we discuss below.

The uncertainties on the threshold radii were derived using a “bootstrap method.” In this method we use random sampling with replacement, based on pairs of the measured R_p and the corresponding calculated p . As discussed before, R_{th} is the value that minimizes the MSE function. In the bootstrapping method, we recalculate MSE and then R_{th} , using subsamples of the original (p, R_p) set. In the procedure, which is repeated 10,000 times, we are sampling with a replacement 83 pairs $(p_i, R_{p,i})$ from the original (p, R_p) sample. The simulated subsample $(p_i, R_{p,i})$ has the same length as the original one, but includes repetitions of a random number of values. For each subsample $(p_i, R_{p,i})$ we calculate the MSE_i , and find a corresponding threshold radius $R_{\text{th},i}$. This procedure creates 10,000 $R_{\text{th},i}$ values of the sample statistics.

The sample statistics may include some unphysical extreme values. Therefore, we exclude the lowest 2.5% and the highest 2.5% of the simulated set for $R_{\text{th},i}$. Then, we construct a cumulative distribution function (CDF) from the simulated $R_{\text{th},i}$ values (see Figure 7). The 50th percentile of the CDF corresponds to the mean value of R_{th} (shown by the dark gray dashed line in Figure 7), while the 16th and 84th percentile

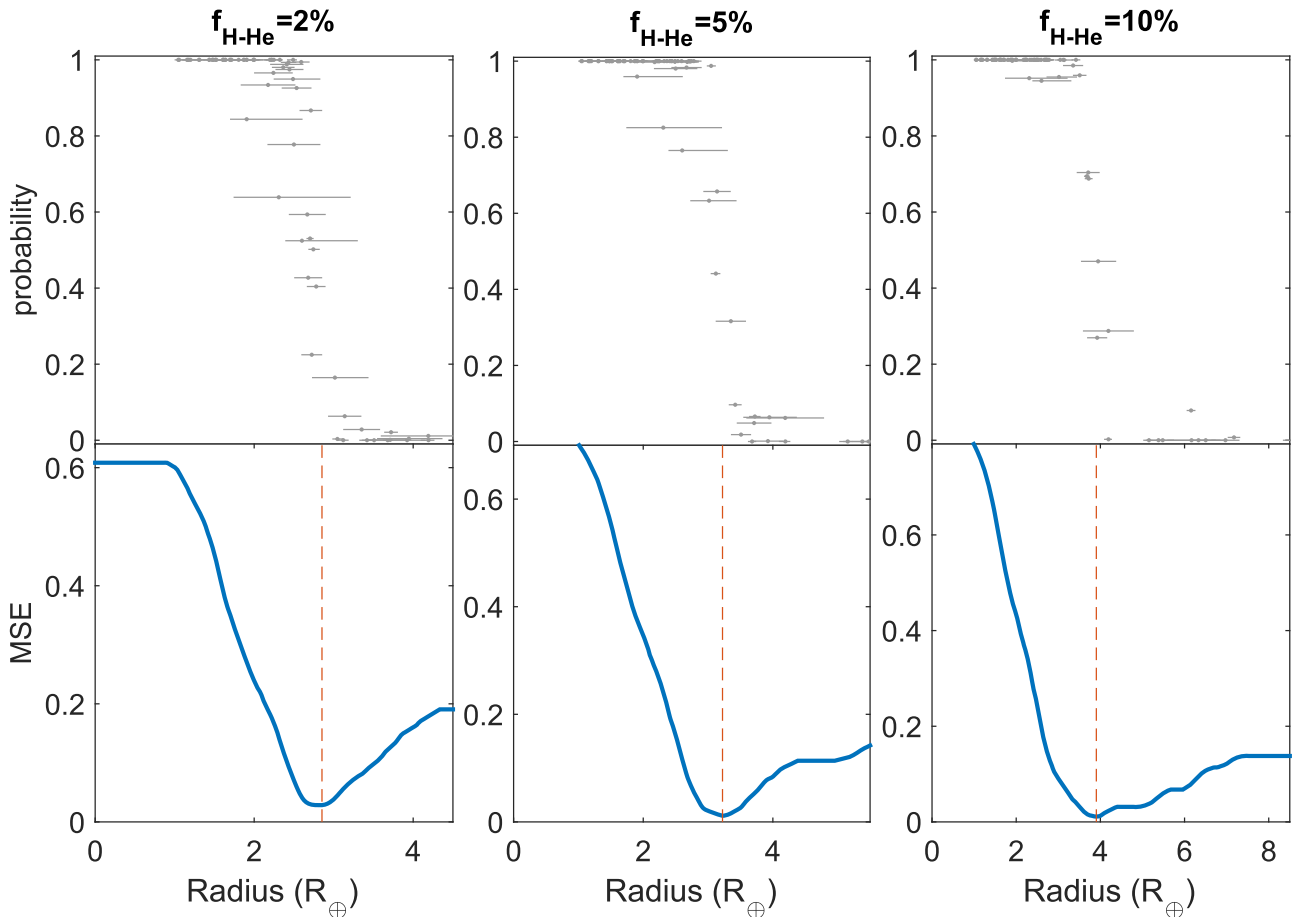


Figure 6. Same as Figure 5, for models with a rocky core surrounded by an envelope consisting of homogeneously mixed H–He and water (*scenario-1*) with $Z = 0.2$. The percentage in the title is the mass fraction of H–He ($f_{\text{H-He}}$).

values (shown by the light gray dashed lines) correspond to the standard error, and therefore to the formal uncertainties on R_{th} . Note that the asymmetric form of the MSE in Figure 6 leads to asymmetric uncertainties ($R_{\text{err}+}$ and $R_{\text{err}-}$ in Figure 7). In a few cases the boundaries are inconclusive due to a lack of separation between the 16th or/and 84th percentile and the 50th percentile. For these cases the method cannot provide a meaningful uncertainty (being referred to as n/a).

3. Results

The calculated threshold radii and their uncertainties for different possible compositions are summarized in Tables 1–6. The results for the threshold radii R_{th} are presented in Figures 5 and 6, where the top panels show the probability of being denser than a given composition p as a function of planetary radius R_p for different theoretical compositions.

3.1. Planets Without Volatiles

First, we confirm the result of Rogers (2015) that planets with radii larger than $1.6 R_{\oplus}$ are not pure rock. We find that planets above $1.4 R_{\oplus}$ cannot have Earth-like compositions (32% iron, 68% silicate) and therefore have to contain a larger fraction of silicates or lighter materials. The threshold radii for volatile-poor planets are summarized in Table 1.

Table 1

Derived Threshold Radii R_{th} (as Defined in Section 2.4) and Their Uncertainties for Various Possible Compositions Without Volatiles

Composition	$R_{\text{th}} (R_{\oplus})$
Earth-like	$1.47^{+0.11}_{-0.01}$
Pure rock	$1.66^{+0.01}_{-0.08}$
Pure water	$2.58^{+0.05}_{-0.05}$

3.2. Planets with Volatile Envelopes

We find that most of the planets larger than $2.6 R_{\oplus}$ cannot be pure-water worlds. As expected, the planetary radius typically increases with $f_{\text{H-He}}$. The results for planetary models with gaseous envelopes are listed in Table 2. We find that R_{th} of planets with homogeneous envelopes (*scenario-1*) is typically lower than R_{th} of the differentiated structure (*scenario-2*) for the same bulk composition (see Section 2.2 for details). That is illustrated in Figure 3, which shows the boundaries derived from the two scenarios.

It is interesting to note that while in the fully mixed case (*scenario-1*) R_{th} decreases with Z , the opposite occurs for the differentiated structure (*scenario-2*). This is a consequence of the way we built our planets: in *scenario-1* increasing the envelope’s metallicity leads to a contraction of the envelope due to self-gravity, and thus to a smaller radius. In *scenario-2*, the volume of pure H–He is constant for a given $f_{\text{H-He}}$, and the

Table 2
Derived Threshold Radii R_{th} (in R_{\oplus}) Assuming Different Envelope Metallicities (Z) and H–He Mass Fractions ($f_{\text{H–He}}$), as Defined in Section 2.2.1

	$f_{\text{H–He}} = 2\%$		$f_{\text{H–He}} = 5\%$		$f_{\text{H–He}} = 10\%$	
	Mixed	Differentiated	Mixed	Differentiated	Mixed	Differentiated
$Z = 0.0$	$2.90^{+0.02}_{-0.01}$	$2.90^{+0.02}_{-0.01}$	$3.49^{+n/a}_{-0.22}$	$3.49^{+0.04}_{-0.13}$	$4.04^{+0.05}_{-0.00}$	$4.06^{+0.03}_{-0.02}$
$Z = 0.1$	$2.90^{+0.02}_{-0.03}$	$2.90^{+0.02}_{-0.01}$	$3.31^{+0.02}_{-0.02}$	$3.53^{+0.03}_{-0.17}$	$3.86^{+0.18}_{-0.01}$	$4.06^{+0.03}_{-0.02}$
$Z = 0.2$	$2.81^{+0.05}_{-0.13}$	$2.92^{+n/a}_{-0.03}$	$3.22^{+0.02}_{-0.01}$	$3.53^{+0.03}_{-0.17}$	$3.85^{+0.01}_{-0.08}$	$4.09^{+n/a}_{-0.03}$
$Z = 0.4$	$2.59^{+0.04}_{-0.03}$	$2.92^{+n/a}_{-0.03}$	$2.94^{+0.03}_{-0.02}$	$3.53^{+0.03}_{-0.04}$	$3.56^{+0.01}_{-0.03}$	$4.33^{+n/a}_{-n/a}$
$Z = 0.7$	$2.52^{+0.03}_{-0.14}$	$2.97^{+0.22}_{-0.08}$	$2.78^{+0.12}_{-0.10}$	$3.65^{+n/a}_{-0.12}$	$3.19^{+n/a}_{-0.20}$	$4.33^{+n/a}_{-n/a}$

Note. Listed are the results for the two structural scenarios of a fully mixed envelope (*scenario-1*) and a fully differentiated planet (*scenario-2*) (see Section 2.2 for details). The planetary albedo is assumed to be $A = 0$ and the semimajor axis 0.1 au. In some cases the statistical test is inconclusive and cannot provide an estimate of the threshold boundaries.

Table 3

Derived Threshold Radii for Pure-water Planets Using Different Water EOSs

H ₂ O EOS	R_{th} (R_{\oplus})
Seager et al. (2007)	$2.58^{+0.05}_{-0.05}$
Wagner et al. (2011)	$2.53^{+0.03}_{-0.08}$
More et al. (1988)	$2.63^{+0.10}_{-0.05}$

Table 4

Derived Threshold Radii R_{th} and Their Uncertainties Assuming Different Semimajor Axes and Albedos

Semimajor Axis (au)	Planetary Temperature, T_{eq}		R_{th} (R_{\oplus})
	(K)		
0.05	1120		$3.48^{+0.03}_{-0.17}$
0.1	770		$3.22^{+0.02}_{-0.01}$
0.5	330		$2.92^{+0.18}_{-0.03}$
Calculated individually, $A = 0.0$	Calculated individually		$3.19^{+0.08}_{-0.22}$
Calculated individually, $A = 0.3$	Calculated individually		$3.19^{+0.08}_{-0.22}$
Calculated individually, $A = 0.9$	Calculated individually		$2.97^{+0.02}_{-0.08}$

Note. $f_{\text{H–He}}$ is set to be 5% with a metallicity of $Z = 0.2$. The planetary envelope is assumed to be fully mixed (*scenario-1*).

Table 5

Derived Threshold Radii R_{th} Using Different Luminosities (in Units of Neptune’s Luminosity)

Luminosity (L_N)	R_{th} (R_{\oplus})
Standard case	$3.22^{+0.02}_{-0.01}$
$\sim 10^{-1}$	$3.19^{+0.01}_{-0.22}$
$\sim 10^2$	$3.27^{+0.01}_{-0.08}$
$\sim 10^3$	$3.53^{+0.03}_{-0.04}$

Note. We present a fully mixed case (*scenario-1*) of $f_{\text{H–He}} = 5\%$ and $Z = 0.2$ with a luminosity set by Equation (A5) as described in the Appendix, versus a range of constant luminosities.

only effect of increasing Z is an increase in the ratio of water to rock (Equation (8)). Since the amount of water, which has a lower density than rock, is increased at the expense of reducing the abundance of rocks in *scenario-2* the planet’s radius increases with increasing Z .

For the various compositions and internal structures we consider (Table 2), we find threshold radii in the range between

Table 6

Summary of a Data Stability Test, Described in Section 4.4

Planet Sample	R_{th} (R_{\oplus})
Original data	$3.24^{+0.03}_{-0.03}$
Modified data sample 1	$3.22^{+0.70}_{-0.38}$
Modified data sample 2	$3.22^{+0.70}_{-0.38}$
Modified data sample 3	$3.31^{+0.61}_{-0.29}$

Note. The results compare R_{th} found from the real data and slightly randomized data sets (modified data). The composition used for the testing is *scenario-1* with $f_{\text{H–He}} = 5\%$ and $Z = 0.2$.

2.5 and 4.3 R_{\oplus} depending on the chosen scenario, $f_{\text{H–He}}$, and Z . We find that planets above 2.6 R_{\oplus} must have significant amounts of H–He and therefore can be classified as mini-Neptunes. There is degeneracy between the threshold for pure-water worlds ($R_{\text{th}} = 2.58^{+0.05}_{-0.05} R_{\oplus}$) and mini-Neptunes with high ($Z = 0.7$) atmospheric metallicity ($R_{\text{th}} = 2.52^{+0.03}_{-0.14} R_{\oplus}$). Although the composition of planets with these sizes cannot be inferred exactly, we can conclude that planets with radii larger than $\sim 2.6 R_{\oplus}$ are likely to consist of H–He atmospheres. In addition, we find that planets with sizes $R_p \gtrsim 4 R_{\oplus}$ are likely to have significant H–He atmospheres (more than 10% of the planetary mass). It should be noted that the planets in our sample are significantly hotter than Uranus and Neptune (see Figure 8). Therefore the planets that are found to have H–He envelopes are hot- and warm-Neptunes/mini-Neptunes, i.e., with compositions similar to that of Uranus/Neptune but with a larger radius due to stellar irradiation (Baraffe et al. 2006).

4. Sensitivity of the Results to the Model Assumptions

4.1. Sensitivity to the EOS of Water

The EOS of water (as well as other elements) is still not perfectly known, especially in the regime of high pressure temperature. In the case of a pure-water planet, the default EOS we use for water is that of Seager et al. (2007). However, differences in the water EOS could lead to differences in the derived radii. As a result, we investigate the sensitivity of R_{th} to the assumed water EOS. A comparison of the M–R relation for pure-water planets using three different water EOSs is presented in Figure 2. Additional curves using the EOSs for water of Wagner et al. (2011) and More et al. (1988) and different effective temperatures are also presented. The difference between the EOSs is mainly linked to the different assumed bulk moduli (K_0 and K'_0) of the planetary ices, but is

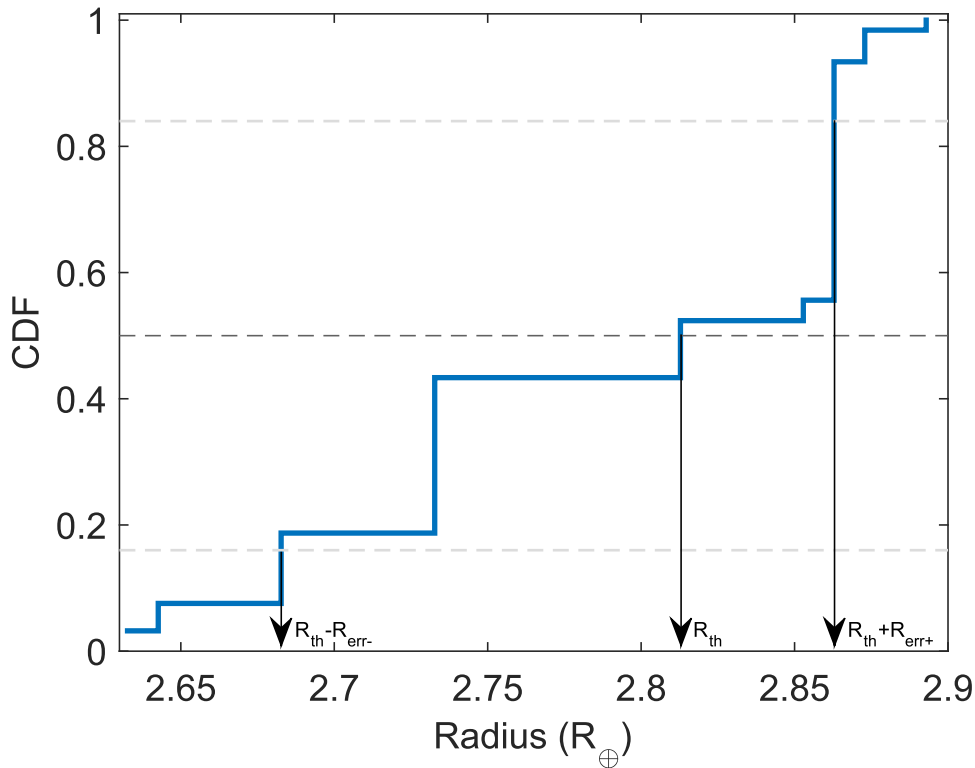


Figure 7. Cumulative distribution function (CDF) of R_{th} for a composition in *scenario-1* with 2% of H–He and $Z = 0.2$. The median of R_{th} is evaluated at the 50th percentile, while the lower ($R_{\text{err-}}$) and upper ($R_{\text{err+}}$) uncertainty values are taken to be the 16th and 84th percentiles.

in general relatively small. Therefore, the inferred threshold radii are relatively insensitive to the water EOS. The derived R_{th} assuming various EOSs for water are presented in Table 3. We find that R_{th} changes from $2.58^{+0.05}_{-0.05} R_{\oplus}$ to $2.63^{+0.10}_{-0.05} R_{\oplus}$ for the most extreme cases. As shown in Figure 2, the differences between ANEOS and QEOS for water are negligible. We can therefore conclude that the differences between the M–R relations derived in *scenario-1* and *scenario-2* are linked to the distribution of elements (i.e., the assumed internal structure) and are not affected by the choice of the water EOS.

4.2. Sensitivity to the Equilibrium Temperature and Planetary Albedo

The M–R relation derived for compositions (and structures) with significant amount of volatiles depends on the planet’s equilibrium temperature. Information on the stellar and orbital properties of the system, such as stellar temperature, stellar radius and semimajor axis, can be used to calculate the planetary equilibrium temperature T_{eq} . The equilibrium temperature is given by

$$T_{\text{eq}} = T_{\odot} (1 - A)^{1/4} \sqrt{\frac{R_{\odot}}{2D}}, \quad (14)$$

where T_{\odot} and R_{\odot} are the stellar temperature and radius, respectively, D is the semimajor axis, and A is the planetary albedo. In this study we set the default case for a Sun-like star, semimajor axis of 0.1 au, and albedo of 0 (i.e., blackbody, full absorption). Our default values correspond to $T_{\text{eq}} = 770$ K.

We next investigate the effect of the semimajor axis on the threshold radius (assuming a constant albedo of $A = 0$). We derive new M–R relations for *scenario-1* with 5% of H–He and repeat the analysis for the new curves using various semimajor

axes. We find that R_{th} varies from $2.9 R_{\oplus}$ to $3.5 R_{\oplus}$ for semimajor axes between 0.05 au and 0.5 au, with the larger radius corresponding to the smaller radial distance. The inferred threshold radii for different assumed temperatures are summarized in the first three rows of Table 4. Since a semimajor axis of ~ 0.1 au corresponds to the majority of the planets in the sample, we use this as the default value.

The planetary albedo depends on many variables such as cloud layers and chemical composition. To explore the sensitivity of the inferred R_{th} to the assumed albedo we perform the analysis assuming three different albedo values: $A = 0$ (low, full absorption), $A = 0.3$ (Earth-like, intermediate), and $A = 0.9$ (high), while keeping a semimajor axis of 0.1 au. A histogram of the derived planetary temperature for the different albedos is presented in Figure 8. As expected, higher albedo leads to a lower equilibrium temperature.

In a second test, instead of using a single M–R relation with a constant temperature for the entire planetary sample, as done in Section 3.2, we adjust the M–R relation for each planet individually using the calculated equilibrium temperature and assumed albedo. We then calculate p for each planet using the individually calculated M–R curve. After the p distribution for the planetary sample is derived, we find R_{th} in the same fashion as described above. The results are summarized in the lower three rows of Table 4. We find that R_{th} is relatively insensitive to the assumed albedo value. The cases with albedos of 0 and 0.3 are essentially identical ($3.19 R_{\oplus}$), and the inferred threshold radius is not very different even when using $A = 0.9$ ($2.97 R_{\oplus}$).

4.3. Sensitivity to Luminosity

The planetary luminosity in our models is calculated using the luminosity fit of Rogers & Seager (2010) (see the Appendix

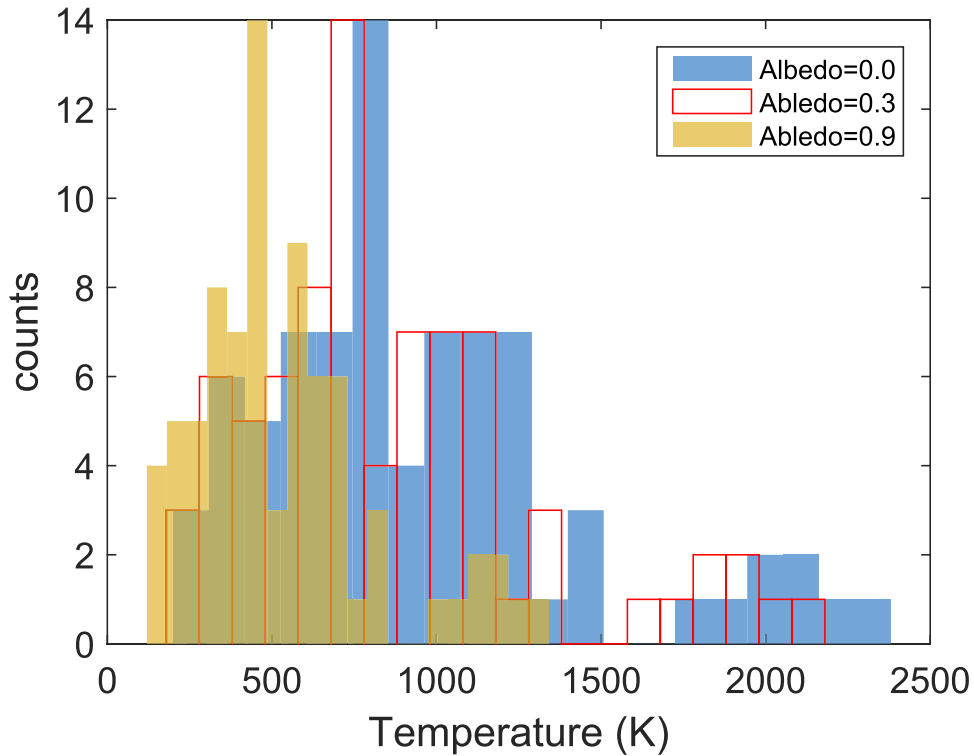


Figure 8. Histograms of planetary equilibrium temperature for three different albedos: $A = 0.0$, $A = 0.3$, and $A = 0.9$.

for details). In this approach the luminosity varies as a function of planetary mass and radius in the range of $\sim 10^0$ – $10^3 L_N$, where L_N is Neptune’s luminosity (377 GW). In order to test the sensitivity of the results to the assumed luminosity, we apply different constant luminosity values (between $\sim 10^{-1} L_N$ and $10^3 L_N$) when constructing the planetary structure and their corresponding M–R relations. We apply the comparison to *scenario-1* with $f_{\text{H-He}} = 5\%$ and $Z = 0.2$. The comparison between the different cases is summarized in Table 5). As can be seen from the table, R_{th} changes only by up to 10% for a range of luminosities that covers several orders of magnitude. We can therefore conclude that the inferred threshold radii are relatively insensitive to the assumed luminosity.

4.4. Sensitivity to the Data Sample Used

Systematic observational biases can influence the determination of the planetary mass and/or radius. In order to investigate the sensitivity of our results to the data used, i.e., exact values of the measured mass and radius (and their uncertainties), we explore the dependence of R_{\oplus} on the exact values of the masses and radii. We then randomly reassign the masses (with their corresponding uncertainties) for 20% of the planets in the sample (16 planets in total), while keeping the same radii and their corresponding uncertainties. In this way we ensure that the distributions of mass and radius of the synthetic data sample are kept the same as the original data. We then repeat the analysis using the new partially randomized sample (modified data sample). This test has been performed three times in order to explore the robustness of the results. The results are presented in Figure 9 and Table 6. We find that the threshold radius changes by less than 2%, and we therefore conclude that our results are robust.

4.5. The Existence of a Threshold Mass M_{th} ?

We find that the existence of threshold radii is statistically significant, and in principle one could expect to have a similar behavior for the mass. Models of planet formation predict that a heavy-element core starts to accrete H–He in significant amounts at around the so-called *critical core mass*. The value of this mass depends primarily on the solid accretion rate, envelope composition, and opacity (e.g., Ikoma et al. 2000; Venturini et al. 2016; Venturini & Helled 2017), and represents the transition between gas-poor and gas-rich planets. At the moment, the estimates for this critical core mass range from less than $1 M_{\oplus}$ to several Earth masses (e.g., Pollack et al. 1996; Brouwers et al. 2018). Therefore, finding a threshold mass could provide important constraints on the physical conditions and dominant processes during planet formation.

We investigate whether we can infer threshold masses M_{th} using a similar analysis to the one used for the radii. We perform the test for the case of a pure-rock composition, which is very robust. The results are presented in Figure 10. Unlike for the radii (left panel), for the masses (right panel) no sharp transition is observed, and therefore the existence of M_{th} is less convulsive. In the figure, we highlight 10 planets with the most accurate determination of radius/mass. The threshold on mass is also less distinct for this subsample.

The inferred distributions of p as a function of planetary mass for all the compositions we considered (not shown) do not resemble a sharp step function in the form of Equation (12). As mentioned above, reducing the sample to 10 planets with the smallest relative measurement uncertainties (orange circles) leads to a similar conclusion. This suggests that the distribution of masses is more continuous than that of the radii and that there are no sharp mass boundaries. Nevertheless, at the

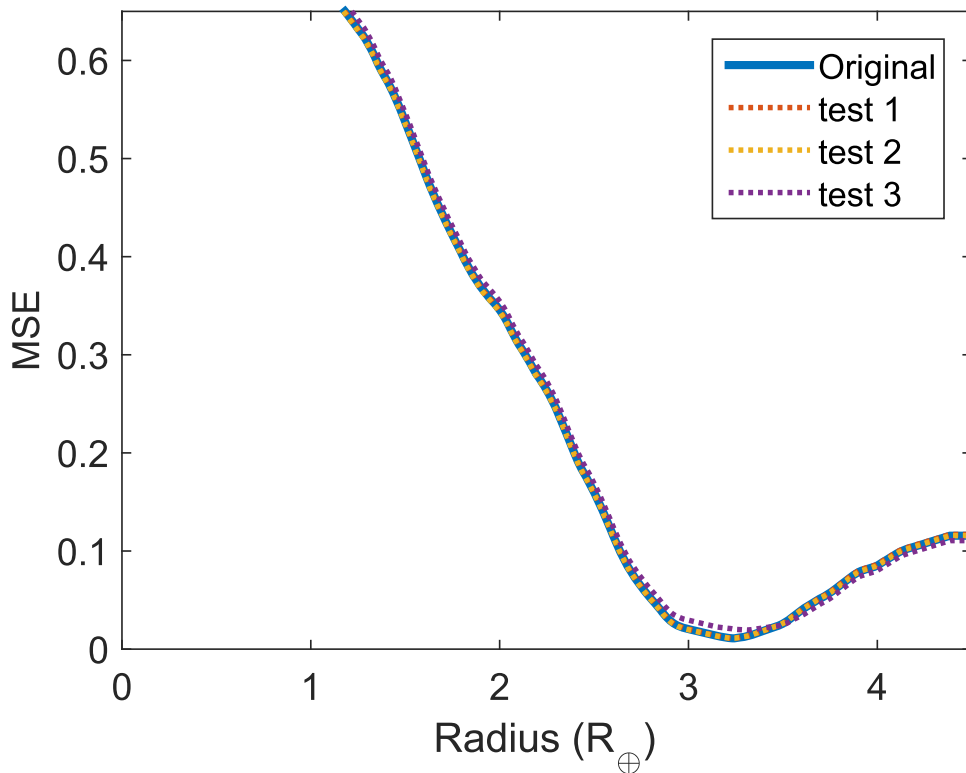


Figure 9. Sensitivity to the data sample used. Shown are the MSE for the original M–R data (solid line) vs. the test (modified) cases (dotted lines). See the text for details.

moment, we cannot exclude the possibility that threshold masses do not exist. In order to do so it is desirable to have a large number of planets with accurate measured masses. At the moment, we can only conclude that our sample of planets cannot be used to determine threshold masses using the same methodology. Accurate measurements with $\sim 5\%$ uncertainty on *both* the planetary mass and radius are expected from *PLATO*, with the masses being determined via ground-based RV follow-ups (Rauer et al. 2014). This could reveal the existence of M_{th} and can be used to further refine the threshold radii.

4.5.1. The Photoevaporation Valley

The bimodal size distribution of the *Kepler* planets reported by Fulton et al. (2017) and Fulton & Petigura (2018) and explained by photoevaporation models (Jin et al. 2014; Lopez & Fortney 2014; Owen & Wu 2017) suggests that most exoplanets originally formed with H–He, but the less massive planets lost it at a later stage due to their low gravity and strong irradiance from the host star. Since the valley falls in the size range of $1.5\text{--}2 R_{\oplus}$, these models adopt a radius of $1.7 R_{\oplus}$ as a boundary between super-Earths and mini-Neptunes. Thus, in the photoevaporation interpretation, the definitions of super-Earth and mini-Neptune do not reflect a difference in the origin of the objects, but in their evolution. In our study, we provide threshold radii based on the existing data, and therefore cannot provide predictions that go beyond the available data. Our method does not exclude the possibility that some planets with radii of $\sim 2 R_{\oplus}$ have H–He atmospheres, as is inferred from the work of Fulton. It is interesting to note that the gap for planets in the size range of $1.5\text{--}2 R_{\oplus}$ is not empty (Fulton &

Petigura 2018). It is hard to reconcile this with a scenario where planets are composed of a purely rocky core surrounded solely by an H–He envelope: planets with a radius in the gap should be unstable towards photoevaporation, losing H–He until they reach the first peak of the distribution (Owen & Wu 2017; Jin et al. 2014). Perhaps this suggests that exoplanetary atmospheres are typically enriched with heavy elements and are not made of pure H–He.

5. Summary and Conclusions

We present a statistical analysis to determine the threshold radii of volatile-rich planets. We show that different assumed compositions and internal structures with fixed $f_{\text{H-He}}$ and Z lead to a range threshold radii. As a result, in order to characterize individual planets information on their orbital properties and atmospheric compositions is required. However, despite the degenerate nature of the problem we suggest that there are characteristic threshold radii for different compositions.

First, we confirm that planets with radii larger than $1.6 R_{\oplus}$ are not rocky, and must consist of lighter elements. This conclusion is consistent with the work of Rogers (2015), despite some differences in the statistical analysis and the planetary sample used. It is found that distinguishing a pure-water planet from a rocky planet with a thin H–He atmosphere is not possible. Therefore, planets that are classified as ocean planets might in reality be planets with a rocky core and a volatile atmosphere (Adams et al. 2008).

Second, we show that most of the planets larger than $\sim 3 R_{\oplus}$ must contain at least 2% of H–He, while most of the planets with radii larger than $\sim 3.6 R_{\oplus}$ and $4.3 R_{\oplus}$ must contain at least 5% and 10% of H–He, respectively.

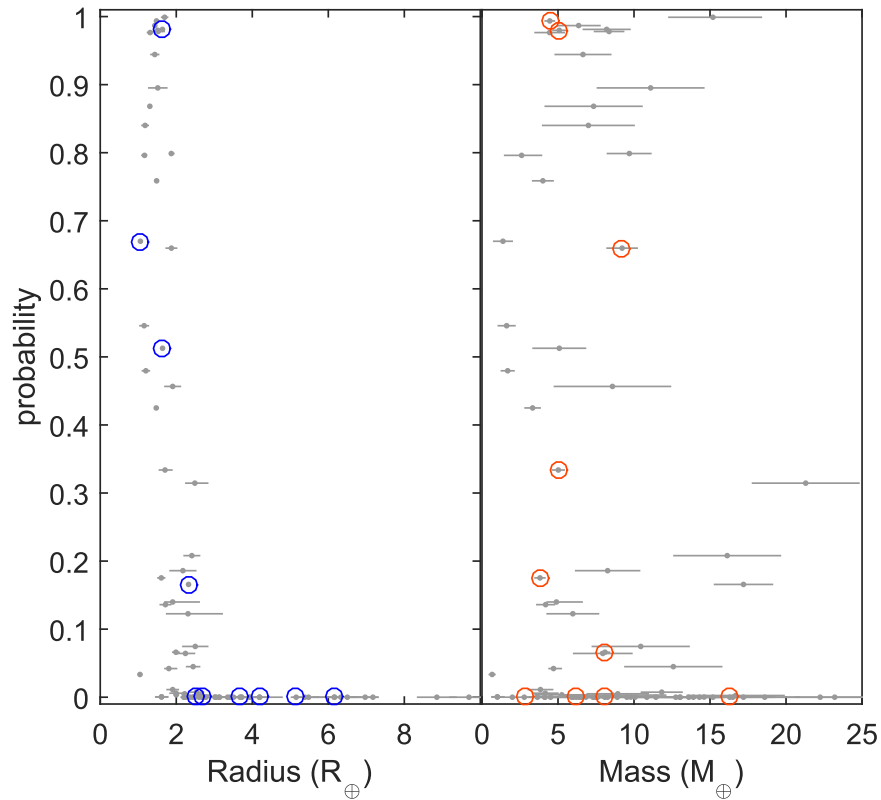


Figure 10. The probability of being denser than pure rock as a function of planetary radius (left) and mass (right). The 10 planets with the best determination of radius or mass are highlighted by the blue and orange circles, respectively. While the threshold on radius is significant, the threshold on mass is less distinct.

While the exact value of R_{th} depends on the model assumptions (i.e., composition, structure, thermal state, EOS), we find a range of threshold radii of $\sim 2.5\text{--}4.3 R_{\oplus}$ for planets with rocky cores and gaseous atmospheres. These radii are typically larger than the threshold radii for pure-water planets ($R_{\text{th}} \sim 2.6 R_{\oplus}$). We find that although the albedo and semimajor axis of the planet affect its temperature, they have a relatively small impact on the inferred R_{th} . For albedos between zero and 0.9, R_{th} varies from $\sim 3 R_{\oplus}$ to $\sim 3.2 R_{\oplus}$, suggesting that assumed albedo has a very small impact on R_{th} . We suggest that high planetary luminosity leads to somewhat larger R_{th} ; in the range of sensible luminosities ($L \sim 10^{-1}\text{--}10^2 L_N$, where L_N is Neptune’s luminosity) the change in R_{th} is very small, suggesting that R_{th} is relatively insensitive to the assumed planetary luminosity.

The key conclusions of our study can be summarized as follows.

1. We confirm that planets with radii larger than $\sim 1.6 R_{\oplus}$ are not purely rocky worlds and must consist of lighter elements.
2. Planets with radii larger than $\sim 2.6 R_{\oplus}$ are not pure-water worlds and must have atmospheres (presumably of H–He).
3. By defining a mini-Neptune (or a Neptune analog) as a planet with at least 2% of H–He by mass, we find that the transition from super-Earths (planets consisting of less than 2% of H–He) to mini-Neptunes occurs at $\sim 3 R_{\oplus}$.
4. Planets with radii larger than $\sim 4 R_{\oplus}$ are expected to consist of at least 10% of H–He and are therefore gas-rich.

Upcoming data from space missions such as *CHEOPS*, *TESS*, and *PLATO* as well as ground-based facilities will

further constrain the possible compositions of exoplanets. Measurements of planets with similar masses at larger radial distances will allow us to extend our scheme and characterize colder planets, and reveal whether the threshold radii are expected to change with distance to the host star. Finally, accurate measurements of both the masses and radii of low- and intermediate-mass exoplanets will allow us to determine whether threshold masses exist. This will significantly improve our understanding of the formation, evolution, and internal structures of planets in the solar neighborhood.

M.L. thanks Omri Harosh and Uzi Vishne for valuable remarks and advice on the statistical analysis. R.H. acknowledges support from SNSF grant 200021_169054. Some of this work has been carried out within the framework of the National Centre for Competence in Research PlanetS, supported by the Swiss National Foundation.

Appendix Atmospheric Model

The mass–radius relations are derived using the standard structural equations of hydrostatic equilibrium, mass conservation, and heat transport for the gaseous envelope:

$$\frac{dP}{dr} = -\frac{Gm\rho}{r^2}, \quad (\text{A1a})$$

$$\frac{dm}{dr} = 4\pi r^2 \rho, \quad (\text{A1b})$$

$$\frac{dT}{dr} = \frac{T}{P} \frac{dP}{dr} \nabla, \quad (\text{A1c})$$

where r is the radius, m is the corresponding cumulative mass, ρ is the density of each shell, P is a pressure, G is the gravitational constant, σ is the Stefan–Boltzmann constant, L is the intrinsic luminosity, and ∇ is the minimum between the adiabatic and radiative gradients (Equation (A4)).

To account for irradiation, we use a semi-gray atmosphere model (Guillot 2010; Jin et al. 2014), in which two opacity sources are included: the visible (κ_v) and infrared (κ_{th}) mean opacities. The optical depth is computed, which is given by

$$\frac{d\tau}{dr} = \kappa_{th}\rho, \quad (\text{A2})$$

κ_{th} being the infrared mean opacity, taken as $\kappa_{th} = 0.01 \text{ g cm}^{-3}$ (Guillot 2010).

The temperature gradient of the irradiated atmosphere is given by (Guillot 2010)

$$\begin{aligned} T^4 = & \frac{3T_{\text{int}}^4}{4} \left[\frac{2}{3} + \tau \right] + \frac{3T_{\text{eq}}^4}{4} \\ & \times \left[\frac{2}{3} + \frac{2}{3\gamma} \left\{ 1 + \left(\frac{\gamma\tau}{2} - 1 \right) e^{-\gamma\tau} \right\} \right. \\ & \left. + \frac{2\gamma}{3} \left(1 - \frac{\tau^2}{2} \right) E_2(\gamma\tau) \right] \end{aligned} \quad (\text{A3})$$

where $\gamma = \kappa_v/\kappa_{th}$ (ratio between visible and infrared opacities), T_{int} is the intrinsic temperature given by $T_{\text{int}} = (L/(4\pi\sigma r^2))^{1/4}$, and $E_2(\gamma\tau)$ is the exponential integral, defined by $E_n(z) \equiv \int_1^\infty t^{-n} e^{-zt} dt$ with $n = 2$. γ is taken from the calibration of Jin et al. (2014). The boundary between the irradiated atmosphere and the envelope is set at $\gamma\tau = 100/\sqrt{3}$ (Jin et al. 2014). For $\gamma\tau$ larger than this, the usual Schwarzschild criterion to distinguish between convective and radiative layers is applied. That is, if the adiabatic temperature gradient is larger than the radiative one, the layer is stable against convection, and the radiative diffusion approximation is used for computing the temperature gradient:

$$\frac{dT}{dr} = -\frac{3\kappa_{th}L\rho}{64\pi\sigma T^3 r^2} \quad (\text{A4})$$

where L is the intrinsic luminosity.

The planetary luminosity uses the luminosity fit of Rogers & Seager (2010), which corresponds to calculations of planet evolution derived by Baraffe et al. (2008), and is given by

$$\begin{aligned} \log \frac{L}{L_\odot} = & a_1 + a_{M_p} \log \frac{M_p}{M_\oplus} + a_{R_p} \log \frac{R_p}{R_{\text{jup}}} \\ & + a_{t_p} \log \frac{t_p}{1 \text{ Gyr}}, \end{aligned} \quad (\text{A5})$$

where L_\odot is the solar luminosity, M_p is the planetary mass in Earth masses, R_p is the planetary radius in Jupiter radii, and t_p is the stellar age (taken to be 5 Gyr). The coefficients are $a_1 = -12.46$, $a_{M_p} = 1.74$, $a_{R_p} = -0.94$, $a_{t_p} = -1.04$. It

should be noted that the atmospheric temperature does not depend on the intrinsic luminosity alone, but also on stellar irradiation.

For the fully mixed models (*scenario-1*) we assume that the water is homogeneously distributed in H–He in a vapor phase. In *scenario-2*, where the planet is assumed to be differentiated, any liquid water is assumed to be isothermal while ice is assumed to be adiabatic (e.g., Dorn et al. 2017). The stellar luminosity is assumed to be solar and the semimajor axis 0.1 au, corresponding to a temperature of ~ 770 K. The sensitivity of the inferred M–R relation to these assumptions is investigated in Sections 4.2 and 4.3. Further details on the structure models can be found in Venturini et al. (2015), Dorn et al. (2017), and references therein.

ORCID iDs

M. Lozovsky  <https://orcid.org/0000-0001-5451-3221>

R. Helled  <https://orcid.org/0000-0001-5555-2652>

References

- Adams, E. R., Seager, S., & Elkins-Tanton, L. 2008, *ApJ*, 673, 1160
 Baraffe, I., Alibert, Y., Chabrier, G., & Benz, W. 2006, *A&A*, 450, 1221
 Baraffe, I., Chabrier, G., & Barman, T. 2008, *A&A*, 482, 315
 Batygin, K., & Stevenson, D. J. 2013, *ApJL*, 769, L9
 Brouwers, M. G., Vazan, A., & Ormel, C. W. 2018, *A&A*, 611, A65
 Dorn, C., Khan, A., Heng, K., et al. 2015, *A&A*, 577, A83
 Dorn, C., Venturini, J., Khan, A., et al. 2017, *A&A*, 597, A37
 Fulton, B. J., & Petigura, E. A. 2018, arXiv:1805.01453
 Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, *AJ*, 154, 109
 Guillot, T. 2010, *A&A*, 520, A27
 Helled, R., Anderson, J. D., Podolak, M., & Schubert, G. 2011, *ApJ*, 726, 15
 Ikoma, M., Nakazawa, K., & Emori, H. 2000, *ApJ*, 537, 1013
 Jin, S., Mordasini, C., Parmentier, V., et al. 2014, *ApJ*, 795, 65
 Lopez, E. D., & Fortney, J. J. 2014, *ApJ*, 792, 1
 Marcus, R. A., Sasselov, D., Hernquist, L., & Stewart, S. T. 2010, *ApJL*, 712, L73
 More, R. M., Warren, K. H., Young, D. A., & Zimmerman, G. B. 1988, *PhFl*, 31, 3059
 Owen, J. E., & Wu, Y. 2017, *ApJ*, 847, 29
 Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, *Icar*, 124, 62
 Rauer, H., Catala, C., Aerts, C., et al. 2014, *ExA*, 38, 249
 Rogers, L. A. 2015, *ApJ*, 801, 41
 Rogers, L. A., & Seager, S. 2010, *ApJ*, 712, 974
 Saumon, D., Chabrier, G., & van Horn, H. M. 1995, *ApJS*, 99, 713
 Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007, *ApJ*, 669, 1279
 Swift, D. C., Eggert, J. H., Hicks, D. G., et al. 2012, *ApJ*, 744, 59
 Thompson, S. 1990, Sandia National Laboratories Report, ANEOS—Analytic Equations of State for Shock Physics Codes, SAND89-2951, <https://prod.sandia.gov/techlib-noauth/access-control.cgi/1989/892951.pdf>
 Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2018, *MNRAS*, 479, 4786
 Vazan, A., Helled, R., Podolak, M., & Kovetz, A. 2016, *ApJ*, 829, 118
 Vazan, A., Kovetz, A., Podolak, M., & Helled, R. 2013, *MNRAS*, 434, 3283
 Venturini, J., Alibert, Y., & Benz, W. 2016, *A&A*, 596, A90
 Venturini, J., Alibert, Y., Benz, W., & Ikoma, M. 2015, *A&A*, 576, A114
 Venturini, J., & Helled, R. 2017, *ApJ*, 848, 95
 Wagner, F., Sohl, F., Hussmann, H., Grott, M., & Rauer, H. 2011, *Icar*, 214, 366
 Weiss, L. M., & Marcy, G. W. 2014, *ApJL*, 783, L6
 Wolfgang, A., & Lopez, E. 2015, *ApJ*, 806, 183
 Zeng, L., & Sasselov, D. 2013, *PASP*, 125, 227
 Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, *ApJ*, 819, 127