The Radius Cliff is a Waterfall: Explaining Sub-Neptunes with Steam Worlds

Aritra Chakrabarty[1,2], **Gijs D. Mulders**[3], **Artyom Aguichine**[4], **Natalie Batalha**[4]

[1]*NASA Ames Research Center, California,* [2]*NASA Postdoctoral Program (NPP) Fellow*

[3]*Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Chile,*

[4]*University of California, Santa Cruz, California*

aritra.astrophysics@gmail.com

The Kepler Radius Valley problem: Evidence of two small-planet populations?

The bias-corrected radius histogram and period-radius distribution of Kepler short period (period < 100 days) exoplanets reveal a bimodality with a radius valley at 1.8-2 $R_\oplus$. This begs the question: do two distinct populations exist between 1 and 4 $R_\oplus$ – super-Earths and sub-Neptunes? Several competing theories attempt to explain this, including: i) atmospheric loss theory – bare rocky planets (without primary H/He atmosphere) vs. planets with H/He atmospheres, and ii) *ab initio* dichotomy in bulk composition – rocky core vs water-ice-rich core, also known as water worlds.

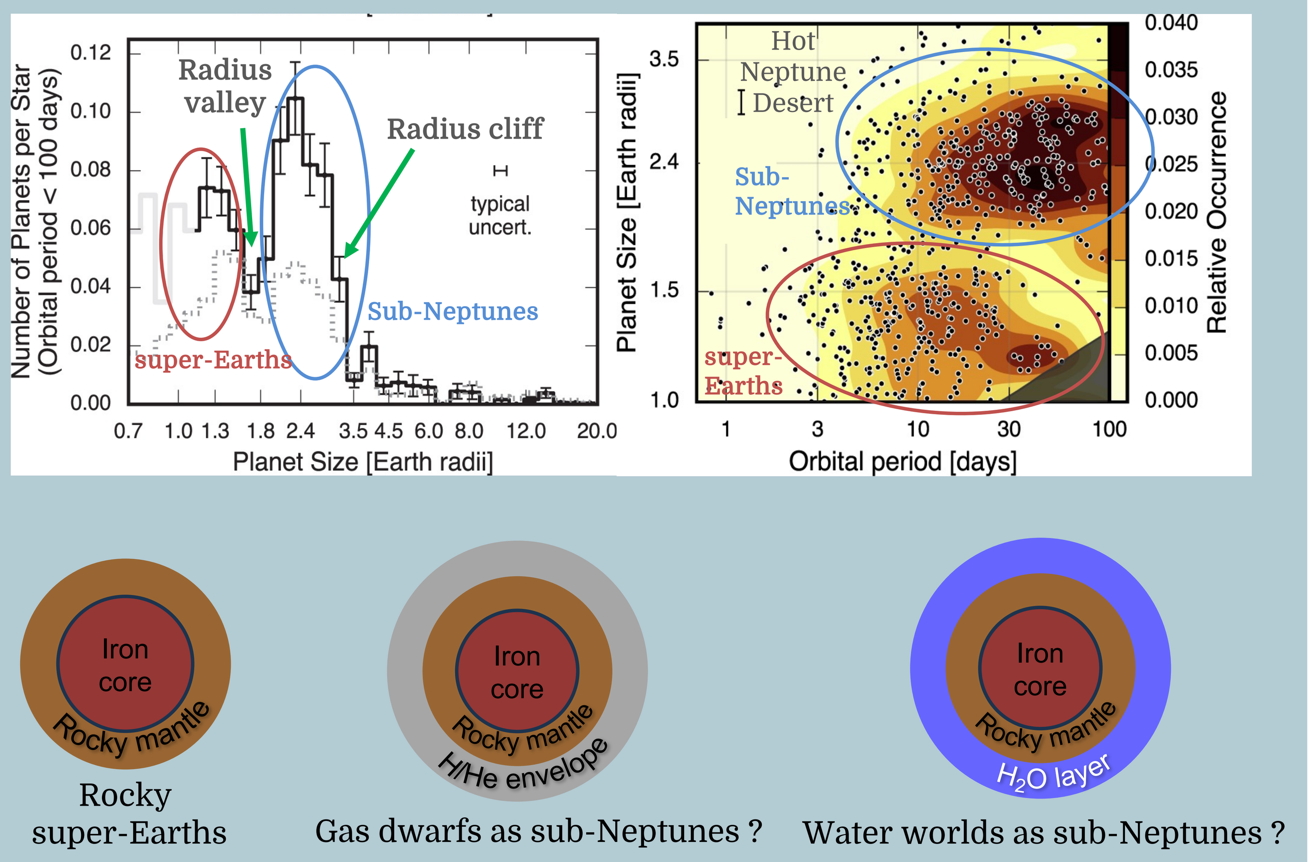


Figure 1 – Top: Bias-corrected radius histogram and period-radius distribution of Kepler planets from CKS catalog (Fulton et al. 2018), showing the radius valley, the radius cliff, and the hot Neptune Desert. Bottom: Possible interior structures of super-Earths and sub-Neptunes

Water Worlds & Steam Worlds

This work presents an agnostic investigation of the bulk composition dichotomy theory. It posits that the rocky planets make up the inner super-Earth population, while the sub-Neptunes are predominantly water-rich. Water-rich planets are thought to originate beyond the snowline and migrate inward into the Kepler-observed region via disk migration (Paardekooper et al. 2010; Izidoro et al. 2017; Mulders et al. 2020; Chakrabarty et al. 2024).

Previous models suggest that bare water worlds with condensed (isothermal) H_2O layers may lie along the radius valley due to their intermediate densities (Burn et al. 2024), or potentially even belong to the super-Earth population (Chakrabarty et al. 2024). In contrast, close-in water-rich planets are expected to possess adiabatic, supercritical H_2O envelopes, forming extended steam layers (Aguichine et al. 2021; Burn et al. 2024). These "steam worlds" are more inflated, making them viable candidates to explain the observed sub-Neptune population.

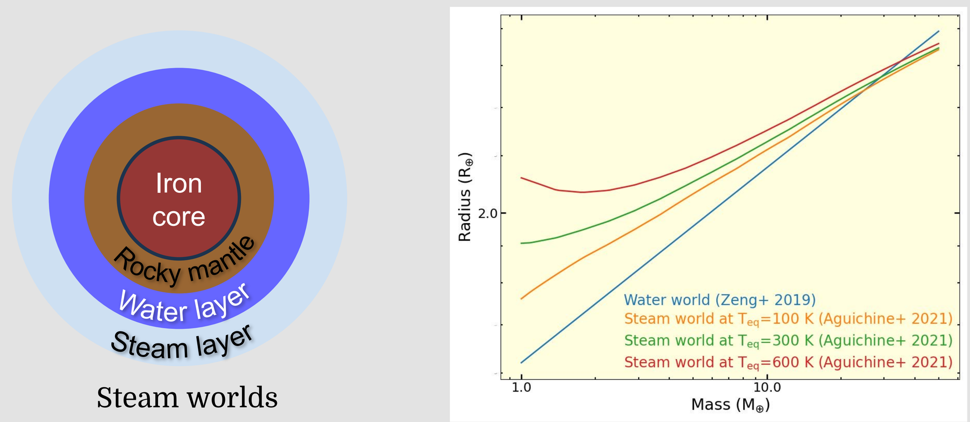


Figure 2 – Interior structure of steam worlds (left) and mass-radius relations of condensed water worlds vs. steam worlds (right)

A Bayesian Hierarchical Model

We perform Bayesian inference on the Kepler period–radius distribution using a two-population model **(Chakrabarty et al. 2025, in prep.)**: rocky planets (as super-Earths) and steam worlds (as sub-Neptunes).

- We adopt the observed Kepler period distributions of the super-Earths (for rocky planets) and sub-Neptunes (for water worlds) using the best-fit parametric approximation of Bergsten et al. (2022).
- Two distinct log-normal mass distributions are used for rocky planets and water worlds, along with a normal distribution for the water mass fraction (WMF) of water worlds.
- Radii are computed using the mass-radius relation of Zeng et al. (2019) for rocky planets and the model grid of Aguichine et al. (2021) for steam worlds.
- We calculate model occurrence rates in period–radius space of the Kepler planets and perform MCMC to maximize the likelihood of detecting the Kepler planets, considering it to be a Poisson point process (Bryson et al. 2020, Rogers et al. 2021)

Quantity	Distribution	Values
Mass of rocky planets	Log-normal	$\mu = 2.58 \pm 0.2$ $\sigma = 0.74 \pm 0.06$
Mass of water worlds	Log-normal	$\mu = 7.78 \pm 2.2$ $\sigma = 1.74 \pm 0.22$
WMF of water worlds	Normal	$\mu = 0.39 \pm 0.06$ $\sigma = 0.09 \pm 0.05$
Avg. occurrence rate of planets per star	--	1.3 ± 0.06

Table 1 – Selected model parameters estimated via MCMC **(Chakrabarty et al. 2025)**.

The fitted Radius histogram and period-radius distribution

The posterior radius and period-radius distributions of our model matches well with the observed distribution of the Kepler planets denoting that the steam worlds alone can well explain the sub-Neptune population. The radius valley is a result of the transition from rocky planets to water worlds across the radius space. We further notice a drop in the occurrence of the water worlds at ~4 $R_\oplus$, suggesting the "radius cliff" observed at ~4 $R_\oplus$ to be a result of this "waterfall".

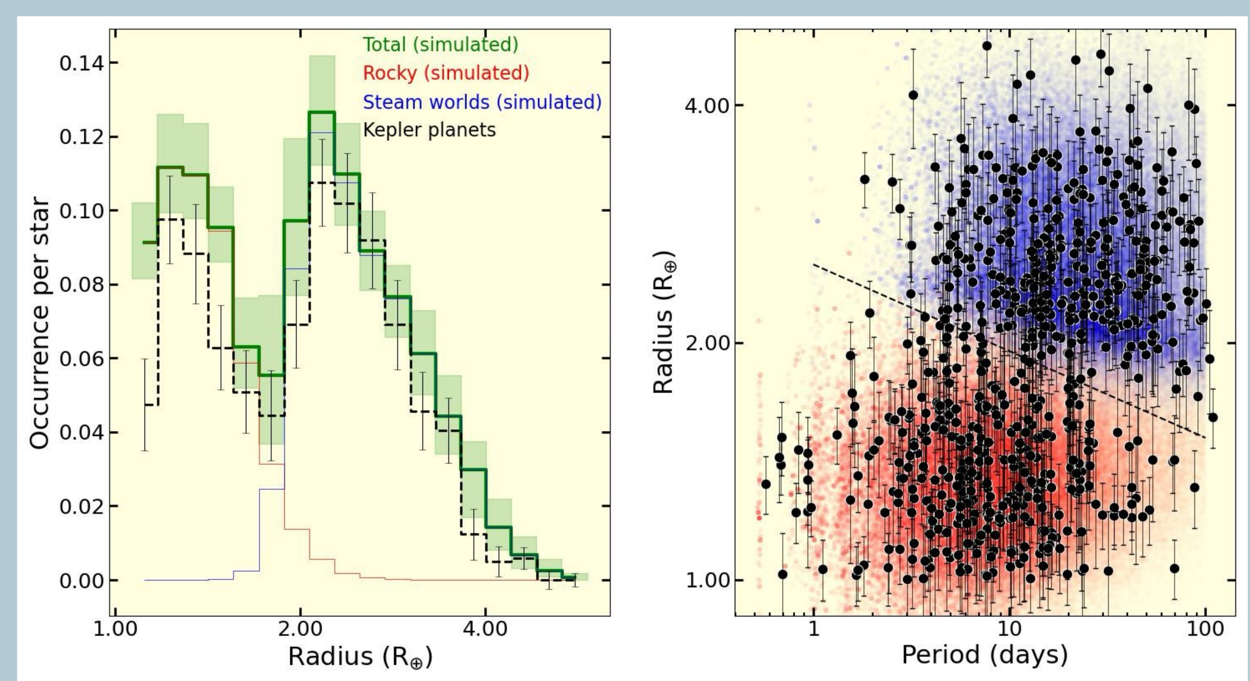
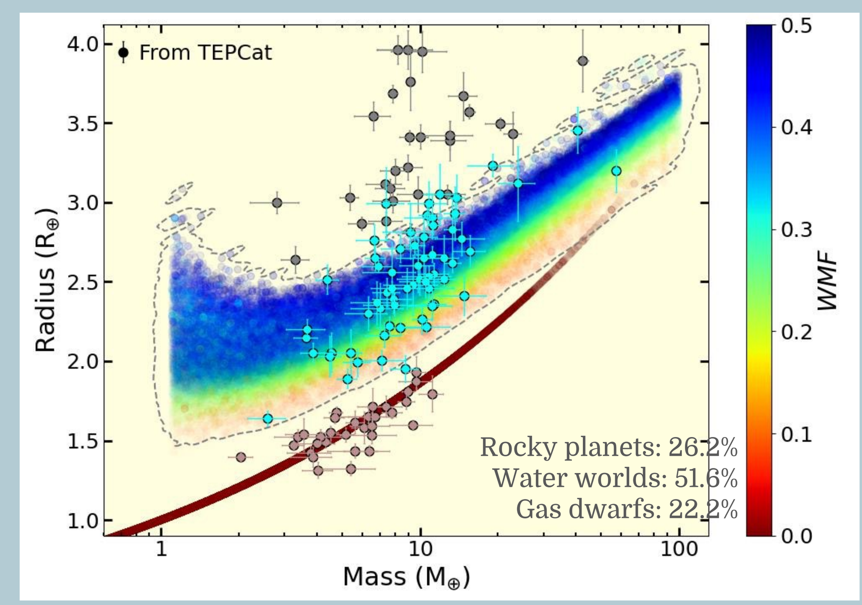


Figure 3 – Left:. The green solid line denotes the total occurrence rate corresponding to the median parameters, while the green shaded region indicates the 1-σ uncertainty range. Bottom: Period–radius distribution of the simulated planets, each of which is represent the full posterior sample of the free parameters, with each point weighted by the Kepler detection completeness and its posterior likelihood **(Chakrabarty et al. 2025).**

Mass-radius distribution: Hint at three small-planet populations?

We compare our model with the observed mass–radius distribution of small planets around G-type stars, selected from the TEPCat (Southworth et al. 2011) catalog with radii < 4 $R_\oplus$ and with mass and radius measurements accurate to better than 25% and 8%, respectively. Although the steam worlds can explain the radius distribution up to ~4 $R_\oplus$, they fail to reproduce the observed mass–radius relationships beyond ~3.5 $R_\oplus$, suggesting the presence of gas dwarfs with significant H/He envelopes.



TEPCat Catalog

Figure 4 – Mass–radius distribution of planets from our model compared with that of Kepler planets from the CKS catalog. The model points represent the full posterior sample of the free parameters, with each point weighted by the Kepler detection completeness and its posterior likelihood. The TEPCat planets are identified as likely rocky planets (brown error-bars), water worlds (cyan error-bars), and gas dwarfs (grey error-bars) by using the contour of the water worlds in the mass-radius plane **(Chakrabarty et al. 2025)**..

Conclusion

- A two-population model of rocky and steam worlds explains both the radius valley (transition) and the radius cliff (waterfall).
- The hot Neptune desert—a paucity of ultra-short-period sub-Neptunes and Neptunes—can be attributed to the scarcity of close-in water worlds. Formation models suggest this may happen in a migration scenario due to resonance chains with inner planets (Mulders et al. 2020), expansion of the magnetospheric cavity during disk dispersal (Liu et al. 2017), among other factors.
- While steam worlds reproduce the sub-Neptune population in the radius histogram and period–radius plane, the mass–radius distribution hints at a population of planets with H/He envelopes at $\gtrsim 3.5\ R_\oplus$. This underscores the need for multidimensional benchmarking, especially using a completeness-corrected mass–radius sample, which could substantially alter interpretations of population models.

References

1. Paardekooper, S. -J. et al. 2010, MNRAS, 401, 1950.
2. Izidoro, A. et al. 2017, MNRAS, 470, 1750.
3. Mulders, G. D. et al. 2020, ApJ, 897, 72.
4. Chakrabarty, A. & Mulders, G. D. 2024, ApJ, 966, 185.
5. Burn, R. et al. 2024, NatAs, 8, 463.
6. Aguichine, A. et al. 2021, ApJ, 914, 84.
7. Chakrabarty, A. et al. 2025, in preparation.
8. Bergsten G. J. et al. 2022, AJ, 164, 190.
9. Bryson S. et al. 2020, AJ, 159, 279.
10. Rogers J. G. & James, E. O. 2021, MNRAS, 503, 1526.
11. Southworth J. 2011, MNRAS, 417, 2166.
12. Liu B. 2017, A&A, 601, A15.

Link to GitHub: GPS (A related project)

Link to paper on GPS