The imprint of global magma oceans on exoplanet demographics

Martin Schlecker[1] and al.

[1] *Department of Astronomy/Steward Observatory, The University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA*

ABSTRACT

. . . magma oceans . . .

Here, we assess the ability of space and ground-based telescopes to test this hypothesis using `Bioverse`, a simulation framework that leverages contextual information from the overall planet population. . . .

1. INTRODUCTION

papers: (, introduce runaway GH) (Turbet et al. 2019, statistical detection of GH transition), (Lichtenberg et al. 2019, (initial) water content of terrestrial planets) (Luger & Barnes 2015, water loss, O2 buildup around M dwarfs), (???, Solubility of H in magma oceans) Hamano et al. (2013, 2015); Barth et al. (2021); Downey et al. (2022).

"Points for the intro: `Bioverse`
Transit survey to measure mass/radius of planets
Does the HZ have an inner edge?
Run-away GH
Magma oceans
Mention missions that are NOT statistical and therefore not considered here: CHEOPS, ARIEL, Roman. Other missions/concepts do not directly measure radii and not considered here: Habitable Worlds Observatory, LIFE.

In contrast, PLATO and *Nautilus* are
Possibility of follow-up to measureme masses "

For observations of rocky exoplanets, the currently best-probed regime is that of warm, close-in planets. These bodies experience climatic conditions that are similar to the environment of the inner Solar System bodies at early stages of their formation. Studies of the geophysical state and evolution of hot exoplanets can thus contribute to our understanding of the early formation stages of Earth and other habitable worlds.

... Venus and Earth, while having accreted from the same mass reservoir and despite their similar bulk properties, evolved into planets with very different environmental conditions on their surfaces. At the formation of the Solar System, both planets underwent a giant impact phase [CITE!] that melted their mantles, leading to a magma ocean stage. Due to these similar early formation phases, it was commonly assumed that the divergence of Venus and Earth – in particular Venus's water loss – occurred late in their evolution (Elkins-Tanton 2013). Hamano et al. (2013) showed that such a late desiccation process is not needed to explain the stark differences: for a planet rich in volatiles and receiving high enough radiation levels from its host star, a steam atmosphere limits the outgoing radiation from the (molten) planetary surface. This runaway greenhouse state prevents the planet's rapid cooling and can extend the magma ocean stage to hundreds of Myr, enough to remove the entire water reservoir from a rocky planet by hydrodynamic escape.

important for motivating the existence of the pattern: "the solidification timescale can become comparable to the main-sequence lifetime of the star (Hamano et al. 2013, 2015)" (Lichtenberg et al. 2022)

Consequently, it is aptly considered to be the inner boundary of the habitable zone (Kasting et al. 1993; Kopparapu et al. 2013).

(Under the right conditions, different orbital distances alone can thus decide about a planet's water content and ultimately its habitability.) If received flux makes the difference between our own habitable planet and dry, dead Venus, there is good reason to believe that such patterns exist in exoplanet systems as well.

...long-term goal/overarching objective: derive the geophysical history of rocky extrasolar planets. More concrete: Constrain the limits of runaway greenhouse transitions, and thus the inner edges of the habitable zone. Also: How close is Earth from this runaway greenhouse irradiation limit? ...

introduce magma oceans and their influence on planetary radii (Dorn & Lichtenberg 2021).

...current/future observations of planets that are currently in the runaway greenhouse phase may constrain properties of their planetary mantles... make connection between interior and atmospheres. ...

The runaway greenhouse transition is a robust prediction from climate models, and its impact on planetary bulk properties has been shown to be in the detectable range of current astronomical instrumentation (Goldblatt 2015). Measuring these impacts would indeed provide the first empirical evidence of the habitable zone concept. On the planetary level, however, such undertaking is hindered by degeneracies: From internal composition to atmospheric loss to volatile content, many factors influence a planet's size. It is therefore prudent to shift the focus to the planet population as a whole and search for a statistical effect in exoplanet demographics (Turbet et al. 2019). Using `Bioverse` to combine the predictions from geophysical models of steam atmospheres with a synthesis of planets informed by observed demographics, we...

Planetary radius changes due to the combined effect of steam atmospheres and water dissolved in the molten mantle of planets within the runaway greenhouse threshold is expected to cause a discontinuity in the radius distribution of small (DEFINITION) exoplanets.

2. BASELINE MODEL

The goal of this study is to determine – for different configurations of near-future exoplanet surveys – the confidence level with which the runaway greenhouse threshold can be detected statistically. Our basic methodology was as follows: We expanded the `Bioverse` framework (Bixel & Apai 2020, 2021)[1] to generate synthetic samples of stars that host planets according to the observed exoplanet demographics. We then adapted planetary bulk properties as predicted from models of runaway greenhouse atmospheres, simulated observations of the planets, and computed Bayesian evidences in favor of the runaway greenhouse hypothesis (see diagram in Fig. 1).

> make sure rgh hypothesis is mentioned before, or rephrase

In this section, we review the source of the stellar sample, the modeled luminosity evolution, the generation of a synthetic planet sample, as well as the orbital parameters of the planets. An overview of our key assumptions and model parameters can be found in Table 1.

2.1. *Stellar sample from Gaia DR3*

> KEVIN: update reference to my paper once accepted

The original `Bioverse` stellar catalog was generated randomly from the Chabrier (2003) stellar mass function. Exquisite parallax and photometric data from the *Gaia* mission made it possible to generate a homogeneous and complete stellar catalog out to about 100 pc, which became the new standard stellar catalog for `Bioverse` (Hardegree-Ullman et al., submitted). Here, we briefly describe how stellar parameters $T_{\rm eff}$, $L_\star$, $R_\star$, and $M_\star$ were derived.

Hardegree-Ullman et al. (submitted) used the Gaia Catalogue of Nearby Stars (hereafter GCNS, Smart et al. 2021) as the basis for deriving stellar parameters for the `Bioverse` catalog. The GCNS identified stars out to 120 pc, and includes *Gaia* DR3 parallaxes and photometry in G, G_{BP}, and G_{RP} bands (Brown et al. 2021) and K_S band photometry from 2MASS (Cutri et al. 2003), all of which were used in stellar classification. From this information, colors and absolute magnitudes were computed, and initial color-magnitude cuts

were made to remove non-main sequence stars. Effective temperatures were computed primarily from color-temperature relations derived using the main sequence stellar parameters table from Pecaut & Mamajek (2013). Next, luminosities were computed from absolute G-band magnitudes and a derived bolometric correction. Stellar radii were then computed with the effective temperatures and luminosities using the Stefan-Boltzmann law, or using absolute K_S-band magnitudes and an empirical radius-luminosity relation from Mann et al. (2015) for targets within the absolute magnitude range of M dwarfs. Finally, masses were derived from the mass-luminosity relation of Torres et al. (2010) for stars with $M_\star \gtrsim 0.7\,M_\odot$, and from that of Mann et al. (2019) for targets within the absolute magnitude range of M dwarfs. The derived stellar parameters from Hardegree-Ullman et al. (submitted) were compared to measured stellar parameters for all known exoplanet hosts from the literature and were found to be consistent within 1%, 3%, and 5.5% for $T_{\rm eff}$, $R_\star$, and $M_\star$, respectively, which are all below the typical measurement uncertainties of 3.3%, 6.8%, and 7.9%, respectively. From this catalog, `Bioverse` samples stars within an isotropic distance from the Solar System as required by the planetary sample size.

2.2. *Stellar luminosity evolution*

Planetary systems are hosted by stars of a wide range of ages, and stellar luminosities evolve with time. Since the emergence of a runaway greenhouse phase on a planet is highly dependent on the level of radiation it receives, and thus on the luminosity of the host star, we assigned age-dependent luminosities to our synthetic stars.

While stellar ages are notoriously poorly constrained (e.g., Adams et al. 2005), the age distribution of planet host stars in the Solar neighborhood was shown to be broadly consistent with uniform (Reid et al. 2007; Gaidos et al. 2023). For our synthetic stars, we thus drew random ages from a uniform distribution from $0\,{\rm Gyr}$ to $10\,{\rm Gyr}$. We then assigned each star a luminosity from the mass-dependent evolutionary models in Baraffe et al. (1998). Figure 2 shows the corresponding luminosity evolution as a function of stellar mass and age.

2.3. *Synthetic planet sample*

Next, we assigned to the stellar sample planetary systems with frequencies, orbital parameters, and bulk properties derived from the *Kepler* mission. We adopted the planetary occurrence rate model from Bergsten et al. (2022), which defines the occurrence of small planets

---

[1] `Bioverse` is actively maintained and documented open source software written in Python. Its latest version and documentation can be found at https://github.com/danielapai/bioverse.
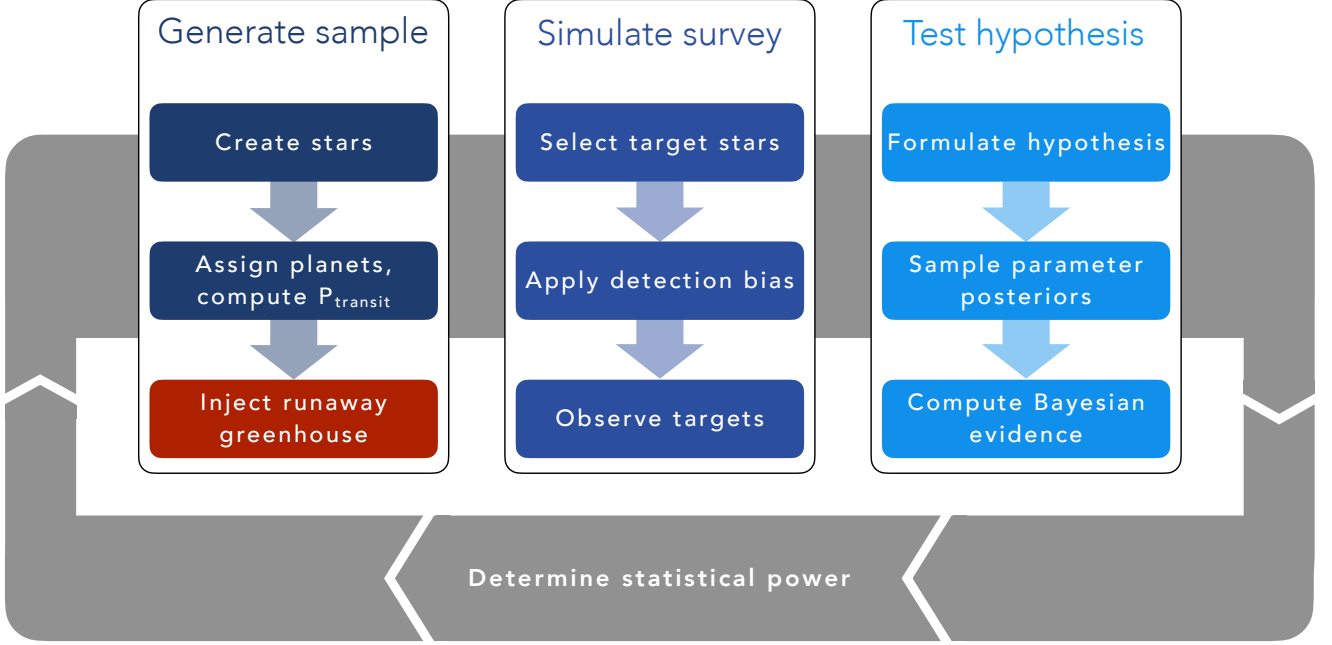
Figure 1. Workflow of our hypothesis testing with `Bioverse`. First, we generate a sample of stars and populate them with planets based on *Kepler* demographics. A fraction of them are then assigned a runaway greenhouse climate based on the model described in Sect. 3. We then simulate an exoplanet survey, whereby selection effects and detection biases are introduced. Finally, we test the runaway greenhouse hypothesis based on data from the survey simulation. By iterating through these steps, we compute the statistical power of testing the hypothesis for different survey designs.
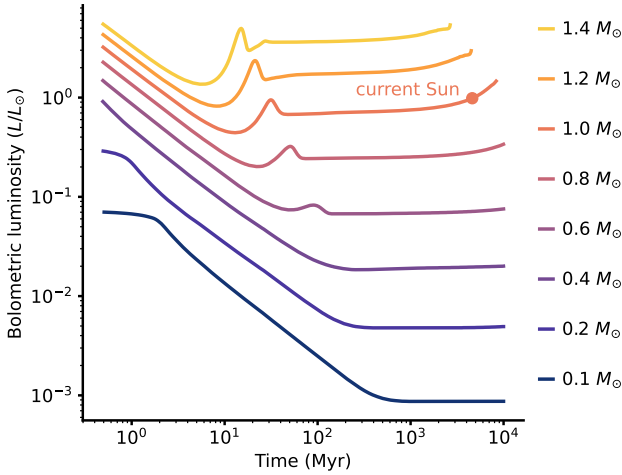
Figure 2. Bolometric luminosity tracks of stars with different masses, computed from stellar evolution models in Baraffe et al. (1998). Low-mass stars, which make up the majority of stars in the solar neighborhood, undergo an extended early phase of several magnitudes higher luminosity before entering a lifetime of relative faintness.

in radius and orbital period. Following Youdin (2011), their inferred occurrence rate density can be expressed in the form:

$$\frac{\mathrm{d}^2 n}{\mathrm{d}R\,\partial P} = F_0 C_n g(R, P), \tag{1}$$

where F_0 represents the average number of planets per star and C_n is a normalization constant. The shape function $g(R, P)$ describing the occurrence distribution is the same as in Bergsten et al. (2022). The functional form and stellar-mass dependent parameters (including a new fit to the sample of *Kepler* planets with hosts below $0.56\,M_\odot$) are included in `Bioverse` as the prominent planet-generation model; the original model based on SAG13 occurrence rates is preserved as a legacy feature.

With this updated distribution function, planets are generated as in the original `Bioverse` paper, but with an additional stellar mass dependence (for a star of a given mass, the relevant model parameters from the bin containing that mass are used). In summary, stars are assigned a number of planets in accordance with F_0, with planet parameters randomly drawn in proportion to the occurrence distribution function evaluated over a grid of radii and orbital periods.

2.4. *Orbit parameters and planet masses*

Eccentric orbits alter the probability of a planet to transit (e.g., Barnes 2007). The distribution of eccentricities e of exoplanets has been found to resemble a Beta function (Kipping 2013), which we chose to draw synthetic eccentricities from. Following Kipping (2013), we used a Beta distribution with parameters $a = 0.867$ and $b = 3.03$, and truncated the distribution at $e = 0.8$. Assuming isotropic alignments of orbits, we assigned each planet an inclination drawn from a distribution uniform in $\cos(i)$.

To assign masses to our planets, we use the semi-empirical mass-radius relationship assuming a pure $MgSiO_3$ composition from Zeng et al. (2016) (see green line in Fig. 3). This represents the baseline bulk density before any climate-related effects are applied.

2.5. *Transit probability*

We model the occurrence of transits by assuming isotropic orientations of planetary orbits and calculating the impact parameters $b = \cos(i)/R_\star$. Following the approach in Bixel & Apai (2021), we further consider only planets with $|b| < 1$. For these cases we calculate the transit depth

$$\delta = \left(\frac{R_P}{R_\star}\right)^2, \qquad (2)$$

which is relevant for the detection probability of the respective planet (see Sect. 4.1). Excluding all non-transiting planets diminishes the sample to 1.5% of its original size.

3. RUNAWAY GREENHOUSE MODEL

TIM: please review

The climate state of a planet has a direct influence on its apparent size measured by transit photometry (Turbet et al. 2019). We use predictions on transit atmospheric thickness from geophysical models to derive the change in transit radius and bulk density that planets with installation-induced runaway greenhouse climates experience. ...

ARNAUD: revise definition of S following our discussion on 2022-12-22. Introduce net installation and clarify what is meant by it.

The power absorbed by the planet from the host star per unit area at its orbital distance or net installation S in units of Earth's insolation is given by

$$\frac{S}{S_\oplus} = \left(\frac{L_\star}{L_\odot}\right)\left(\frac{au}{a}\right)^2. \qquad (3)$$

We assume global redistribution of incoming flux and a wavelength and spherically integrated albedo, or Bond albedo, of $0.3$, comparable to Earth's (Haar & Suomi 1971). The absolute value of $S_\oplus$, by which all synthetic planets are scaled, is then $238\,\mathrm{W\,m^{-2}}$. We do not take into account additional heating sources such as tidal effects (e.g. Barnes et al. 2013).

While we search for the runaway greenhouse pattern in demographic quantities such as average planet radii, the injected changes happen on the planetary level: We changed each planet's radius based on its individual set of properties and the associated predictions from steam atmosphere and water incorporation models. Relevant properties are a planet's mass M, its net installation S, and its bulk water inventory expressed as a water mass fraction x_{H_2O}. We consider the following cases (see Fig. 3):

Non-runaway planets retain the radius assigned based on exoplanet occurrence rates (see Sect. 2.3). This case serves as our null hypothesis.

For all planets receiving a dayside-averaged installation exceeding a threshold of $S_{\mathrm{thresh}} = 280\,\mathrm{W\,m^{-2}}$, we assume an inflated radius due to a steam atmosphere. While the actual installation threshold for a runaway climate depends on planetary albedo, surface gravity, and clouds (Pierrehumbert 2022), this value was found to be a typical limit for the flux a planet can emit in a runaway greenhouse situation (Abe & Matsui 1988; Kasting 1988; Nakajima et al. 1992; Goldblatt et al. 2013; Leconte et al. 2013; Boukrouche et al. 2021). To quantify the radius change, we applied the mass-radius relationships derived by Turbet et al. (2020) using a 1D inverse radiative-convective model (Turbet et al. 2019). Their calculations rely on the same mass-radius relations for rocky interiors that we apply for our non-runaway planets (Zeng et al. 2016). For each planet above the installation threshold, we assigned the predicted radius for the given water mass fraction and planet mass. This scenario does not take into account the expected additional radius change due to molten interiors (Bower et al. 2019). It further assumes a *dry melt* without dissolved volatiles.

Wet melt planets follow the rules above, but we take into account a radius decrease from water partitioning in the melt: In the situation of a magma ocean below a steam atmosphere, part of the water will partition into the melt. This removes it from the steam atmosphere, which in turn decreases the radius inflation. The effect is generally small compared to the radius inflation from a steam atmosphere but resulting radius changes are predicted to reach measurable values (Dorn & Lichtenberg 2021). The magnitude of this effect is again dependent on the planet mass and on the water mass fraction. We follow Dorn & Lichtenberg (2021) and use their computed radius deviations between a wet magma ocean and
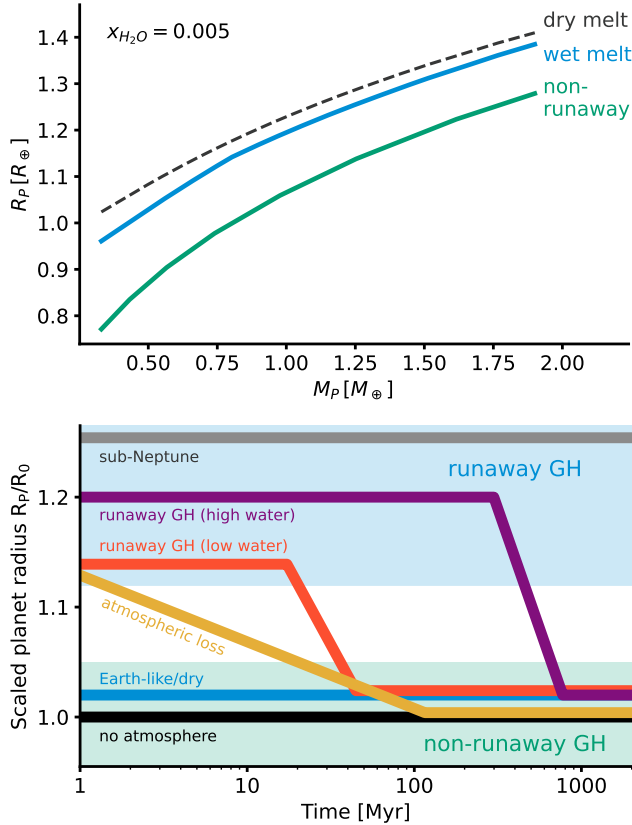
Figure 3. *Top*: Mass-radius relationships for a water mass fraction $x_{H_2O} = 0.005$ and different planet states. Green: planets with a solid mantle and no steam atmosphere. Dashed: planets with steam atmospheres. Blue: planets with steam atmospheres and including the effect of water incorporation in the melt. Steam atmospheres cause a significant radius increase, which is slightly reduced when water incorporation in the melt is considered. *Bottom*: Radius evolution of different planet types, illustrating degeneracies and potential for confusion among planet classes. Shown is a schematic time evolution of the transit radius normalized to the atmosphere-free radius R_0 for different scenarios. Planets can move between planet classes through processes such as atmospheric loss and desiccation, which ultimately ends a runaway greenhouse phase on a timescale dependent on a planet's water content.

a solid mantle for a tropopause pressure $P_{\mathrm{iso}} = 0.1\,\mathrm{bar}$. We then add the (in almost all cases negative) radius deviations to the planet radii computed for the dry melt case.

We illustrate the mass-radius relations of the three cases in Figure 3 where a fiducial water mass fraction of $x_{H_2O} = 0.005$ is assumed. In the following, we only distinguish between the non-runaway greenhouse and wet melt scenarios.

To account for planets unable to sustain a steam atmosphere as well as evolutionary effects such as desic-

cation (see bottem panel of Fig. 3), we introduce a dilution parameter f_{rgh}. It represents the fraction of planets above the threshold installation that are currently inflated due to a runaway greenhouse climate. Our simulation setup is such that all planets receiving a net installation $S < S_{\mathrm{thresh}}$ follow the non-runaway greenhouse relation, and a fraction f_{rgh} of the planets with $S > S_{\mathrm{thresh}}$ follow the wet melt relation. In the following, we test if and under what conditions this parametrization causes a demographic trend that is large enough to be detected with high significance.

4. EXOPLANET SURVEY SIMULATIONS AND HYPOTHESIS TESTING

The survey module of `Bioverse` converts the synthetic planet sample into a set of uncertainty-laden measurements on a subset of that sample. This task includes selection of the targets, application of detection biases, and conducting simulated measurements, all of which are specific to the particular survey. For each planet-level measurement such as radius or installation, we draw the measured value from a normal distribution centered on the true value with a standard deviation set by the survey's precision. We then follow a Bayesian hypothesis testing approach to assess various realizations of simulated surveys in terms of their ability to detect and characterize the runaway greenhouse transition.

4.1. *Detection bias, target selection, and sensitivity*

Not all transiting planets are detectable with the same likelihood and detection biases have an impact on the demographic measurements we are interested in. A detailed characterization of the detection biases of individual missions would not be justifiable given the uncertainties of the theoretical predictions. Instead, we derived generic observing limits that reflect the limitations of state-of-the-art transit surveys.

A successful transit detection requires a sufficient signal-to-noise ratio (S/N), which is sensitive on the achieved photometric precision. *PLATO* (PLAnetary Transits and Oscillation of stars) is an ESA mission designed to characterize terrestrial planets in the habitable zones of Sun-like stars, a goal to be achieved via long-term high-precision photometric monitoring of a sample of bright stars (Rauer et al. 2016). In line with this requirement, *PLATO* is designed to enable the detection of a 80 ppm transit signal (ESA 2017, Matuszewski et al., in prep.). To reflect its sensitivity, we chose a minimum transit depth of 80 ppm as a detection limit and consider only measurements of planets exceeding this threshold. We further exclude target stars with *Gaia* magnitudes $M_{\mathrm{G}} > 16$.

The runaway greenhouse effect becomes obsolete both for very small instellations and where no atmosphere can be maintained due to proximity to the host star and resulting atmospheric erosion. Ensuring to stay well clear of such regions, we clear our sample from all planets with a net instellation $S < 10\,\mathrm{W\,m^{-2}}$ or $S > 2000\,\mathrm{W\,m^{-2}}$. We further consider only rocky planets with masses below $2\,\mathrm{M_\oplus}$.

4.2. *Measurements and their uncertainties*

Under real-world conditions, the planetary properties in question can only be probed with a finite precision that is specific to each exoplanet mission. *PLATO*'s definition study report (ESA 2017) states precision requirements for planet radii ($3\,\%$), planet masses through radial velocity (RV) follow-up ($10\,\%$), and stellar masses, radii, and ages ($10\,\%$). We adopted these estimates and assumed a $10\,\%$ error on instellation measurements.

Since planetary bulk density $\rho \propto R_\mathrm{P}^{-3}$, we expect a stronger runaway greenhouse signal when measured through bulk density instead of radius. We thus simulated measurements of planetary densities assuming the mass-radius relation defined above. For uncertainties in bulk density measurements, we propagated the errors of the mass measurements assuming $\sigma_{\mathrm{M_P}} = 10\,\%$.

4.3. *Hypothesis tests*

We now turn to quantifying the ability of the simulated surveys to detect the runaway greenhouse transition and to constrain parameters of this trend. To do this, we rely on a Bayesian hypothesis testing approach where we quantify the evidence of a hypothesis over another based on the (simulated) data. For our specific problem, this implies comparing evidences for a demographic imprint of the runaway greenhouse effect to its absence. As a null hypothesis, we consider the case where the planetary radius distribution is independent of the instellation,

$$H_0(\theta, S) = \theta, \qquad (4)$$

where θ is the set of parameters defining the radius distribution. We further define an alternative hypothesis that describes radius changes due to steam atmospheres and water in the melt. As motivated above, this hypothesis takes the form of a step function in net instellation S, where the step occurs at the outer edge of the runaway greenhouse region. Our main observable shall be the average planet radius in the planet population inside and outside this threshold. The runaway greenhouse hypothesis is then defined as

$$H_\mathrm{rgh}(\theta, S) = \begin{cases} H_0, & S \le S_\mathrm{thresh} \\ \langle R_\mathrm{P}\rangle(f_\mathrm{rgh}, \Delta R_\mathrm{stm}, \Delta R_\mathrm{wtr}), & S > S_\mathrm{thresh}. \end{cases} \qquad (5)$$

Here, f_rgh is the fraction of planets experiencing a runaway greenhouse effect. ΔR_stm and ΔR_wtr are predicted radius changes from the steam atmosphere and water incorporation models, respectively. They are assumed to act additively on the planet radii and thus on their average $\langle R_\mathrm{P}\rangle$.

The only free parameter of the null hypothesis, which assumes the average radius to be independent of instellation, is the predicted mean radius $\langle R_\mathrm{P}\rangle$. The functional form of the runaway greenhouse hypothesis is more complex: Besides the mean radius of planets outside the threshold $\langle R_\mathrm{P}\rangle_\mathrm{out}$, which is a nuisance parameter necessary to define the hypothesis, it relies on the threshold instellation for the "step" S_thresh, the planetary water mass fraction x_{H_2O}, and the dilution factor f_rgh. For hypothesis tests based on bulk density instead of radius, we proceeded in the same way and substituted R_P by the bulk density ρ.

A sensible choice of priors is central for evidence estimation via nested sampling. As the parameters of interest are poorly constrained by previous data, we used relatively uninformative priors to sample the entire physically plausible parameter space. For S_thresh, we chose a uniform prior in $[10, 1000]\,\mathrm{W\,m^{-2}}$. We sampled x_{H_2O} from a log-uniform distribution to imply scale-invariant ignorance. Its boundaries $[10^{-5}, 0.1]$ are motivated by the water mass fractions covered by the geophysical models (Sect. 3). For f_rgh, we chose a uniform prior in $[0, 1]$. Finally, we adopted a broad, uniform prior for $\langle R_\mathrm{P}\rangle_\mathrm{out}$ bound by $[0.1, 15]\,\mathrm{R_\oplus}$. In the case of measuring bulk densities instead of radii, we drew uniformly from $[1, 6]\,\mathrm{g\,cm^{-3}}$.

The measured radii $R_\mathrm{P,i}$ or bulk densities ρ_i cannot be directly used for the hypothesis tests as they include intrinsic scatter that is not caused by measurement errors. H_rgh and H_0 should thus be tested against a statistical estimator that represents the population mean. To avoid binning and the artificial patterns it may introduce, we chose to test our hypotheses against a simple moving average SMA along the instellation axis with a window of size 25 centered around each measurement.[2] We further computed the uncertainty of this moving average

---

[2] See Appendix A.1 for a robustness test using a different estimator.

by propagating the individual measurement errors and applying a rolling standard error of the mean.

As our procedure involves random sampling of the model parameters θ, we need to define the probability of obtaining a dataset given the model parameters, i.e., a likelihood function. We assumed here that the individual moving averages SMA_i are measured with a normally distributed uncertainty σ_{SMA_i} and adopted a normal distribution

$$\mathcal{L}(SMA \mid \boldsymbol{\theta}) = \prod_i^N \frac{1}{\sqrt{2\pi\sigma_{SMA_i}^2}} \quad (6)$$

$$\times \exp\left(-\frac{(SMA_i - H(\boldsymbol{\theta}, S_i))^2}{2\sigma_{SMA_i}^2}\right). \quad (7)$$

Here, $H(\boldsymbol{\theta}, S_i)$ corresponds to the functional form of the runaway greenhouse or null hypothesis.

4.4. *Bayesian model comparison*

We are now ready to assess the relative plausibility of H_{rgh} and H_0 given the synthetic data we have generated, assigning equal a priori probabilities to these models. This is done by comparing the Bayesian evidence $\mathcal{Z}$ of the models, which we estimated with the nested sampling (Skilling 2004) algorithm *dynesty* (Speagle 2020). We initialized the sampler with the priors defined above to let it estimate the evidence and sample the posterior distributions. Our criterion to reject the null hypothesis is

$$\Delta \ln \mathcal{Z} = \ln \mathcal{Z}_{\mathrm{rgh}} - \ln \mathcal{Z}_0 > 3. \quad (8)$$

5. RESULTS

5.1. *Statistical signature of the runaway greenhouse threshold*

To characterize the population-level imprint of individual radius changes, we generated a generic planet population with an injected runaway greenhouse effect assuming a water fraction $x_{\mathrm{H_2O}} = 0.005$. Figure 4 shows the resulting planetary radii. When ordered in orbital period space, the different planet types overlap, diluting the demographic imprint. With net installation as an independent variable, planets above and below the runaway greenhouse threshold separate: The runaway greenhouse-induced radius inflation introduces a discontinuity of average planet radii and bulk densities as a function of stellar irradiation. We also show the predictions of observable average planet radii from the statistical hypotheses defined above. Within the runaway greenhouse regime, an average radius change of 15 % occurs. This pattern is consistent with the injected radius
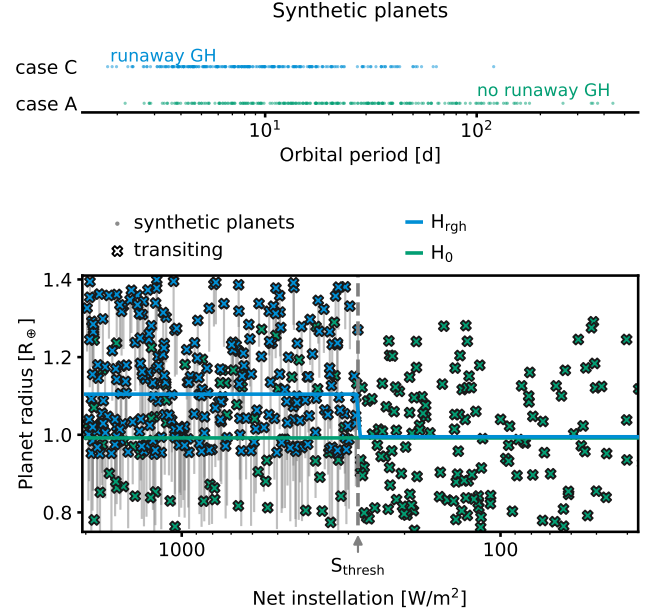
Figure 4. Synthetic planets above and below the runaway greenhouse threshold. *Top*: Planet state as a function of orbital period. Planets with and without a runaway greenhouse climate mix and are not distinguishable in period space. *Bottom*: Radii of synthetic planets with injected radius deviation as a function of net installation. Only the planets marked as transiting are observable. Above the runaway greenhouse threshold $S_{\mathrm{thresh}} = 280\,\mathrm{W\,m^{-2}}$, some planets maintain their original radii (green crosses) whereas some have their radii inflated (blue crosses) by the amount indicated with gray lines. The sharp boundary at S_{thresh} causes a discontinuity in the average planet radius (blue line). This runaway greenhouse hypothesis can be tested against the null hypothesis H_0 (green line), where average radii are independent of installation.

inflation as predicted from the atmospheric models (see Appendix A.2 for an investigation of the interplay between model predictions and our synthetic planet population).

5.2. *Testability of the runaway greenhouse hypothesis*

Figure 5 shows a prototypical statistical detection of the runaway greenhouse transition. For the sake of clarity, we chose an optimistic scenario where the sample of characterized planets is large ($N = 497$) and the measurement uncertainties are small ($\sigma_R = 2\,\%, \sigma_S = 5\,\%$). We assumed that the fraction of those planets irradiated stronger than S_{thresh} that have runaway greenhouse climates is 0.8, and we chose a water mass fraction of 0.005 for each planet. In this case, the pattern was detected with high significance ($\Delta \ln \mathcal{Z} \approx 100$).

With such a strong signal, we can attempt an inference of the parameters defining the injected effect. Figure 6 shows the posterior distributions of S_{thresh}, x_{H_2O}, and

Table 1. Key assumptions and model parameters used in our simulation setup

Parameter	Value	Unit	Description
Stellar sample			
G_{max}	16		Maximum Gaia magnitude
$M_{\star,max}$	1.5	$M_\odot$	Maximum stellar mass
Luminosity evolution			Baraffe et al. (1998)
Planetary parameters			
M_P	$0.1 - 2.0$	$M_\oplus$	Planetary mass range
$R_{P,min}$	0.75	$R_\oplus$	Minimum planet radius
Baseline mass-radius relation			Zeng et al. (2016) 100 % $MgSiO_3$[a]
δ_{min}	80	ppm	Minimum transit depth
P_{max}	500	d	Maximum orbital period [d]
S	$10 - 2000$	$W\,m^{-2}$	Net instellation range
S_{thresh}	280	$W\,m^{-2}$	Threshold instellation for runaway greenhouse
Runaway greenhouse model			
Runaway greenhouse atmospheric models			Turbet et al. (2020); Dorn & Lichtenberg (2021)
x_{H_2O}	10^{-5}–0.1		Water mass fraction (fiducial case: 0.005)
f_{rgh}	0–1		Dilution factor (fiducial case: 0.8)
Priors			
$\Pi(S_{thresh})$	[10, 1000]	$W\,m^{-2}$	uniform
$\Pi(x_{H_2O})$	$[10^{-5}, 0.1]$		log-uniform
$\Pi(f_{rgh})$	[0, 1]		uniform
$\Pi(\langle R_P \rangle_{out})$	[0, 15]	$R_\oplus$	Mean radius of non-runaway planets, uniform

[a]For a comparison with alternative interior compositions, see Appendix A.3.

f_{rgh} as determined by the nested sampler. The threshold instellation can be accurately constrained with a median $\hat{S}_{thresh} = 276\,W\,m^{-2}$. Both a higher water mass fraction and a higher dilution factor lead to larger average radii, thus these parameters are strongly correlated.

Figure 7 explores the statistical power of the hypothesis test achieved in the above scenario for different combinations of the poorly constrained parameters x_{H_2O} and f_{rgh}. It is highest for large water inventories and large dilution factors. For all but very low water fractions, f_{rgh} dominates this trend: It enters linearly into the average planet radius, whereas the contribution of x_{H_2O} - as predicted by the geophysical models - is sublinear with a power-law exponent of ~ 0.3. As long as f_{rgh} is larger than ~ 0.2, a sample size of 497 is sufficient for a 50 % detection rate even for water ratios as low as 10^{-3}.

5.3. *Detecting the runaway greenhouse transition with PLATO*

For a volume-limited sample, considering all host star spectral types, we find that a yield of 288 planets enables a significant detection if the fraction of runaway greenhouse planets is larger than ~ 0.1 (see Fig. 8). The minimum needed fraction rises to 0.2 for $N = 100$. For much smaller samples, only a very optimistically strong signal is likely to be detected.

5.4. *Statistical power of different mission designs*
5.4.1. *Additional planet mass measurements*

Comparing a measurement of the signal in radius space with a measurement in density space (which requires planetary mass measurements), a stronger detection occurs in the latter case: With an optimistic choice of geophysical parameters (see above), the average measured radius change is 15 % whereas the average density change is -33%.

As can be seen in Fig. 8, the detectability is enhanced when radius constraints can be augmented with mass
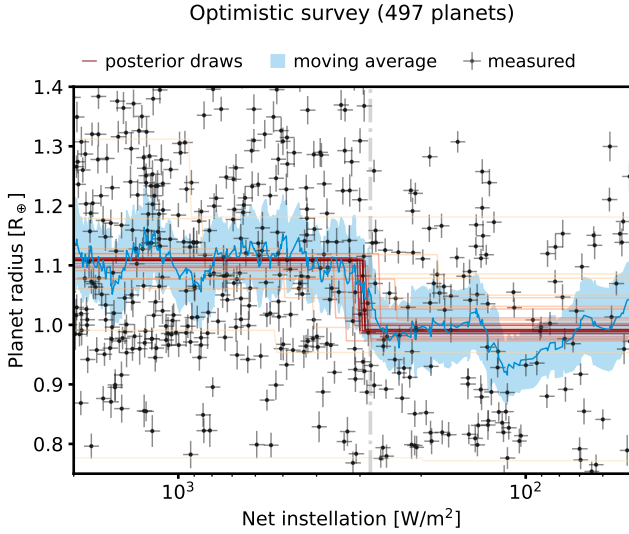
**Figure 5.** Detection of the runaway greenhouse threshold in the optimistic case. From simulated radius and instellation measurements of a large ($N = 497$) survey, we compute the moving average (blue confidence intervals) and fit the runaway greenhouse hypothesis to it (Eqn. 5, random draws from the posterior in red). The pattern is detected with high significance.

measurements: At unchanged sample size, the difference in Bayesian evidence can be up to an order of magnitude larger. Consequently, we achieve a statistically significant detection with smaller samples or lower dilution factors $f_{\rm rgh}$. A density-based survey of 100 targets is roughly equivalent to a radius-based survey of 288 targets. At $N = 100$, pure radius measurements require $f_{\rm rgh} \gtrsim 0.2$ whereas bulk density measurements enable a detection from $f_{\rm rgh} \gtrsim 0.1$.

5.4.2. *Dependence on host star spectral type*

Since the receiving radiation is sensitive on the spectral type of the host star, the share of planets on either side of $S_{\rm thresh}$ is different for FGK and M dwarfs. We tested the detectability of the runaway greenhouse signal when only FGK or only M dwarfs are considered (see Fig. 8). The samples are volume and magnitude-limited to reflect the target counts of *PLATO*'s provisional Long-duration Observation Phase fields (15996 FGK stars in the P1 and P2 samples, 33948 M stars in the P4 sample, Nascimbeni et al. 2022). The resulting M dwarf planet sample is significantly larger with 228 ± 14 planets compared to 40 ± 6 planets in the FGK sample. No significant detection is possible in the pure FGK sample, independent of the assumed geophysical parameters. In the M dwarf sample, the evidence threshold is reached around $f_{\rm rgh} \sim 0.2$, similar to the case above where all spectral types are considered.
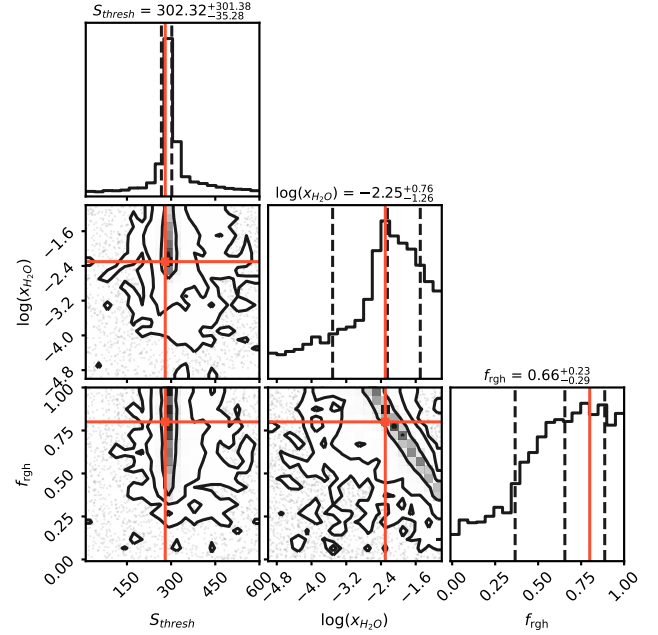

Figure 6. Retrieved posterior distribution of key parameters in the optimistic scenario. The density maps in each panel show relationships between and marginalized distributions of the threshold instellation $S_{\rm thresh}$, the water mass fraction x_{H_2O}, and the dilution factor $f_{\rm rgh}$ as they could be retrieved with a high-precision transit survey and a sample of 497 planets. True values of the parameters for the injected effect are shown in orange. The threshold instellation can be reasonably constrained; the predominant water fraction and the fraction of planets with runaway greenhouse climates are degenerate.

5.4.3. *Constraining the threshold instellation*

A key constraint resulting from a successful detection is a measurement of S_{thresh}. Here, we assess the ability of different mission concepts explored above to constrain this parameter. Figure 9 shows posterior distributions from a grid of inferences together with the true values of the injected signal for different fractions $f_{\rm rgh}$. We consider three cases: only radius measurements, radius and mass measurements, and radius and mass measurements of only planets orbiting M dwarfs. The simulations otherwise represent the simulated PLATO example described above. We chose a planetary sample size of $N = 100 \pm 10$, as this was found to be a threshold case in Sect. 5.3.

We find that retrievals from radius measurements alone require high fractions of greenhouse climate-bearing planets to achieve an accurate constraint on $S_{\rm thresh}$: Dilution factors $f_{\rm rgh} \gtrsim 0.5$ yield posterior probability distributions that are condensed at the order of magnitude of the true value; accurate constraints to
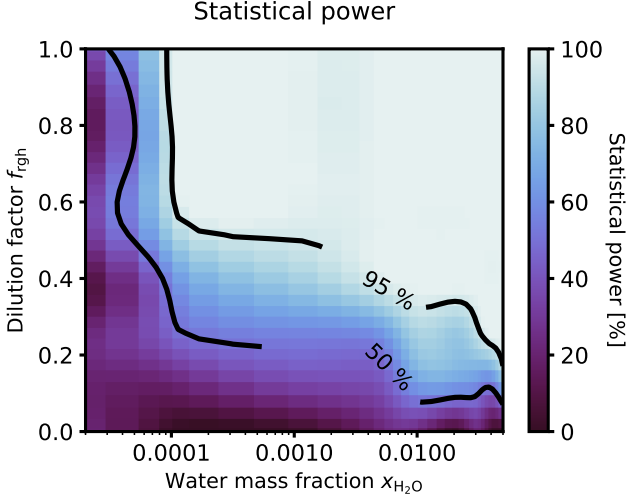
Statistical power

Figure 7. Statistical power of the runaway greenhouse hypothesis test as a function of model parameters. For a sample size $N = 497$, the color code shows the fraction of simulations resulting in a sound detection ($\Delta \ln \mathcal{Z} > 3$) for different combinations of water mass fraction and dilution factor. Higher values in either parameter result in a more reliable detection. For water mass fractions $\gtrsim 10^{-4}$, the statistical power largely depends on the fraction of greenhouse climate planets in the sample.

within ± 0.25 dex of the truth are reached only from $f_{\mathrm{rgh}} \approx 0.8$.

In contrast, if planet masses are available and the hypothesis test is conducted in bulk density space, useful constraints emerge already from about $f_{\mathrm{rgh}} \approx 0.35$. Accuracy and precision of the retrievals are improved.

A sample containing only planets around M dwarf yields still better performance and result in the overall best accuracy and precision. The reason is that these planets contribute most to the statistical power of the retrieval.

6. DISCUSSION

> **Potential additional discussion points:**
> (How close is Earth to runaway? (-> Goldblatt & Watson (2012); Popp et al. (2016); what is the max. expected change in Earth's albedo due to deglaciation?)

6.1. *Statistical imprint of exoplanet climates*

We showed in Sect. 5.1 that injecting the theoretically predicted radius inflation effect into a synthetic planet population following the currently known demographics leaves a distinct pattern in the radius and density distribution as a function of instellation. Such a statistical imprints of the runaway greenhouse transition was suggested before (Turbet et al. 2019), and our simulations show quantitatively how the contributions of the individual radius changes combine to create a significant demographic signal. Its strength depends on largely un-

known factors such as the planetary volatile content, but turned out to be well in the detectable range under reasonable assumptions of these factors.

Our finding that the expected trend is strong except for extremely low water mass fractions gives reason for optimism regarding its detection. It also presents a potential for constraints on the water inventory of terrestrial planets in the case of a non-detection. Such constraints will shed light on whether the initial water content is variable (Lichtenberg et al. 2019) or there is a predominant quantity. Ultimately, such measurements will shed light on the question of if planets like Earth received their water from primordial sources during their formation or if it was supplied later (e.g., Raymond et al. 2004). Overall, searching for the runaway greenhouse discontinuity appears to be the most promising approach toward a first empirical test of the habitable zone concept.

6.2. *Detectability of the runaway greenhouse transition*

We showed in Sect. 5.2 that under favorable conditions, a sufficiently large (497 planets) photometric survey is likely to detect the demographic imprint of the runaway greenhouse transition and accurately constrain the associated threshold instellation. However, even assuming capable instrumentation and an optimal survey design, geophysical and demographic unknowns may jeopardize this undertaking. The detectability of the runaway greenhouse threshold is thus a function of both survey design and astrophysical circumstances.

6.2.1. *Key factors influencing tests of the runaway greenhouse hypothesis*

What influences the probability of correctly rejecting a false null hypothesis? We identified the following drivers of the diagnostic power for detecting the runaway greenhouse transition with a transit survey.
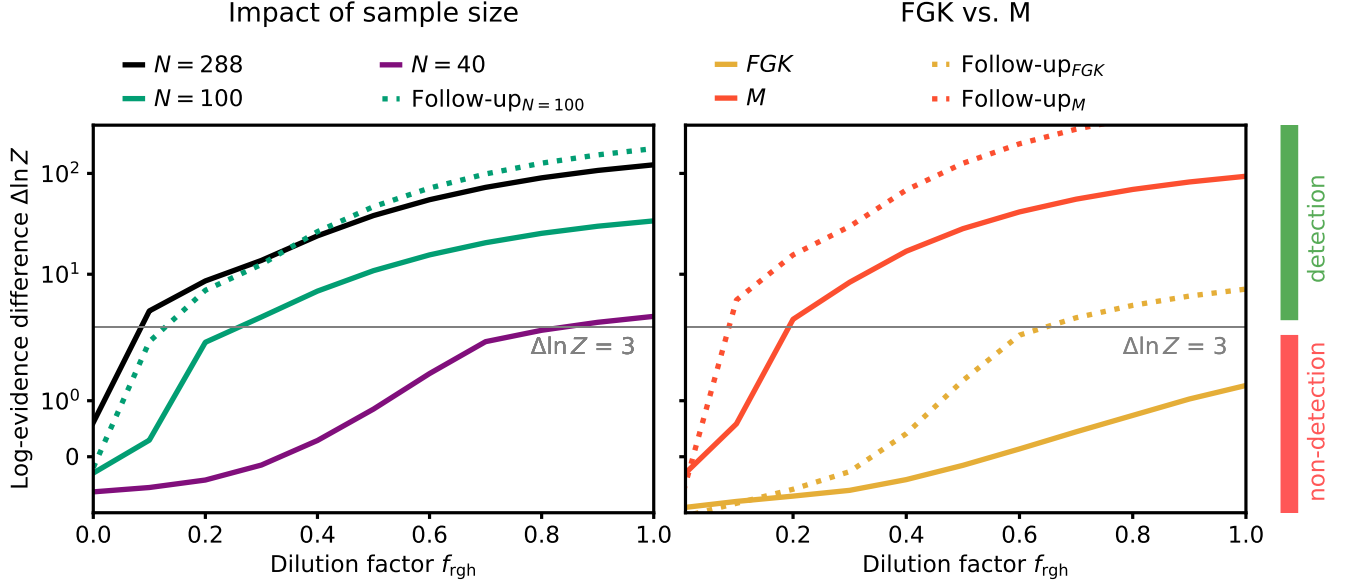
Occurrence rate of planets forming steam atmospheres: The runaway greenhouse climate relies on the production of large amounts of water vapor that then acts as a greenhouse gas. This process works only on planets with an atmosphere and of sufficiently volatile-rich composition

> TIM: how much?

. The fraction of planets fulfilling these requirements has an impact on the amplitude of the demographic imprint of the runaway greenhouse transition.

Planetary evolution and duration of the steam atmosphere phase: An inflated steam atmosphere can only be sustained until the planet has lost its water. Depending on host star spectral type, planetary mass, and composition, planets can spend from a few Myr to several Gyr in runaway greenhouse climates (Hamano et al.

Impact of sample size

- $N = 288$
- $N = 100$
- $N = 40$
- Follow-up$_{N=100}$

FGK vs. M

- FGK
- M
- Follow-up$_{FGK}$
- Follow-up$_{M}$

Figure 8. Expected delta-evidences as a function of the fraction of planets with runaway greenhouse climates for different versions of the *PLATO* survey. The median values of randomized survey simulations are shown; $\Delta \ln Z > 3$ (gray horizontal line) is considered sufficient evidence to reject the null hypothesis. *Left:* For a large planet yield of $N = 288$, even small dilution factors ~ 0.1 allow a detection. A sample of 100 planets is sufficient if their masses are constrained to within $10\,\%$ (dotted green line). Without such follow-up measurements, sufficient diagnostic power can only be achieved with this sample if $f_{\rm rgh} \gtrsim 0.2$. Even smaller samples are unlikely to yield a significant detection. *Right:* Evidences when only FGK or only M dwarfs are considered. Only M dwarfs host enough planets on both sides of the threshold instellation to allow a reliable detection of the runaway greenhouse signal.

2015; Luger & Barnes 2015), and only planets observed during this phase will contribute to the signal.

Prevalent water inventory: The magnitude of radius change at the runaway greenhouse threshold is sensitive on the water mass fraction. As a result, the statistical abundance of water in terrestrial planets impacts the strength of the demographic pattern: The higher the water content and the higher the fraction of planets in runaway greenhouse climates, the greater the prospect of a statistically sound detection.

Size and composition of the planetary sample: Evidently, the significance of a statistical trend increases with a larger sample size. In addition, the sample must include planets on both sides of the instellation threshold, which, due to the strongly distance-dependent detection bias associated with the transit method, is only sufficiently given for low-mass host stars.

Radius measurement precision: The more precise individual planet radii can be determined, the less smeared out the pattern will be. Good *accuracy* is less important, as long as it does not have a systematic error scaling with stellar irradiance.

Availability and precision of mass measurements: For simple geometric reasons ($\rho \propto R^{-3}$), the expected trend is stronger when measured in bulk density than it is in planet radius space. If transiting planets can be fol-

lowed up to obtain mass measurements, the statistical significance increases.

Besides these main factors, uncertainties in the measured instellations can influence the result, although they are typically small due to the very precise orbital period measurements available for transiting planets. This can be different for young host stars when their ages can not be well constrained; in particular the long pre-main sequence phase of M dwarfs shows a large variation in bolometric luminosity (see Fig. 2).

6.2.2. *False positive scenarios*

Runaway greenhouse climates are not the only physical mechanism that may cause a change in transit radius for a subset of planets. Some of the most prominent alternative scenarios involve atmospheric loss due to either photoevaporation through high-energy radiation by the host star (e.g., Owen & Wu 2013; Jin et al. 2014; Mordasini 2020) or due to residual heat from the planet's interior shortly after formation (Ginzburg et al. 2016, 2018; Gupta & Schlichting 2019). Both processes are being traded as potentially sculpting the observed bimodality in the radius distribution of small exoplanets (Fulton et al. 2017; Van Eylen et al. 2018). Either of them leads to a decrease of planet radius for planets
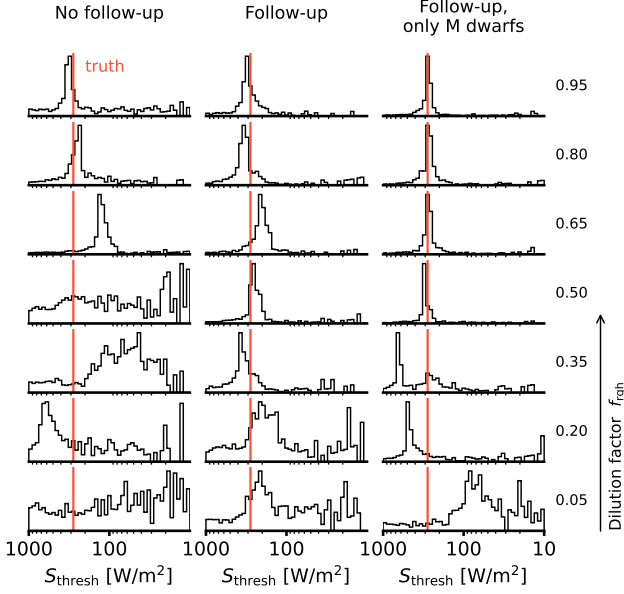
Figure 9. Retrieved posterior distributions of the threshold installation for different survey realizations. All cases assume $x_{H_2O} = 0.005$ and a planet sample size $N = 100 \pm 10$; The fraction of planets with runaway greenhouse climates varies across rows. Orange lines show the true value of the injected signal. Accuracy and precision of the constraint on S_{thresh} generally improve with higher f_{rgh}. Bulk density-based inferences improve the constraints, and M dwarf samples yield the highest accuracy and precision.

close to their host star, which is distinct from the radius inflation introduced by runaway greenhouse climates.

Another false positive contribution may come from potential unknown variations in the occurrence rate gradients in radius-installation space. They can impact the statistical inference of the transitions, especially if these variations are similar to the expected pattern. Although an abrupt pattern at the expected runaway greenhouse transition seems unlikely, examples of steep occurrence rate density changes exist. An example is the "Neptune desert", a triangular region of low planet occurrence density of close-in planets in period-radius space (Szabó & Kiss 2011; Mazeh et al. 2016; Dreizler et al. 2020). The shape of this region is such that smaller planets become less frequent the closer to the star they are, which to some degree resembles the pattern introduced by the installation dependency of the runaway greenhouse transition. However, the Neptune desert occurs at smaller orbital periods and its location depends on a planet's size (Szabó & Kiss 2011), which is not expected for the runaway greenhouse transition.

Luque & Pallé (2022) found that small planets around orbiting red dwarfs can be classified into three density regimes with a particularly strong separation between planets consistent with a pure rocky and those consistent with a water-rich composition. This trend does not represent a false positive scenario for the runaway greenhouse transition, since no strong dependency on installation has been found or is expected. However, a population of "water worlds" with low bulk densities on a wide range of orbits would dilute the statistical runaway greenhouse imprint. If the dichotomy forms primordially through migrated planets that accreted from different regions of their protoplanetary disk (**?**Burn et al. 2021; Schlecker et al. 2021a,b), systems of all ages can be affected by this dilution.

6.2.3. *OPTIONAL...: Atmospheric signatures*

> TIM: Ideas?
> discuss potential atmospheric signatures of magma oceans. e.g.: H/O ratios of sub-Neptunes could be low (because H is in the melt) (e.g. Tim's recent talk)

> "Can a smaller sample size's diagnostic value be increased if atmospheric compositional information can be added, i.e., PLATO + ARIEL (or similar combo)?"

> also: use atmospheric characterization to identify small planet cores with thick H/He envelopes that contribute to diluting the signal (aka background population). These will be amenable for transmission spec due to large scale heights (Moran2018, Turbet et al. (2019)).
> Spectral water features << rgh-induced radius change (Turbet et al. 2019, end of Sect. 4).

6.2.4. *Detection in multi-planet systems*

As suggested by Turbet et al. (2019), an alternative approach for detecting the runaway greenhouse-induced radius inflation is to search for its "local" imprints in multi-planet systems. Systems harboring planets on both sides of the transition, such as TRAPPIST-1 (Gillon et al. 2016, 2017; Luger et al. 2017; Agol et al. 2021), may show the predicted abrupt radius and density change, provided the initial volatile content was sufficient and complete desiccation has not yet occurred. While degeneracies remain in interpreting bulk density fluctuations within individual systems (e.g., Turbet et al. 2020; Dorn & Lichtenberg 2021), the detection of a consistent pattern in several such systems could be a convincing statistical evidence of the runaway greenhouse transition. The current sample of suitable systems is sparse: The California-Kepler Survey catalog (Fulton & Petigura 2018) contains only six planets with installations $< 2\,S_\oplus$ and smaller than $2\,R_\oplus$ in five systems. Future additions to the multi-planet sample through missions such as *PLATO* are needed.

6.3. *Diagnostic power of near-future exoplanet missions*

The fidelity of a future detection or falsification of a runaway greenhouse signature will depend on the significance with which the null hypothesis can be excluded. As a consequence, instrumentation and survey strategy aiming at probing the runaway greenhouse threshold

should aim at maximizing the probability of a true positive detection.

As discussed above, key drivers from a mission design perspective are sample size, photometric precision, and the availability of planets around low-mass host stars. We found that along these axes, *PLATO* will be the most favorable among the upcoming transit missions. The *PLATO* team has released an estimate on the number of exoplanets that will be characterized in the course of the main survey mission (Matuszewski, priv. comm.). With an expected transit radius precision of 3 % (ESA 2017) and a yield of > 100 planets studied with such high precision (Rauer 2021), *PLATO* is comparable to the optimistic survey we introduced in Sect. 5.2. If successful, it should readily detect the predicted statistical imprint or, in case of a non-detection, provide strong upper limits on the occurrence rate of runaway greenhouse planets. The latter depends on the lifetimes of runaway greenhouse phases, which are a function of the initial water inventory of the planets (Hamano et al. 2015). Overall, it seems feasible to derive the typical water content of rocky planets from these occurrence estimates.

What other planned missions are suited to test the runaway greenhouse hypothesis? *Kepler* and *K2* have contributed a large number of discovered terrestrial-sized planets, but few of them are in the habitable zone and their host stars are typically too faint for RV follow-up (Dressing & Charbonneau 2015).

Similarly, the Transiting Exoplanet Survey Satellite (*TESS*, Ricker et al. 2014) planet sample lacks temperate, small planets around bright host stars (Ment & Charbonneau 2023), as was expected from planet yield calculations (Barclay et al. 2018). As of 2023 March 22, the NASA Exoplanet Archive[3] lists 40 *TESS* candidate or confirmed planets smaller than $4\,R_\oplus$ with lower estimated instellation than Earth's.

The ongoing *CHEOPS* mission was designed as a follow-up mission to search for transits of planets discovered with other techniques, in particular with radial velocity measurements (Benz et al. 2021). As such, it will provide precise radius constraints on a sample of small planets; however, only a small number of planets with periods >50 d are being observed. This largely limits *CHEOPS*' coverage to planets within the runaway greenhouse regime, preventing a detection of the transition.

As *CHEOPS*, the Atmospheric Remote sensing Infrared Exoplanet Large survey (*Ariel*, Puig et al. 2016)

---

[3] https://exoplanetarchive.ipac.caltech.edu

will be a follow-up mission that is not designed to provide a large number of new radius measurements. *Ariel*'s primary targets are larger planets in the range of sub-Neptune to Jupiter-like planets. We thus do not expect a significant contribution to statistically exploring the inner edge of the habitable zone for Earth-sized planets.

While not primarily designed to detect transiting planets, the Galactic Bulge Time Domain Survey of the *Nancy Grace Roman Space Telescope* (Spergel et al. 2015) is expected to yield $\sim 10^5$ transiting planets on short orbits and constrain their radii in the course of its mission (Montet et al. 2017). $\mathcal{O}(1000)$ planets smaller than Neptune could be found around early to mid-M dwarfs, however, only a small fraction of them will reach into the habitable zone (Tamburo et al. 2023). We thus conclude that the *Nancy Grace Roman Space Telescope* could provide a useful sample to explore the runaway greenhouse transition, albeit with a predominant focus on water-rich (sub-)Neptunes (e.g., Pierrehumbert 2022).

Looking further ahead, the *Nautilus* Space Observatory concept (Apai et al. 2019) represents a statistical mission able to provide precise radius measurements of a large sample of small exoplanets. It employs a swarm of large telescopes using ultralight diffractive-refractive optical elements (Milster et al. 2020) with the primary goal to study the atmospheres of transiting exoplanets. Operating in an array mode, *Nautilus* would achieve the equivalent light-collecting area of a 50 m telescope. Its expected 1 ppm photometric precision (Apai et al. 2022) would enable precise radius measurements of a large sample, also through a low number of required visits per object. If realized, *Nautilus* will be a valuable instrument for characterizing the runaway greenhouse transition.

Other mission concepts exist that focus on characterizing exoplanet habitability, most notably the *Habitable Worlds Observatory* (Gaudi et al. 2020; **?**) and the Large Interferometer For Exoplanets (*LIFE*, Quanz et al. 2021). Such direct imaging surveys, however, do not directly measure planetary radii and are therefore most useful for providing context through atmospheric measurements of individual planets.

TIM: feel free to add something about LIFE if you see fit

As for exoplanets missions in their implementation phase, *PLATO* overall remains to be the most promising mission for an empirical confirmation or falsification of the runaway greenhouse transition at this time.

6.4. *Mission design trades*

To explore the impact of mission trade-offs on the detectability of the runaway greenhouse transition, we sim-

ulated different survey designs and strategies and measured their capability to recover the trend and constrain its parameters. We assessed this capability based on two determinants: the likelihood that the mission is able to detect the injected trend, and the precision with which it can constrain the parameters of that trend.

6.4.1. *The value of follow-up campaigns*

The dependencies of a detection change when additional information beyond planet radii is available for the characterized planet population. As runaway greenhouse phases leave a stronger imprint in bulk density than in planet radius (see Sect. 5.4.1), it would be beneficial to obtain constraints on planetary masses and test the runaway greenhouse hypothesis in density space instead of radius space. This way, useful results can be obtained under more pessimistic conditions, e.g., a low predominant water content of planetary surfaces and atmospheres or a smaller available planet sample. For a mission design similar to PLATO, a density-based hypothesis test on about a third of the overall sample is equivalent to a pure radius-based analysis. At a fixed sample size, key parameters of the runaway greenhouse models can be more narrowly constrained when additional mass measurements are available.

Precise ground-based radial velocity measurements will be needed to provide these data, and a number of instruments are already successfully employed in characterizing terrestrial-sized exoplanets (e.g., Queloz et al. 2001; Pepe et al. 2010; Johnson et al. 2010; Ribas et al. 2023). Recently, the National Aeronautics and Space Administration (NASA) and the National Science Foundation (NSF) commissioned an "Extreme Precision Radial Velocity (EPRV) Initiative" (Crass et al. 2021) to develop methods and facilities for precise mass measurements of temperate terrestrial planets. Their findings highlight that such measurements are costly, and therefore follow-up efforts may only be available for a subsample of the targets of a mission of *PLATO*'s scale. The diagnostic power of the hypothesis tests we demonstrated here may be improved by simultaneously fitting for the trend in radius space in the subsample without RV follow-up. An optimized mission in search for the inner edge of the habitable zone will further enhance its information content via a wise selection of follow-up targets, i.e., balancing objects located on either side of the expected installation threshold S_{thresh}.

6.4.2. *The importance of M dwarfs in the target list*

To date, the majority of planets with radius measurements orbit FGK dwarfs, and, based on the installation they receive, most of them lie in the runaway greenhouse regime (Thompson et al. 2018). Obviously, a step-

function-like signal in the exoplanet demographics like the one investigated above can not be constrained well if only one side of the discontinuity is being sampled. This, however, is the situation for planetary systems around Sun-like stars – their habitable zones are so distant that transiting planets within them are very rare due to pure geometrical reasons. It was, among other reasons, the sharp drop in transit probability with orbital distance that has prompted a number of recent transit surveys to specifically target M dwarfs (e.g., Irwin et al. 2009; Obermeier et al. 2016; Delrez et al. 2018; Sebastian et al. 2021; Dietrich et al. 2023), but the sample of terrestrial planets orbiting them is still small (e.g., Berger et al. 2020; Hardegree-Ullman et al. 2020).

M dwarf systems are also key for detecting the runaway greenhouse transition: Our calculations with different spectral types (Sect. 5.4.2) show that the information content of M dwarfs in a sample dominates the hypothesis tests. Besides their large number in a volume/magnitude-limited sample, transiting M dwarf planets are more likely to be located near the threshold installation and in particular on orbits further out, i.e., in the optimistic habitable zone. In fact, we showed that the FGK part of the planet sample barely contributes to the statistical power. Furthermore, the transit depth difference at the transition is expected to be larger for M dwarfs ($\sim 100\,\mathrm{ppm}$ for early, $\sim 1000\,\mathrm{ppm}$ for late M stars, Turbet et al. 2019), enhancing the demographic signal it leaves. An additional advantage of targeting M dwarfs are their extended runaway greenhouse phases, whose duration can reach the order of $100\,\mathrm{Myr}$ (Luger & Barnes 2015). This increases the probability of observing any given planet in the sample during the runaway greenhouse phase, essentially driving f_{rgh} to higher values. Therefore, in addition to the high scientific value of boosted detections of potentially habitable planets, M dwarfs are also indispensable for the discovery and characterization of the runaway greenhouse transition. As with a pure volume-limited sample, a targeted M dwarf survey, too, profits from follow-up measurements of planetary masses with an order of magnitude increase in evidence.

6.5. *Constraining planetary habitability*

A potential for liquid water on the surface of a planet is commonly used as an environmental marker to assess its habitability (Huang 1959; Hart 1978; Kasting et al. 1993; Kaltenegger & Sasselov 2011; Kopparapu et al. 2013). The runaway greenhouse transition represents an upper bound on received irradiation for this condition. Its detection would thus not only empirically confirm the habitable zone concept but also help to locate

it in the observationally available planetary parameter space. In Sect. 5.4.3, we show that the threshold instellation at which the runaway greenhouse transition occurs can be reasonably constrained without imposing overly optimistic conditions on the underlying planet population, instrumentation, or survey strategy. A mission like *PLATO* is well equipped to perform this measurement; the chances for success are directly proportional to the proportion of characterized planets around M dwarfs and to the number of planets for which masses can be determined.

The situation is different for the planetary water inventory and the fraction of planets with runaway greenhouse climates: Since these parameters are degenerate, they cannot be well constrained without independent measurements. This degeneracy could be lifted if independent measurements of atmospheric compositions can be made. For example, detections of water vapor in planets above the threshold instellation, combined with precise radius measurements, would constrain the predominant water content of terrestrial planets. A follow-up campaign with the *Ariel* mission of *PLATO* planets could be considered to make this contribution.

Once a runaway greenhouse region is identified in the parameter space, the community will will have a tool at hand to discern potentially habitable planets from Venusian worlds on an empirical basis. Together with atmospheric measurements (see Sect. 6.2.3), we will be able to put a number on the probability of an individual planet to harbor sufficient surface water to sustain life.

6.6. *Impact of assumptions on our findings*

> Caveats
> - not yet explored: how sensitive are we on the assumed planet population?
> ∃ danger of false positive detection on occurrence rate "ridges" like the Neptune desert? Currently we have a relatively smooth occurrence rate density in $R_P - S$ space based on the power-law fits in Bergsten et al. (2022). We could, with relatively moderate effort, run a sanity check on an alternative sample (e.g., the more structured Bern model populations, or the population in the Genesis database)

The prospects for detecting the runaway greenhouse transition depends on astro- and geophysical factors, as well as on the specific instrumentation and survey strategy of a particular mission. We can only approximate all this here, and we therefore seek to assess the sensitivity of our results on the assumptions we have made.

6.6.1. *Structures in the planet occurrence rate density*

In principle, the baseline occurrence rate density in radius-period space that governs the generation of our synthetic planet sample might influence our findings, especially if it contains any features that coincide with the injected demographic feature. This is not the case in the model from Bergsten et al. (2022) that we adopted: Its occurrence rate density varies smoothly in the domain relevant for the runaway greenhouse hypothesis; transi-

tions only occur at smaller instellations ($< 50\,\mathrm{W\,m^{-2}}$) and larger radii ($> 1.6\,R_\oplus$). We thus do not expect the model underlying our planet sample to affect the recovery of the transition. However, it should be kept in mind that any unknown sharp features in the true distribution of terrestrial planets could complicate a detection or lead to a false positive result (see Sect. 6.2.2).

6.6.2. *Baseline mass-radius relationship*

Our baseline mass-radius relationship assuming pure MgSiO$_3$ interiors (Zeng et al. 2016) might not be representative of the rocky planet population. However, while interior composition may change the magnitude of radius inflation and could introduce an offset to its demographic imprint, we do not expect a change of its structure. The magnitude of the radius inflation effect is in fact expected to be larger for an Earth-like interior composition with an iron core-silicate mantle structure.

> TIM: do you have a reference for this statement?

We thus consider our mass-radius relationship a conservative case and do not expect a negative impact on the detectability of the runaway greenhouse transition due to deviations from it. Nevertheless, we performed a sanity check to assess the influence varying our baseline model on our results (see Appendix A.3), and found general agreement between different interior composition models.

The models in this study do not into account a predicted additional radius increase on the percent level due to a molten interior (Bower et al. 2019), and we neglect any potential effects stemming from a deviating gas exchange between atmosphere and interior in runaway greenhouse planets (Ikoma et al. 2018). Future theoretical work that self-consistently models interior-atmosphere interactions is needed to take these effects into account (Turbet et al. 2019).

6.6.3. *Simplifications in the atmosphere model*

> TIM: your expertise will be valuable here.
> Mention we only consider instellation-induced rgh.
> - we ignore water loss by H$_2$O photolysis in the upper atmosphere and subsequent H escape, which would eventually cool even planets within the runaway greenhouse regime (Lichtenberg et al. 2022).
> - we ignore tidal heating, which could extend the magma ocean phase of close-in planets (and change their orbits via tidal orbital evolution)(Jackson et al., 2008; Barnes et al., 2013 (?)).
> - assumption of a single albedo. "if there is a range of albedos, will that wash away the signal?"

As illustrated in Fig. 3, evolutionary processes and a background population of planets unable to sustain a steam atmosphere complicate the classification of a planet into the runaway or non-runaway category. For instance, both a Hydrogen/Helium-dominated "sub-Neptune" and a terrestrial planet currently hosting a steam atmosphere will appear as a planet in our runaway greenhouse category, diluting the expected demographic signature. Further, a planet in a runaway green-

house state will eventually lose its steam atmosphere and move to the non-runaway greenhouse category on a timescale of 10^8 yr for a planet with one Earth ocean of water (Watson et al. 1981). We marginalized over these degeneracies by introducing the dilution factor $f_{\rm rgh}$, but in a real survey contextual information about a planet's age or likely formation history could be taken into account.

6.6.4. *Water mass fraction*

The predominant mass fractions of water, which sensitively controls the atmospheric state of a rocky exoplanet, is poorly constrained. Inferred water contents in the literature range from upper limits on the order 10^{-5} to "water worlds" with tens of percent mass fraction (e.g., Rogers & Seager 2010; Unterborn et al. 2018; **?**; Agol et al. 2021; Luque & Pallé 2022), all of which are within the realm of theoretical predictions (Mulders et al. 2015; Sato et al. 2016; Jin & Mordasini 2018; Bitsch et al. 2019; **?**; Emsenhuber et al. 2021; Schlecker et al. 2021a). Our nominal case assumes a water mass fraction of $x_{\rm H_2O} = 0.005$. This can be considered a conservative choice that is unlikely to introduce a systematic overestimation of the statistical radius inflation trend, also since the hypothesis tests are only marginally sensitive on $x_{\rm H_2O}$ as long as it is $\gtrsim 10^{-4}$ (see Sect. 5.2). Cases of pure rocky composition and very low volatile contents can be considered absorbed by the dilution factor $f_{\rm rgh}$. Assuming a distribution of water mass fractions instead of a fixed value would thus not significantly change our results.

6.6.5. *"Sharpness" of the demographic trend*

The real counterpart to our idealized step-function central to our hypothesis tests could be blurred by several factors. Its amplitude is affected by "sub-Neptunes" with significant Hydrogen/Helium envelopes, which can disguise as inflated rocky planets. They might not contribute to the demographic signal we are looking for, although a similar transition was suggested for water-dominated sub-Neptunes (Pierrehumbert 2022). A subset of such gas-rich planets will experience atmospheric loss via photoevaporation (Owen & Wu 2013) or core-powered mass loss (Ginzburg et al. 2018), turning them from planets appearing inflated into much smaller (super-)Earths. Further dilution comes from the population of planets that lack an atmosphere or sufficient volatiles that can create a steam atmosphere. There is also a time dependence, both because stellar luminosities change with age and because runaway greenhouse planets with limited water reservoirs ultimately desiccate, ending their steam atmosphere phase.

Intrinsic variation in the threshold installation is caused by differences in planetary features influencing the onset of a runaway climate such as albedo, atmospheric composition, or surface gravity (Pierrehumbert 2022). The choice of statistical estimator for the hypothesis tests may further blur the recovered discontinuity. We compare our nominal running mean approach with a binned statistic in Appendix A.1.

All of the above mechanisms dilute the demographic imprint of the runaway greenhouse transition with contributions that are challenging to quantify. The "dilution factor" we introduced embodies our ignorance of these contributions. Importantly, while this may offset our estimates of the signal's amplitude, its shape remains the same.

7. CONCLUSIONS

Runaway greenhouse climates on rocky planets cause significant inflation of their radii. Using `Bioverse`, a quantitative hypothesis testing framework, we have explored the potential of contemporary exoplanet missions to statistically detect a radius discontinuity resulting from this inflation in the exoplanet population. Our key findings are as follows:

1. The predicted runaway greenhouse transition causes a discontinuity in the radius and density distribution of small exoplanets with respect to the incident host star irradiation.

2. The demographic imprint of the runaway greenhouse threshold should be detectable with high-precision transit measurements. For a large sample $\gtrsim 100$, a detection is likely if radius inflation occurs on at least $10\,\%$ of the observed planets and if typical water mass fractions are above $\sim 10^{-3}$.

3. We find that the planned *PLATO* transit survey will provide a sufficient sample and the required precision to detect the predicted trend. Assuming the projected photometric precision, *PLATO* will be able to test the runaway greenhouse hypothesis for planet yields $\gtrsim 100$.

4. The diagnostic power of transit missions in testing this hypothesis can be increased through a follow-up campaign providing planet mass measurements. This can reduce the required planet yield by about a factor of three. Only an adequate sample of planets orbiting M dwarfs will ensure sufficient targets on both sides of the expected threshold installation.

5. Testing the runaway greenhouse hypothesis on a population level can provide constraints on the wa-

ter inventory of rocky exoplanets and thus make an important contribution to assessing their habitability. A detection will provide an empirical confirmation of the habitable zone concept and localize its inner edge.

The habitable zone is widely employed in target prioritization for exoplanet missions, and it will provide context for interpreting potential signatures of life. As we have demonstrated, it appears realistic that an empirical confirmation of the habitable zone concept is imminent. This will be a key contribution to the search for extraterrestrial life in the Universe.

Author contributions:

DATA AVAILABILITY

All data sets and software required to reproduce our results and this article are available through GitHub and Zenodo. The code to reproduce the figures can be accessed via the icon links associated with the respective figure caption.

Software: `Bioverse` (Bixel & Apai 2021), Astropy (Astropy Collaboration et al. 2018), NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), corner.py (Foreman-Mackey 2016), dynesty (Speagle 2020).

APPENDIX

A. ROBUSTNESS TESTS

A.1. *Alternative statistics for the average radius or bulk density*

The hypothesis tests introduced in Sect. 4 rely on a statistical estimator for the variation of planetary radii or bulk densities as a function of net instellation, and we chose a moving average for this estimator in our nominal setup. Here, we explore how robust our results are against this choice by demonstrating the recovery of the runaway greenhouse signal in the case of the optimistic survey (Sect. 5.2) with an alternative estimator: instead of computing moving averages, we used a binned statistic.

We first binned the data of the simulated survey in instellation space, choosing the number of bins via the rule of Freedman & Diaconis (1981) and using logarithmic binning. In each bin, we computed the arithmetic mean of the planet radius and its standard deviation. Then, we assigned each planet the mean radius according to the instellation bin it occupies and used this as the measure for testing the runaway greenhouse hypothesis.

Figure 10 shows the simulated data together with binned, average planet radii and draws from the posterior of the hypothesis test. A clear detection resulted, although with somewhat lower significance ($\Delta \ln \mathcal{Z} \approx 30$) compared to the nominal setup. The accuracy of the recovered instellation threshold is comparable. This test demonstrates that our results are not sensitive on the choice of statistical estimator to test the hypotheses against.

It is conceivable that with very large sample sizes and a very sharp runaway greenhouse transition a binned solution would perform better. For a search with real

Figure 10. Detection of the runaway greenhouse threshold with a binned statistic. Using simulated data from the optimistic survey case ($N = 497$, compare Fig. 5), we tested the runaway greenhouse hypothesis. Instead of using a running mean, we computed a binned statistic (arithmetic mean and standard deviation in blue) to assign planets the average radius or bulk density in their neighborhood in instellation space. Random draws from the posterior of the runaway greenhouse hypothesis (Eqn. 5) are shown in red. As in the nominal case, the pattern is detected with high significance.

data, both approaches, and possibly other alternatives, should be considered.

A.2. *Statistical imprint of runaway greenhouse atmospheres*

The predicted runaway greenhouse-induced planet radius changes are a function of instellation, planet mass, and water mass fraction (compare Sect. 3). In order to better understand the interaction between the planetary populations underlying our simulations and these pre-

dictions, we compared the latter to the average radius and bulk density changes we measured in the synthetic population. We used the "optimistic" scenario with a sample size of 497.

Figure 11 shows this comparison for a range of planetary masses and water mass fractions. The complex dependence of the radius inflation on these parameters is evident, but no significant abrupt changes capable of causing spurious signals occur. Differences between model prediction and population can be explained by the wide dispersion in planet mass in the population.

A.3. *Influence of different mass-radius relationships on our results*

We assessed how a different choice of baseline mass-radius relation influences our results. Focussing on the detectability of the statistical runaway greenhouse signal and the ability to constrain the threshold instellation, we repeated the hypothesis test in Sect. 5.2 with alternative mass-radius relationships. Instead of assuming a pure $MgSiO_3$ composition, we assigned planet masses using either the probabilistic relationship in Wolfgang et al. (2016) or a semi-empirical, two-layer relation assuming

an Earth-like (32.5% Fe + 67.5% $MgSiO_3$) composition (Zeng et al. 2016).

Figure 12 shows how the two alternative baseline mass-radius relations influence the significance of a detection and the ability to constrain the threshold instellation. The relation of Wolfgang et al. (2016) includes intrinsic scatter, which impedes a detection at low dilution factors. In this regime, a layered, Earth-like composition leads to more significant detections and a narrower, although biased, constraint on S_{thresh}. Both mass-radius relations agree and recover the injected value where the null hypothesis can be rejected with high significance. This is consistent with our nominal mass-radius relation (compare Sect. 2.4). We conclude that the underlying core and mantle composition of planets may affect the detectability of the transition if the fraction of planets with runaway greenhouse climates is low; however, the overall trends that our experiments revealed appear robust.

We caution that this analysis may serve only as a sanity check and should not be taken as a result in itself: The atmospheric model from Turbet et al. (2020) we adopted relies on a silicate interior composition for its transit radius prediction. Therefore, only our nominal procedure throughout the main body of the paper represents a self-consistent treatment.

REFERENCES

Abe, Y., & Matsui, T. 1988, Journal of the Atmospheric Sciences, 45, 3081, doi: 10.1175/1520-0469(1988)045<3081: EOAIGH>2.0.CO;2

Adams, F. C., Bodenheimer, P., & Laughlin, G. 2005, Astronomische Nachrichten, 326, 913, doi: 10.1002/asna.200510440

Agol, E., Dorn, C., Grimm, S. L., et al. 2021, The Planetary Science Journal, 2, 1, doi: 10.3847/PSJ/abd022

Apai, D., Milster, T. D., Kim, D. W., et al. 2019, The Astronomical Journal, 158, 83, doi: 10.3847/1538-3881/ab2631

Apai, D., Milster, T. D., Kim, D., et al. 2022, in Optical Manufacturing and Testing XIV, Vol. 12221 (SPIE), 59–71, doi: 10.1117/12.2633184

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, The Astronomical Journal, 156, 123, doi: 10.3847/1538-3881/aabc4f

Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, Astronomy and Astrophysics, v.337, p.403-412 (1998), 337, 403

Barclay, T., Pepper, J., & Quintana, E. V. 2018, The Astrophysical Journal Supplement Series, 239, 2, doi: 10.3847/1538-4365/aae3e9

Barnes, J. W. 2007, Publications of the Astronomical Society of the Pacific, 119, 986, doi: 10.1086/522039

Barnes, R., Mullins, K., Goldblatt, C., et al. 2013, Astrobiology, 13, 225, doi: 10.1089/ast.2012.0851

Barth, P., Carone, L., Barnes, R., et al. 2021, Astrobiology, 21, 1325, doi: 10.1089/ast.2020.2277

Benz, W., Broeg, C., Fortier, A., et al. 2021, Experimental Astronomy, 51, 109, doi: 10.1007/s10686-020-09679-4

Berger, T. A., Huber, D., Gaidos, E., van Saders, J. L., & Weiss, L. M. 2020, The Astronomical Journal, 160, 108, doi: 10.3847/1538-3881/aba18a

Bergsten, G. J., Pascucci, I., Mulders, G. D., Fernandes, R. B., & Koskinen, T. T. 2022, The Demographics of Kepler's Earths and Super-Earths into the Habitable Zone

Bitsch, B., Raymond, S. N., & Izidoro, A. 2019, Astronomy & Astrophysics, 624, A109, doi: 10.1051/0004-6361/201935007
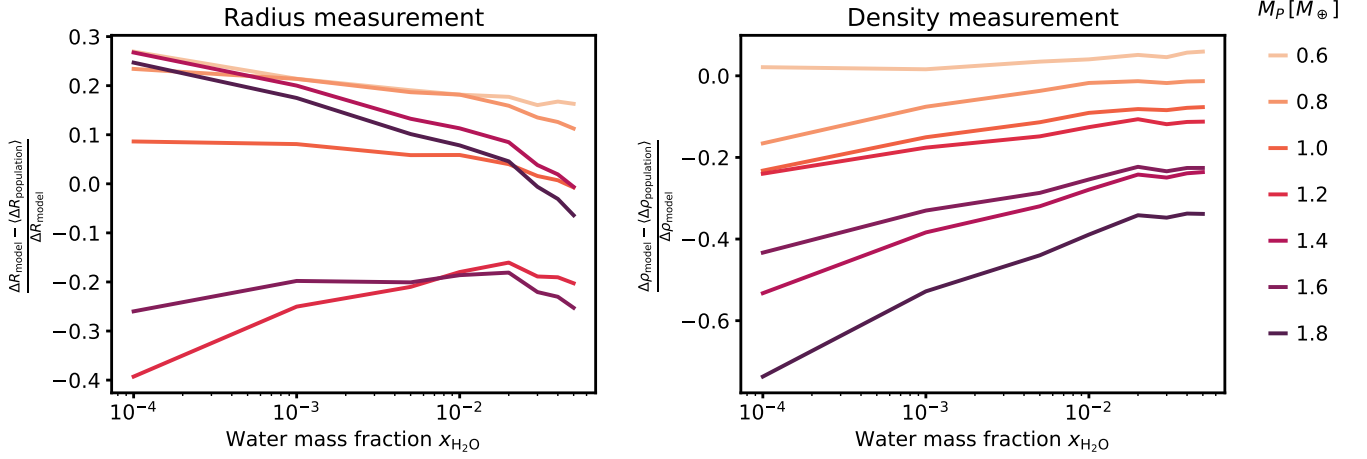
Figure 11. Comparison of radius and bulk density changes predicted by the atmospheric models to the average changes measured in the synthetic planet population. Model predictions $\Delta R_{\mathrm{model}}$ and $\Delta \rho_{\mathrm{model}}$ depend on individual planet masses; $\langle \Delta R_{\mathrm{population}} \rangle$ and $\langle \Delta \rho_{\mathrm{population}} \rangle$ are averaged measurements of the overall population. Significant differences are thus expected.
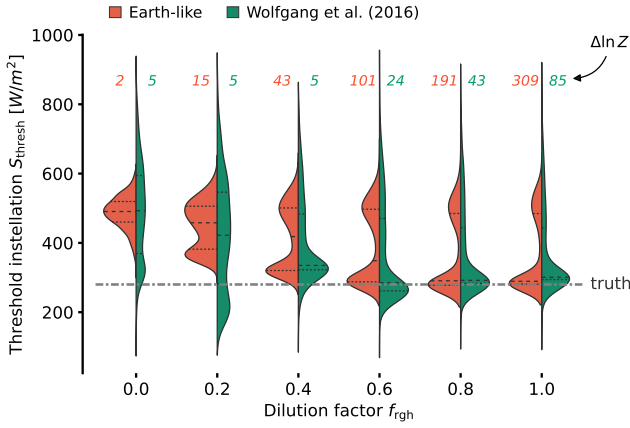
Figure 12. Posterior threshold instellations as a function of dilution factor for two alternative baseline mass-radius relations. For each grid step in f_{rgh}, we show kernel density estimates of retrieved posteriors (averaged over 50 iterations) assuming an Earth-like (32.5 % Fe + 67.5 % MgSiO$_3$) composition or the probabilistic relation from Wolfgang et al. (2016) and otherwise following the optimistic scenario in Sect. 5.2. Lines within the violins show quartiles of the distributions, and the gray line indicates the injected threshold instellation of 280 W m^{-2}. Higher log-evidence differences $\Delta \ln Z$ correspond to more significant rejections of the null hypothesis. In the regime of strong detections, both mass-radius relations lead to similar, accurate constraints on S_{thresh}. Differences occur at low dilution factors, where the Earth-like relation leads to narrower estimates.

Bixel, A., & Apai, D. 2020, The Astrophysical Journal, 896, 131, doi: 10.3847/1538-4357/ab8fad

—. 2021, The Astronomical Journal, 161, 228, doi: 10.3847/1538-3881/abe042

Boukrouche, R., Lichtenberg, T., & Pierrehumbert, R. T. 2021, The Astrophysical Journal, 919, 130, doi: 10.3847/1538-4357/ac1345

Bower, D. J., Kitzmann, D., Wolf, A. S., et al. 2019, Astronomy & Astrophysics, 631, A103, doi: 10.1051/0004-6361/201935710

Brown, A. G. A., Vallenari, A., Prusti, T., et al. 2021, Astronomy & Astrophysics, 649, A1, doi: 10.1051/0004-6361/202039657

Burn, R., Schlecker, M., Mordasini, C., et al. 2021, Astronomy & Astrophysics, 656, A72, doi: 10.1051/0004-6361/202140390

Chabrier, G. 2003, Publications of the Astronomical Society of the Pacific, 115, 763, doi: 10.1086/376392

Crass, J., Gaudi, B. S., Leifer, S., et al. 2021, eprint arXiv:2107.14291, arXiv:2107.14291, doi: 10.48550/arXiv.2107.14291

Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, 2MASS All Sky Catalog of point sources.

Delrez, L., Gillon, M., Queloz, D., et al. 2018, in Ground-Based and Airborne Telescopes VII, Vol. 10700 (SPIE), 446–466, doi: 10.1117/12.2312475

Dietrich, J., Apai, D., Schlecker, M., et al. 2023, EDEN Survey: Small Transiting Planet Detection Limits and Constraints on the Occurrence Rates for Late M Dwarfs within 15 Pc

Dorn, C., & Lichtenberg, T. 2021, The Astrophysical Journal Letters, 922, L4, doi: 10.3847/2041-8213/ac33af

Downey, B. G., Nimmo, F., & Matsuyama, I. 2022, The Thermal-Orbital Evolution of the Earth-Moon System with a Subsurface Magma Ocean and Fossil Figure

Dreizler, S., I., J., et al. 2020, Astronomy & Astrophysics

Dressing, C. D., & Charbonneau, D. 2015, The Astrophysical Journal, 807, 45, doi: 10.1088/0004-637X/807/1/45

Elkins-Tanton, L. T. 2013, Nature, 497, 570, doi: 10.1038/497570a

Emsenhuber, A., Mordasini, C., Burn, R., et al. 2021, Astronomy & Astrophysics, 656, A70, doi: 10.1051/0004-6361/202038863

ESA. 2017, PLATO DEFINITION STUDY REPORT (RED BOOK), Tech. Rep. ESA-SCI(2017)1, European Space Agency

Foreman-Mackey, D. 2016, Journal of Open Source Software, 1, 24, doi: 10.21105/joss.00024

Freedman, D., & Diaconis, P. 1981, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 57, 453, doi: 10.1007/BF01025868

Fulton, B. J., & Petigura, E. A. 2018, The Astronomical Journal, 156, 264, doi: 10.3847/1538-3881/aae828

Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, The Astronomical Journal, 154, 109, doi: 10.3847/1538-3881/aa80eb

Gaidos, E., Claytor, Z., Dungee, R., Ali, A., & Feiden, G. A. 2023, Monthly Notices of the Royal Astronomical Society, stad343, doi: 10.1093/mnras/stad343

Gaudi, S., Seager, S., Mennesson, B., et al. 2020, arXiv e-prints, arXiv:2001.06683

Gillon, M., Jehin, E., Lederer, S. M., et al. 2016, Nature, 533, 221, doi: 10.1038/nature17448

Gillon, M., Triaud, A. H., Demory, B. O., et al. 2017, Nature, 542, 456, doi: 10.1038/nature21360

Ginzburg, S., Schlichting, H. E., & Sari, R. 2016, The Astrophysical Journal, 825, 29, doi: 10.3847/0004-637x/825/1/29

—. 2018, Monthly Notices of the Royal Astronomical Society, 476, 759, doi: 10.1093/mnras/sty290

Goldblatt, C. 2015, Astrobiology, 15, 362, doi: 10.1089/ast.2014.1268

Goldblatt, C., Robinson, T. D., Zahnle, K. J., & Crisp, D. 2013, Nature Geoscience, 6, 661, doi: 10.1038/ngeo1892

Goldblatt, C., & Watson, A. J. 2012, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 370, 4197, doi: 10.1098/rsta.2012.0004

Gupta, A., & Schlichting, H. E. 2019, Monthly Notices of the Royal Astronomical Society, 487, 24, doi: 10.1093/mnras/stz1230

Haar, T. H. V., & Suomi, V. E. 1971, Journal of the Atmospheric Sciences, 28, 305, doi: 10.1175/1520-0469(1971)028<0305: MOTERB>2.0.CO;2

Hamano, K., Abe, Y., & Genda, H. 2013, Nature, 497, 607, doi: 10.1038/nature12163

Hamano, K., Kawahara, H., Abe, Y., Onishi, M., & Hashimoto, G. L. 2015, The Astrophysical Journal, 806, 216, doi: 10.1088/0004-637X/806/2/216

Hardegree-Ullman, K. K., Zink, J. K., Christiansen, J. L., et al. 2020, The Astrophysical Journal Supplement Series, 247, 28, doi: 10.3847/1538-4365/ab7230

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357, doi: 10.1038/s41586-020-2649-2

Hart, M. H. 1978, Icarus, 33, 23, doi: 10.1016/0019-1035(78)90021-0

Huang, S.-S. 1959, American Scientist, 47, 397

Ikoma, M., Elkins-Tanton, L., Hamano, K., & Suckale, J. 2018, Space Science Reviews, 214, 76, doi: 10.1007/s11214-018-0508-3

Irwin, J., Charbonneau, D., Nutzman, P., & Falco, E. 2009, 253, 37, doi: 10.1017/S1743921308026215

Jin, S., & Mordasini, C. 2018, The Astrophysical Journal, 853, 163, doi: 10.3847/1538-4357/aa9f1e

Jin, S., Mordasini, C., Parmentier, V., et al. 2014, Astrophysical Journal, 795, doi: 10.1088/0004-637X/795/1/65

Johnson, J. A., Howard, A. W., Marcy, G. W., et al. 2010, Publications of the Astronomical Society of the Pacific, 122, 149, doi: 10.1086/651007

Kaltenegger, L., & Sasselov, D. 2011, The Astrophysical Journal, 736, L25, doi: 10.1088/2041-8205/736/2/L25

Kasting, J. F. 1988, Icarus, 74, 472, doi: 10.1016/0019-1035(88)90116-9

Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, Icarus, 101, 108, doi: 10.1006/icar.1993.1010

Kipping, D. M. 2013, Monthly Notices of the Royal Astronomical Society: Letters, 434, L51, doi: 10.1093/mnrasl/slt075

Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, The Astrophysical Journal, 765, 131, doi: 10.1088/0004-637X/765/2/131

Leconte, J., Forget, F., Charnay, B., Wordsworth, R., & Pottier, A. 2013, Nature, 504, 268, doi: 10.1038/nature12827

Lichtenberg, T., Golabek, G. J., Burn, R., et al. 2019, Nature Astronomy, 3, 307, doi: 10.1038/s41550-018-0688-5

Lichtenberg, T., Schaefer, L. K., Nakajima, M., & Fischer, R. A. 2022, Geophysical Evolution During Rocky Planet Formation

Luger, R., & Barnes, R. 2015, Astrobiology, 15, 119, doi: 10.1089/ast.2014.1231

Luger, R., Sestovic, M., Kruse, E., et al. 2017, Nature Astronomy, 1, 1, doi: 10.1038/s41550-017-0129

Luque, R., & Pallé, E. 2022, Science, 377, 1211, doi: 10.1126/science.abl7164

Mann, A. W., Feiden, G. A., Gaidos, E., Boyajian, T., & von Braun, K. 2015, ApJ, 804, 64, doi: 10.1088/0004-637X/804/1/64

Mann, A. W., Dupuy, T., Kraus, A. L., et al. 2019, ApJ, 871, 63, doi: 10.3847/1538-4357/aaf3bc

Mazeh, T., Holczer, T., & Faigler, S. 2016, Astronomy and Astrophysics, 589, doi: 10.1051/0004-6361/201528065

Ment, K., & Charbonneau, D. 2023, The Occurrence Rate of Terrestrial Planets Orbiting Nearby Mid-to-late M Dwarfs from TESS Sectors 1-42, doi: 10.48550/arXiv.2302.04242

Milster, T. D., Sik Kim, Y., Wang, Z., & Purvin, K. 2020, Applied Optics, 59, 7900, doi: 10.1364/AO.394124

Montet, B. T., Yee, J. C., & Penny, M. T. 2017, Publications of the Astronomical Society of the Pacific, 129, 044401, doi: 10.1088/1538-3873/aa57fb

Mordasini, C. 2020, Astronomy and Astrophysics, 638, 1, doi: 10.1051/0004-6361/201935541

Mulders, G. D., Ciesla, F. J., Min, M., & Pascucci, I. 2015, The Astrophysical Journal, 807, 9, doi: 10.1088/0004-637X/807/1/9

Nakajima, S., Hayashi, Y.-Y., & Abe, Y. 1992, Journal of the Atmospheric Sciences, 49, 2256, doi: 10.1175/1520-0469(1992)049<2256: ASOTGE>2.0.CO;2

Nascimbeni, V., Piotto, G., Börner, A., et al. 2022, Astronomy & Astrophysics, 658, A31, doi: 10.1051/0004-6361/202142256

Obermeier, C., Koppenhoefer, J., Saglia, R. P., et al. 2016, Astronomy & Astrophysics, 587, A49, doi: 10.1051/0004-6361/201527633

Owen, J. E., & Wu, Y. 2013, Astrophysical Journal, 775, 1, doi: 10.1088/0004-637X/775/2/105

Pecaut, M. J., & Mamajek, E. E. 2013, ApJS, 208, 9, doi: 10.1088/0067-0049/208/1/9

Pepe, F. A., Cristiani, S., Rebolo Lopez, R., et al. 2010, 7735, 77350F, doi: 10.1117/12.857122

Pierrehumbert, R. T. 2022, The Runaway Greenhouse on subNeptune Waterworlds

Popp, M., Schmidt, H., & Marotzke, J. 2016, Nature Communications, 7, 10627, doi: 10.1038/ncomms10627

Puig, L., Pilbratt, G. L., Heske, A., Sanz, I. E., & Crouzet, P.-E. 2016, in Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave, Vol. 9904 (SPIE), 649–657, doi: 10.1117/12.2230964

Quanz, S. P., Ottiger, M., Fontanet, E., et al. 2021, Large Interferometer For Exoplanets (LIFE): I. Improved Exoplanet Detection Yield Estimates for a Large Mid-Infrared Space-Interferometer Mission

Queloz, D., Mayor, M., Udry, S., et al. 2001, The Messenger, 105, 1

Rauer, H. 2021, PLATO Science Objectives and PLATO Mission Consortium, doi: 10.5281/zenodo.5585341

Rauer, H., Aerts, C., Cabrera, J., & PLATO Team. 2016, Astronomische Nachrichten, 337, 961, doi: 10.1002/asna.201612408

Raymond, S. N., Quinn, T., & Lunine, J. I. 2004, Icarus, 168, 1, doi: 10.1016/j.icarus.2003.11.019

Reid, I. N., Turner, E. L., Turnbull, M. C., Mountain, M., & Valenti, J. A. 2007, The Astrophysical Journal, 665, 767, doi: 10.1086/519001

Ribas, I., Reiners, A., Zechmeister, M., et al. 2023, The CARMENES Search for Exoplanets around M Dwarfs. Guaranteed Time Observations Data Release 1 (2016-2020)

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003, doi: 10.1117/1.JATIS.1.1.014003

Rogers, L. A., & Seager, S. 2010, The Astrophysical Journal, 712, 974, doi: 10.1088/0004-637X/712/2/974

Sato, T., Okuzumi, S., & Ida, S. 2016, Astronomy & Astrophysics, 589, A15, doi: 10.1051/0004-6361/201527069

Schlecker, M., Mordasini, C., Emsenhuber, A., et al. 2021a, Astronomy & Astrophysics, 656, A71, doi: 10.1051/0004-6361/202038554

Schlecker, M., Pham, D., Burn, R., et al. 2021b, Astronomy & Astrophysics, 656, A73, doi: 10.1051/0004-6361/202140551

Sebastian, D., Gillon, M., Ducrot, E., et al. 2021, Astronomy & Astrophysics, 645, A100, doi: 10.1051/0004-6361/202038827

Skilling, J. 2004, AIP Conference Proceedings, 735, 395, doi: 10.1063/1.1835238

Smart, R. L., Sarro, L. M., Rybizki, J., et al. 2021, Astronomy & Astrophysics, 649, A6, doi: 10.1051/0004-6361/202039498

Speagle, J. S. 2020, Monthly Notices of the Royal Astronomical Society, doi: 10.1093/mnras/staa278

Spergel, D., Gehrels, N., Baltay, C., et al. 2015, Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report, doi: 10.48550/arXiv.1503.03757

Szabó, G. M., & Kiss, L. L. 2011, Astrophysical Journal Letters, 727, 2, doi: 10.1088/2041-8205/727/2/L44

Tamburo, P., Muirhead, P. S., & Dressing, C. D. 2023, Predicting the Yield of Small Transiting Exoplanets around Mid-M and Ultra-Cool Dwarfs in the Nancy Grace Roman Space Telescope Galactic Bulge Time Domain Survey, doi: 10.48550/arXiv.2303.09959

Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, The Astrophysical Journal Supplement Series, 235, 38, doi: 10.3847/1538-4365/aab4f9

Torres, G., Andersen, J., & Giménez, A. 2010, A&A Rv, 18, 67, doi: 10.1007/s00159-009-0025-1

Turbet, M., Bolmont, E., Ehrenreich, D., et al. 2020, Astronomy & Astrophysics, Volume 638, id.A41, <NUMPAGES>10</NUMPAGES> pp., 638, A41, doi: 10.1051/0004-6361/201937151

Turbet, M., Ehrenreich, D., Lovis, C., Bolmont, E., & Fauchez, T. 2019, Astronomy & Astrophysics, 628, A12, doi: 10.1051/0004-6361/201935585

Unterborn, C. T., Hinkel, N. R., & Desch, S. J. 2018, Research Notes of the AAS, 2, 116, doi: 10.3847/2515-5172/aacf43

Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2018, Monthly Notices of the Royal Astronomical Society, 479, 4786, doi: 10.1093/mnras/sty1783

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: 10.1038/s41592-019-0686-2

Watson, A. J., Donahue, T. M., & Walker, J. C. G. 1981, Icarus, 48, 150, doi: 10.1016/0019-1035(81)90101-9

Wolfgang, A., Rogers, L. A., & Ford, E. B. 2016, The Astrophysical Journal, 825, 19, doi: 10.3847/0004-637X/825/1/19

Youdin, A. N. 2011, Astrophysical Journal, 742, doi: 10.1088/0004-637X/742/1/38

Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, The Astrophysical Journal, 819, 127, doi: 10.3847/0004-637X/819/2/127