arXiv:2310.11613v1 [astro-ph.EP] 17 Oct 2023

No Evidence for More Earth-sized Planets in the Habitable Zone of *Kepler*'s M versus FGK Stars

Galen J. Bergsten (ID),[1] Ilaria Pascucci (ID),[1] Kevin K. Hardegree-Ullman (ID),[2] Rachel B. Fernandes (ID),[3, 4, 1, *]
Jessie L. Christiansen (ID),[5] and Gijs D. Mulders (ID)[6, 7]

[1]*Lunar and Planetary Laboratory, The University of Arizona, Tucson, AZ 85721, USA*
[2]*Steward Observatory, The University of Arizona, Tucson, AZ 85721, USA*
[3]*Department of Astronomy & Astrophysics, Center for Exoplanets and Habitable Worlds, The Pennsylvania State University, University Park, PA 16802, USA*
[4]*Center for Exoplanets and Habitable Worlds, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA, 16802, USA*
[5]*NASA Exoplanet Science Institute, IPAC, MS 100-22, Caltech, 1200 E. California Blvd, Pasadena, CA 91125*
[6]*Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Av. Diagonal las Torres 2640, Peñalolén, Santiago, Chile*
[7]*Millennium Institute for Astrophysics, Chile*

ABSTRACT

Reliable detections of Earth-sized planets in the habitable zone remain elusive in the *Kepler* sample, even for M dwarfs. The *Kepler* sample was once thought to contain a considerable number of M dwarf stars ($T_{\rm eff} < 4000\,{\rm K}$), which hosted enough Earth-sized ($[0.5, 1.5]\,{\rm R}_\oplus$) planets to estimate their occurrence rate ($\eta_\oplus$) in the habitable zone. However, updated stellar properties from *Gaia* have shifted many *Kepler* stars to earlier spectral type classifications, with most stars (and their planets) now measured to be larger and hotter than previously believed. Today, only one partially-reliable Earth-sized candidate remains in the optimistic habitable zone, and zero in the conservative zone. Here we performed a new investigation of *Kepler*'s Earth-sized planets orbiting M dwarf stars, using occurrence rate models with considerations of updated parameters and candidate reliability. Extrapolating our models to low instellations, we found an occurrence rate of $\eta_\oplus = 8.58^{+17.94}_{-8.22}\%$ for the conservative habitable zone (and $14.22^{+24.96}_{-12.71}\%$ for the optimistic), consistent with previous works when considering the large uncertainties. Comparing these estimates to those from similarly comprehensive studies of Sun-like stars, we found that the current *Kepler* sample does not offer evidence to support an increase in $\eta_\oplus$ from FGK to M stars. While the *Kepler* sample is too sparse to resolve an occurrence trend between early and mid-to-late M dwarfs for Earth-sized planets, studies including larger planets and/or data from the *K2* and *TESS* missions are well-suited to this task.

Keywords: Exoplanets (498) — Habitable planets (695)

1. INTRODUCTION

In the ongoing search for life in the Universe, one standing question is the occurrence rate (or frequency) of Earth-sized planets in the habitable zone. Studies to this point are often focused on Sun-like (FGK) stars, in part because they are most reflective of our own solar system and thus life as we know it. Sharing this fo-

cus, the *Kepler* mission (Borucki et al. 2010; Borucki 2016) sought to detect Earth-sized habitable zone planets, with long-term monitoring preferentially targeting Sun-like stars. Mechanical failures rendered the *Kepler* telescope unable to continue viewing its target field before it could provide sufficiently numerous confident detections of Earth-sized planets at $\sim 1\,{\rm year}$ orbital periods to enable robust occurrence calculations. As such, many estimates for the occurrence rate of Earth-sized habitable zone planets ($\eta_\oplus$) around FGK stars are based on extrapolations from the close-in population where *Kepler*'s survey completeness for small planets is com-

Corresponding author: Galen J. Bergsten
gbergsten@arizona.edu

* President's Postdoctoral Fellow

parably higher than in the habitable zone. Estimates for $\eta_\oplus$ around Sun-like stars span from $\sim 1 - 100\%$ depending on various physical and statistical considerations, with recent estimates tending towards $10 - 40\%$ (see e.g., discussions in Kunimoto & Matthews 2020; Bergsten et al. 2022).

The habitable zone for M dwarfs occurs at shorter orbital periods ($\sim 30 - 140$ days), meaning Earth-sized habitable zone planets around these stars would have ideally been detected within *Kepler*'s lifespan. Yet the survey's selection of fewer M dwarfs for monitoring, coupled with its generally poor sensitivity to fainter stars, meant that the *Kepler* population contained only a few detections of these planet candidates around M dwarfs. Nonetheless, Dressing & Charbonneau (2015) leveraged the available *Kepler* information to identify a sample of 156 planets around M dwarf ($T_{\rm eff} < 4000\,{\rm K}$, $\log g > 3$) stars, and estimated (via the inverse detection efficiency method) the occurrence rate of Earth-sized habitable zone planets. For $[1.0, 1.5]\,R_\oplus$ planets, they estimated an occurrence of $\eta_\oplus = 15.82^{+16.60}_{-6.54}\%$ for a conservative habitable zone (between the moist and maximum greenhouse boundaries, Kopparapu et al. 2013) and $24.28^{+17.58}_{-8.39}\%$ for an optimistic habitable zone (between the recent Venus and early Mars boundaries). Separately, Mulders et al. (2015a) used the *Kepler* Q1-Q16 planet candidate sample (Mullally et al. 2015) to find that M dwarfs have ~ 3.5 times more small planets than FGK stars. While the Mulders et al. (2015a) estimate was derived from the close-in ($P < 50$ days) population, it has since been applied to scale $\eta_\oplus$ estimates from FGK stars as an approximation for Earth-sized habitable zone planets around M dwarfs (see e.g., Hardegree-Ullman et al. 2023).

Several recent advancements have since changed the field of *Kepler* demographics, including the release of the final *Kepler* DR25 catalog (Thompson et al. 2018) and stellar parameter revisions informed by the *Gaia* mission (Gaia Collaboration et al. 2018, e.g., Berger et al. 2020b). The work of Hsu et al. (2020) incorporated these advancements by using DR25 in conjunction with information from *Gaia* DR2 and the 2MASS point source catalogue (Skrutskie et al. 2006). They estimated an occurrence rate of $33^{+10}_{-12}\%$ for $[0.75, 1.5]\,R_\oplus$ planets in the conservative habitable zone around M dwarfs, and found that $\eta_\oplus$ estimates are sensitive to the choice of modeling priors in this regime of sparse detections. Hsu et al. (2020) also found that the small planet occurrence rates around M dwarfs are comparable to those of FGK stars (Hsu et al. 2019) when evaluated at similar instellations, thus raising concerns that the factor of 3.5 offset

between M and FGK stars might not be applicable to the habitable zone.

A third advancement which merits consideration are new statistical treatments of candidate reliability (Bryson et al. 2020a,b), which serves to better account for false positives and false alarms. To date, no study has incorporated *Kepler* candidate reliability alongside DR25 and *Gaia*-revised parameters to revisit the question of Earth-sized habitable zone planets around *Kepler*'s M dwarfs. Here we seek to do just that, providing an updated look into the occurrence rate of such planets around the most abundant type of star in the Galaxy.

In Section (2), we outline the motivation for a new study by investigating how previous works and planet samples have been impacted by recent advancements in the field. We describe the current *Kepler* sample of Earth-sized planets around M dwarfs in Section (3), and introduce our approaches to calculating planet occurrence and fitting population models in Section (4). We discuss our results in Section (5) and describe the habitable zone occurrence rates in Section (5.1). In Section (6), we provide an updated comparison of Earth-sized planet occurrence rates within and across spectral types, before summarizing our work in Section (7).

2. NEED FOR AN UPDATED INVESTIGATION

Before performing an updated analysis with the current *Kepler* sample, we examined how the properties of planets originally studied in Dressing & Charbonneau (2015) (hereafter DC15) had been impacted by subsequent studies. Of the 156 planets presented in DC15 (their Table 9), there were 145 planets[1] around 85 stars with planet parameters within $[0.5, 4]\,R_\oplus$ and $[0.2, 400]\,I_\oplus$, where $I_\oplus$ denotes the instellation flux of (present-day) Earth. Cross-matching with the *Kepler* DR25 catalog (Thompson et al. 2018), we found that five planet candidates are now considered false positives according to automated classification via Robovetter (Coughlin 2017). We include a brief discussion of objects with contrasting dispositions between DR25 and subsequent literature in Appendix (A).

For the remaining 140 planets around 83 stars, we cross-matched host stars with the *Gaia-Kepler* Stellar Properties Catalog (Berger et al. 2020b). We found that

---

[1] The candidate planet K04427.01 appeared with $1.56^{+0.25}_{-0.23}\,R_\oplus$ and $0.17^{+0.06}_{-0.05}\,I_\oplus$ in DC15, which fell outside of the original $0.2\,I_\oplus$ lower bound. After updating properties with *Gaia*, the host star has $T_{\rm eff} = 3895\,K$, and the candidate has a habitable zone instellation of $0.32 \pm 0.03\,I_\oplus$. However, the planet radius of $1.79^{+0.12}_{-0.09}\,R_\oplus$ exceeds the "Earth-sized" definition at 1σ, so it is excluded from this discussion (but is considered for our analysis of the current *Kepler* sample; Section 3).

seven stars (hosting fourteen planets) were not included in the updated catalog (typically because of low-quality photometry or lacking a *Gaia* DR2 parallax), and we thus excluded these objects from the following analysis. This included one of six planets which, within the original 1σ uncertainties, could have fallen within the DC15 bounds for an Earth-sized habitable zone planet ($[1, 1.5]\,R_\oplus$, $[0.23, 1.54]\,I_\oplus$) for an optimistic habitable zone. This planet was detected in DC15 with a radius of $1.03\,R_\oplus$ and subsequently appeared as K01681.04 in *Kepler* DR24 ($0.77\,R_\oplus$, Coughlin et al. 2016). However, K01681.04 later appeared in *Kepler* DR25 with a radius of $10.39\,R_\oplus$, so it is unlikely that this planet would now contribute to $\eta_\oplus$ (barring an uncharacteristic order of magnitude decrease in stellar radius with *Gaia*).

From the remaining 126 planets orbiting 76 stars, we used the *Gaia*-updated stellar mass (M_*) estimates from Berger et al. (2020b) along with DR25 orbital periods (P) to calculate planet semi-major axes (a) with Kepler's law for a negligible planet mass:

$$a = \left[\frac{GM_* P^2}{4\pi^2} \right]^{1/3}, \qquad (1)$$

and used these in conjunction with Berger et al. (2020b) luminosities (L_*) to estimate updated planetary instellations (I):

$$I = \frac{L_*}{4\pi a^2}. \qquad (2)$$

We also updated planet radii by multiplying the DR25 planet-to-star radius ratios with the Berger et al. (2020b) stellar radii. For all above calculations, we propagated uncertainties with Monte Carlo sampling of split-normal distributions using the 1σ uncertainties provided in Berger et al. (2020b) for stellar parameters and in DR25 for planetary parameters.[2]

Because the original DC15 sample was defined for cool stars with $T_{\rm eff} < 4000\,$K, it is important to note which stars did or did not meet this classification after receiving updated temperature estimates in Berger et al. (2020b). Thus, when plotting the DC15 planet sample with updated parameters in Figure (1), we split the sample into those with host stars still considered cool, and those with host stars now considered too warm ($T_{\rm eff} > 4000\,$K) to be relevant. We found that two of the DC15 Earth-like planets are excluded from the cool star sample with this change (K00571.05 with $T_{\rm eff} = 4023\,$K and K02650.01 with $T_{\rm eff} = 4096\,$K). Of the still-cool

stars, the aforementioned revisions of planetary parameters had caused the remaining three DC15 Earth-like planets (K00463.01, K02418.01, K03284.01) to no longer fall within the $[1, 1.5]\,R_\oplus$, $[0.23, 1.54]\,I_\oplus$ regime.

The above changes left the original DC15 sample with no Earth-sized planets in the habitable zones of cool stars (lower left panel of Figure 1). We note that the DC15 study took place before the release of *Kepler*'s final catalog in DR25, such that it could not have made use of the full *Kepler* sample. More recently, Hsu et al. (2020) revisited this subject employing both DR25 and their own revised stellar properties informed by *Gaia* and 2MASS (i.e., separate from the catalog of Berger et al. 2020b). They found a similar paucity of relevant planets at habitable zone separations (see e.g., their Figure 3). More generally, they note that the uncertainties on *Kepler*'s M dwarf occurrence rate estimates are larger than previously believed: for example, for planets with $P < 50\,$days and $[1, 2.5]\,R_\oplus$, DC15 found an occurrence rate of $\eta = 1.38^{+0.11}_{-0.09}$ while Hsu et al. (2020) found $\eta = 1.13^{+0.20}_{-0.19}$.

While Hsu et al. (2020) provided an excellent evaluation of *Kepler*'s M dwarf occurrence rates, their work preceded an analysis on the importance of per-candidate reliability in *Kepler* demographic studies (Bryson et al. 2020b). Reliability generally accounts for a candidate's odds of being a false alarm or false positive, and reliability scores (as defined in Bryson et al. 2020a) are bounded between 0 and 1. In an occurrence calculation including reliability, a candidate's contribution will either remain at full weight (if perfectly reliable) or be downweighted (less reliable candidates contribute less), meaning that occurrence rates with reliability are typically lower than those without (see e.g., Bryson et al. 2021; Bergsten et al. 2022). Even in the scenario where a regime of interest is unpopulated (i.e., zero relevant candidates), accounting for reliability can help to ensure more robust extrapolations from populated regimes. As such, to make full use of available *Kepler* statistics, it is necessary to revisit the sample of M dwarfs incorporating DR25, *Gaia*-revised parameters, and candidate reliability in pursuit of updated occurrence rates.

3. CURRENT SAMPLE

Having found the DC15 planet sample sufficiently changed from that of their original study, we elected to perform a new analysis using the current *Kepler* sample. Beginning from the Berger et al. (2020b) catalog, we employed the same effective temperature cut ($T_{\rm eff} < 4000\,$K) used in DC15. We also adopted the empirical selection criterion of Huber et al. (2016) where a

---

[2] Our planet parameters are functionally identical to those of Berger et al. (2020a), who use a similar methodology and set of input parameters, with only small ($< 1\%$) variations attributable to our Monte Carlo sampling.
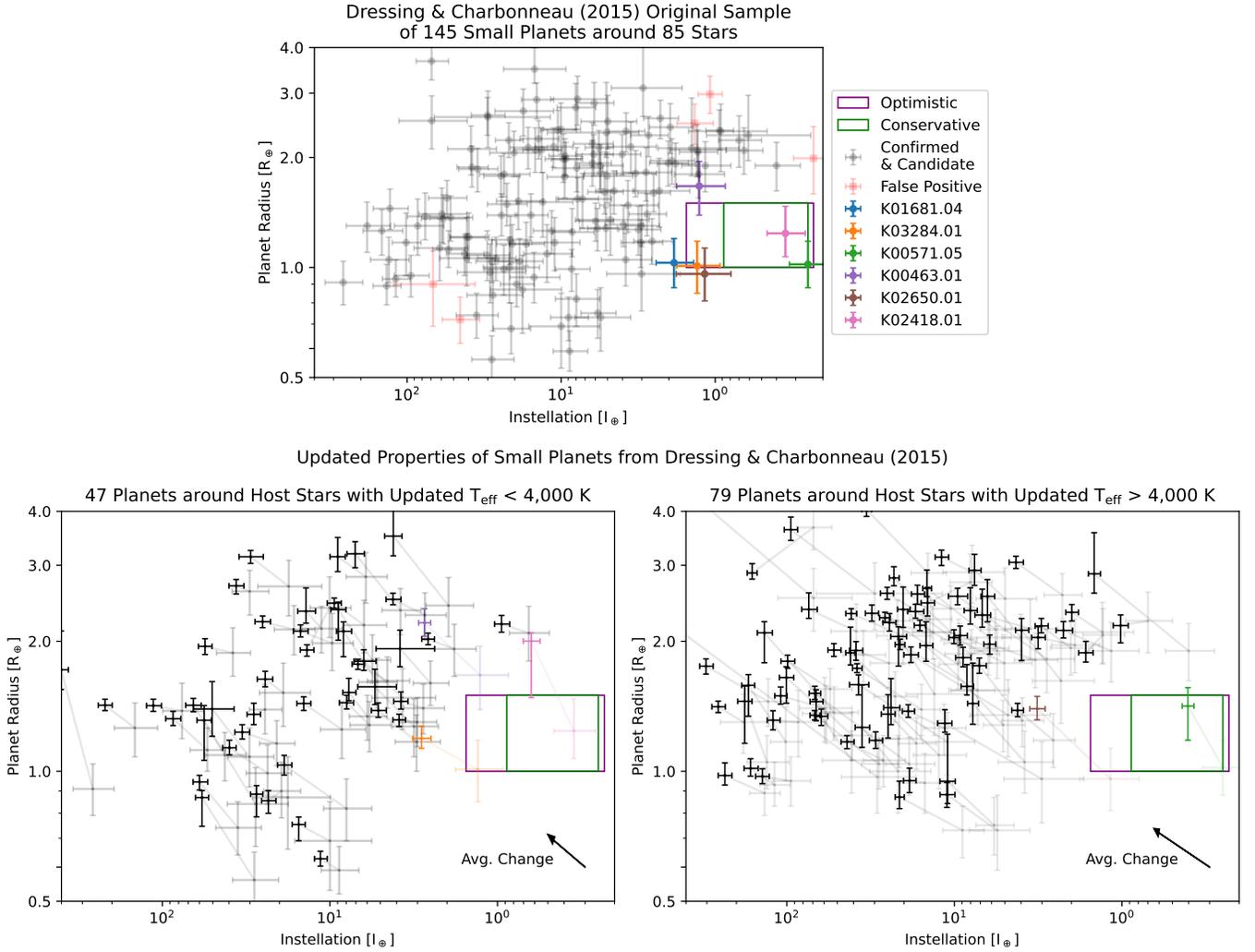
Figure 1. Top: Original properties of small planets included in the Dressing & Charbonneau (2015) study of *Kepler*'s M dwarfs ($T_{\rm eff} < 4000\,{\rm K}$). Rectangles represent the optimistic (purple) and conservative (green) habitable zone boundaries used in Dressing & Charbonneau (2015). Translucent red points denote planets later labeled as false positives in *Kepler* DR25 (Thompson et al. 2018), while opaque colored points represent "Earth-sized" candidates considered relevant to the habitable zone. **Bottom:** current planet properties with updates from *Gaia* DR2 (Gaia Collaboration et al. 2018; Berger et al. 2020b), split between **left:** stars that are still below 4000 K and **right:** stars that are now warmer than 4000 K. Translucent points represent the original properties from Dressing & Charbonneau (2015), while opaque points denote the updated properties of those planets, with lines lines connecting the two for each individual planet. Black arrows denote the average ($\log_{10}$) change in planet properties, measured independently along either axis. The host star of one of the six original habitable zone planets (K01681.04, blue) was not included in the *Gaia-Kepler* Stellar Properties Catalog (Berger et al. 2020b), and thus does not appear in either of the updated parameter panels.

star is considered a dwarf if:

$$\log g > \frac{1}{4.671} \arctan\left(\frac{T_{\rm eff} - 6300}{-67.172}\right) + 3.876. \qquad (3)$$

This left a sample of 2807 *Kepler* M dwarf stars; the distribution of stellar effective temperatures is shown in Figure (2). We then matched these stars with the *Kepler* DR25 planet catalog to identify 60 confirmed and 26 candidate planets orbiting 62 stars in this M dwarf sample, and calculated revised radii and instellations for these

planets with the methodology described in the previous section. We also used the candidate reliability scores calculated in Bergsten et al. (2022), which followed the approach of Bryson et al. (2020b) to calculate reliability in terms of a candidate's false alarm reliability ($R_{\rm FA}$, Bryson et al. 2020b) and false positive probability (Morton et al. 2016). All candidates in our sample had high reliability against being a false alarm ($R_{\rm FA} \gtrsim 95\%$), such that any low reliability scores were driven by a fairly high probability that a candidate was a false positive.
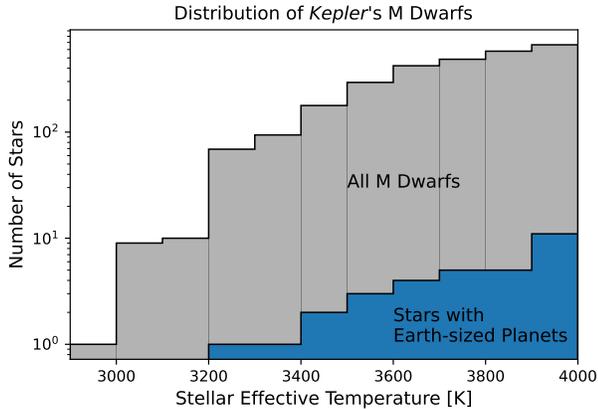
Figure 2. Distribution of stellar effective temperatures (from Berger et al. 2020b) for our sample of *Kepler* M dwarf stars ($T_{\text{eff}} < 4000$ K). The distribution of stars hosting Earth-sized planets is highlighted in blue, and is not significantly distinct from the overall distribution (a two-sample Kolmogorov-Smirnov provides a KS statistic of 0.16 and a p-value of 0.38).

Previous works have indicated that the distribution of occurrence rates (and trends therein) evaluated in orbital period may differ from those evaluated in instellation for the same planet sample (see e.g., Hsu et al. 2020; Petigura et al. 2022). This may be especially prevalent when studying a small number of planets, where the dependence on individual stellar properties in converting between dimensions may introduce enough variation to alter the shape of the resulting occurrence distributions. To enable an assessment of how the chosen dimension might impact occurrence rate calculations, we defined two separate samples: one in orbital period, and one in instellation. For both samples, we adopted a radius range of $[0.5, 1.5] R_\oplus$ to focus on Earth-sized planets. Adopting the same separation bounds as DC15 ($[0.5, 200]$ days; $[0.2, 400] I_\oplus$) yielded a sample of 40 planets around 32 stars in orbital period and 39 planets around 31 stars in instellation *when considering only median parameter values*. The difference in sample size is caused by one planet (K02542.01) with an instellation of $I = 434^{+64}_{-71} I_\oplus$, though the errors are such that this planet may be included when accounting for input uncertainties.

In general, we note that sampling within parameter uncertainties can affect sample size and thus inferred occurrence rates. Relevant for this study, there are several planets whose median radii exceed $1.5 R_\oplus$ but may fall into the Earth-sized classification within their uncertainties, including some with $P > 50$ days where we otherwise lack planets. As such, the samples shown in Figure (3) include a larger number and range of planets

than what is described above, as some may be included when we consider input uncertainties in Section (4.2).

While the *Kepler* completeness contours used to calculate planet occurrence are defined in the planet radius/orbital period plane, we needed to convert these contours to the planet radius/instellation plane to calculate planet occurrence in instellation-space. Following the methodology of Bryson et al. (2021), we took the per-star completeness maps[3] and applied a one-dimensional conversion, translating each orbital period point to a corresponding instellation flux value via Equations (1) and (2) using that star's median mass and luminosity from Berger et al. (2020b). We then interpolated each star's two-dimensional completeness map onto a uniform grid, and added the results for all stars to create the summed completeness grid necessary for Section (4.2). Because the orbital period maps are all defined on the same grid, we may simply add them together as-is without re-interpolating onto a common grid. The average completeness maps (obtained by dividing the summed maps by the total number of stars) are shown in Figure (3).

4. OCCURRENCE RATE METHODOLOGY

We adopted two approaches to calculating planet occurrence rates: direct evaluation via the inverse detection efficiency method, and indirect evaluation by way of a population forward model. The former is straightforward but can only constrain occurrence rates where there are exoplanet detections. Extrapolations from a forward model are thus needed to evaluate occurrence rates in poorly- or un-populated regimes, such as the longer-period (or lower-instellation) habitable zones that we were interested in studying. The inverse detection efficiency method is still a useful way of comparing model predictions in regimes where there is a significant number of exoplanet detections, and so we performed both evaluations where we could.

All subsequent calculations were performed first in orbital period P with the corresponding period sample, and then again in instellation I (where I was substituted for P as needed) with the instellation sample.

4.1. *Inverse Detection Efficiency Method*

We employed the classical inverse detection efficiency method, modified to incorporate the reliability of planet candidates as in Bergsten et al. (2022). For a survey

---

[3] We created per-star completeness maps using a modified version of `epos` (Mulders et al. 2018, updated in Bergsten et al. 2022), which includes the per-star detection efficiency contours generated via `KeplerPORTs` (Burke & Catanzarite 2017).
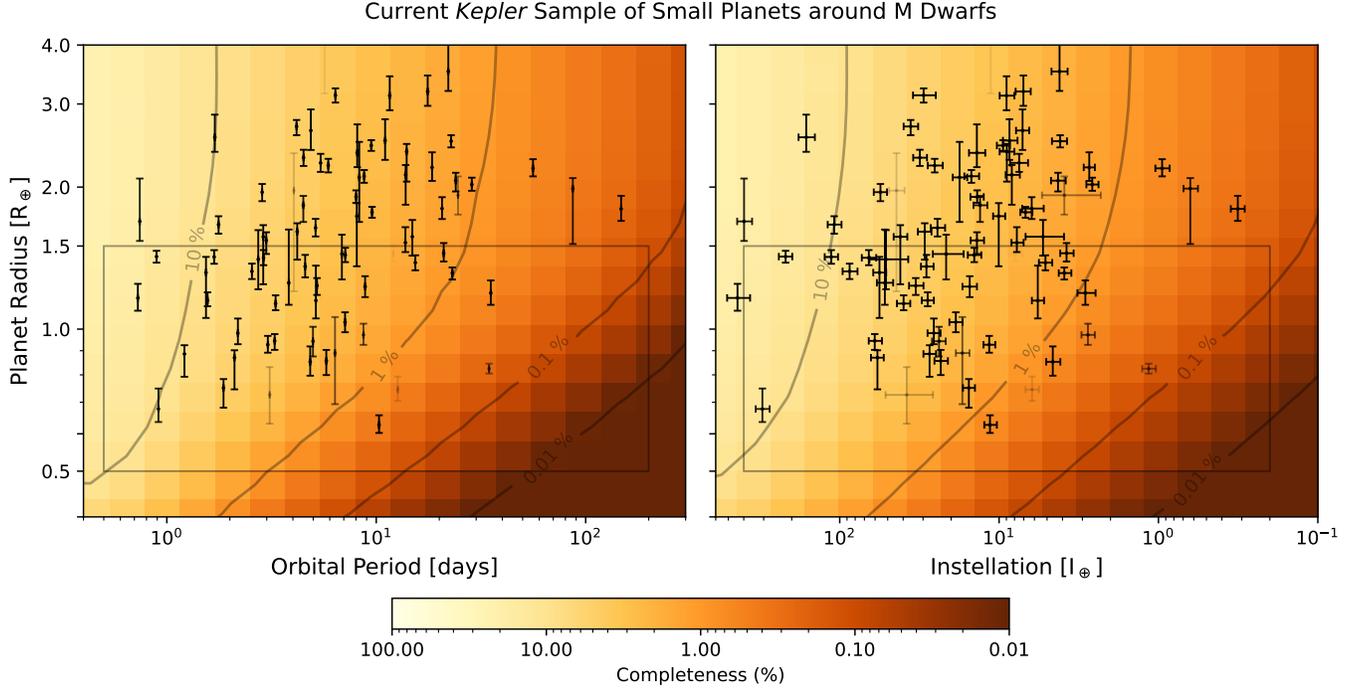
Figure 3. The current sample of *Kepler* small confirmed and candidate planets with *Gaia*-updated properties, plotted in **left:** orbital period and **right:** instellation. Error bars represent 1σ uncertainties, and opacity is determined by reliability (fully opaque corresponds to a reliability score of 1, Bryson et al. 2020a). The average *Kepler* completeness maps for these M dwarfs are plotted in the background, with contours denoting relevant orders of magnitude (0.01% to 10%). Grey rectangles denote the parameter range(s) of interest for this study.

with N_* stars, the occurrence η in a given bin of planetary parameters with $N_{\rm pl}$ planet candidates is given by:

$$\eta_{\rm bin} = \frac{1}{N_*} \sum_j^{N_{\rm pl}} \frac{{\rm rel}_j}{{\rm comp}_j}, \qquad (4)$$

with uncertainties given as $\sigma = \eta_{\rm bin}/\sqrt{N_{\rm pl}}$. Here, the $j^{\rm th}$ planet in a bin provides a weight proportional to its reliability ${\rm rel}_j$ and inversely proportional to the host star-specific completeness ${\rm comp}_j$ evaluated at that planet's parameters.

4.2. *Population Modeling*

Following the approach of Youdin (2011) and Burke et al. (2015), we adopted a population distribution function of the form:

$$\frac{{\rm d}^2 f}{{\rm d}P {\rm d}R} = F_0 C_{\rm n} g(P, R). \qquad (5)$$

The normalization factor $C_{\rm n}$ is defined such that the integral of Equation (5) over the entire sample domain (i.e., $[0.5, 1.5]\,R_\oplus$ and $[0.5, 200]\,{\rm days}$ or $[0.2, 400]\,I_\oplus$) equals the scaling parameter F_0, which represents the average number of planets per star. The shape function $g(P, R)$ describes how the planet population behaves in

orbital period and radius. Previous demographic studies have often made use of (broken) power laws in these dimensions to match observed occurrence rate distributions (see e.g., Youdin 2011; Howard et al. 2012; Dong & Zhu 2013; Burke et al. 2015). We opted for a conventional broken power law in orbital period and a single power law in radius, defined by:

$$g(P, R) = R^\alpha \begin{cases} (P/P_{\rm break})^{\beta_1} & {\rm if}\ P < P_{\rm break} \\ (P/P_{\rm break})^{\beta_2} & {\rm if}\ P \geq P_{\rm break}. \end{cases} \qquad (6)$$

This contains four free parameters in addition to F_0: the exponents β_1 and β_2 control the slope of the occurrence distribution in orbital period space on either side of the break $P_{\rm break}$, and α determines the slope of the occurrence distribution in planet radius space.

The work of Bergsten et al. (2022) studied *Kepler*'s small planets around Sun-like (FGK) stars and found evidence of a period-dependent radius distribution that is not well-described by the simple power laws above. In their model, the small planet population shifts from being dominated by super-Earths at shorter periods to sub-Neptunes at longer periods, switching around some transition orbital period, and potentially signifying the effects of atmospheric mass loss processes. A large number of small planet detections was required to constrain

the fairly complex functional form that described this behavior.

Bergsten et al. (2022) found the transition period to scale with stellar mass, and extrapolations of that trend would place the transition at very short orbital periods ($\sim 5 \pm 1$ days) for the median stellar mass of our sample ($0.50\,\mathrm{M}_\odot$). This coincides with typical values of P_{break} in Equation (6) for M dwarfs, such that the steep occurrence slope at shorter periods may explain the lack of detected planets needed to resolve the closest-in side of the transition. However, assuming a uniformly varying behavior across FGKM stars such that this extrapolation holds, this also means that the relative fractions of super-Earths and sub-Neptunes should be roughly flat with orbital period for the majority of our domain of interest. As such, the simple power laws of Equations (5) and (6) may be suitable for modeling the current M dwarf sample especially given the presently limited number of detected planets; exploration of more complicated forms may be enabled by larger datasets in the future.

4.2.1. *Model Fitting*

The population model was fit to the planet sample by minimizing the corresponding Poisson likelihood function (Eqn. 9 in Burke et al. 2015) through a Markov-chain Monte Carlo process using the `emcee` (Foreman-Mackey et al. 2013) package. Note that this likelihood function requires an evaluation of the sum of each star's completeness over an integral (or grid) of planet radii and orbital periods. We achieved this by evaluating each individual star's completeness on a uniform grid, and then taking the sum of all stars' completeness at each combination of planet parameters. We adopted uniform priors, and specifically fit the log of the power law break ($\log_{10} P_{\mathrm{break}}$) to ensure proper sampling over several orders of magnitude when repeating this process in instellation.

We adopted three different methodologies to determine the input population and their parameters. In our first method (M1), we used the median values of each planetary parameter, thus considering only the confirmed and candidate planets with median values within the specified bounds. We used 64 walkers run for 20,000 steps and discard the first 1,000 for burn-in; justifications for optimization-related choices are included in Appendix (B). While this method is conventional, it cannot account for the reliability of planet candidates nor the uncertainty in their input parameters (which themselves depend on the uncertainty of the stellar parameters used to calculate updated radii and instellations).

In our second method (M2), we followed the approach of Bryson et al. (2020a) to implement reliability by per-

forming separate inferences where the input planet population is drawn according to their reliability (e.g., a 50% reliable candidate is included in 50% of fits). We performed 100 inferences, each using 64 walkers run for 20,000 steps (discarding the first 1,000 for burn-in), and concatenated the posteriors to represent the global parameter distributions.

In our third and most complete method (M3), we implemented both reliability and parameter uncertainties. For a given fit, we first took the entire planet sample around M dwarfs (not yet restricting to a specific parameter regime) and drew each candidate's parameters from within their (presumed independent) split-normal orbital period and radius distributions. If a candidate's drawn parameters were within the relevant regime, *and* the candidate passed a separate draw according to its reliability (as in M2), then that candidate was included in the fit using the sampled parameters. We performed 400 of these inferences, each using 64 walkers run for 20,000 steps (discarding the first 1,000 for burn-in), and concatenated the posteriors.

While we sampled within the uncertainties of our planet parameters, a more inclusive approach may also sample within stellar effective temperature uncertainties to include stars (and their planets) near the $4000\,\mathrm{K}$ boundary. However, as noted in Bryson et al. (2021), because the number of stars would vary between inferences, the summed completeness components would also need re-evaluation for each inference, which is computationally expensive and beyond the scope of this work.

5. RESULTS

The occurrence rate grids from the reliability-weighted inverse detection efficiency method are presented in Figure (4) for both orbital period and instellation. These occurrence rates were then marginalized along either axis to provide the one-dimensional occurrence rates seen in Figure (6) for orbital period and Figure (7) for instellation. To determine the distribution of occurrence rates predicted by our population model(s), we evaluated Equation (5) with the `emcee`-output parameter vectors to produce occurrence grids like those in Figure (5), and further integrated along either axis to present the modeled marginalized distributions in Figures (6) and (7). The best-fit parameter values are presented in Table (1), where we treated the 16th and 84th percentiles as 1σ bounds about the median (50th percentile) value.

We found values from the inverse detection efficiency method to fall within model predictions for results in both orbital period and instellation (left panels of Figures 6 and 7, respectively). For the M2 and M3 meth-
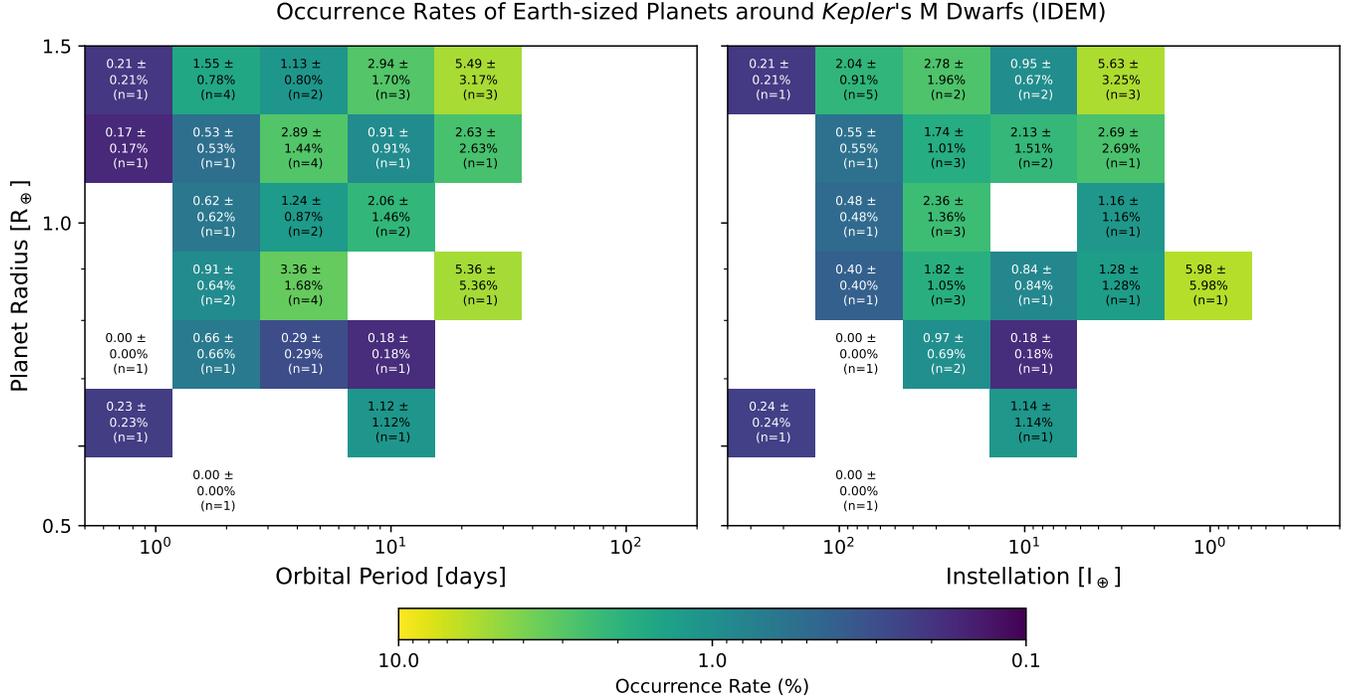
Figure 4. Occurrence rates of Earth-sized planets around M dwarfs distributed in **left:** orbital period and **right:** instellation, calculated via the reliability-weighted inverse detection efficiency method (Eqn. 4). White bins with a printed occurrence of $\eta = 0\%$ indicate bins that are technically populated, but by a candidate(s) with a reliability of 0.

Dimension	Method	F_0	α	P_{break}	β_1	β_2
Orbital Period	M1	$1.04^{+0.47}_{-0.30}$	$-0.32^{+0.61}_{-0.58}$	$4.10^{+4.55}_{-2.03}$	$0.76^{+1.00}_{-0.69}$	$-1.06^{+0.37}_{-0.69}$
	M2	$0.72^{+0.37}_{-0.22}$	$0.45^{+0.74}_{-0.69}$	$3.63^{+2.84}_{-1.57}$	$0.90^{+1.03}_{-0.57}$	$-1.08^{+0.37}_{-0.55}$
	M3	$\mathbf{0.77^{+0.39}_{-0.24}}$	$\mathbf{0.33^{+0.77}_{-0.73}}$	$\mathbf{3.64^{+3.10}_{-1.57}}$	$\mathbf{0.91^{+1.02}_{-0.57}}$	$\mathbf{-1.07^{+0.37}_{-0.58}}$
				I_{break}		
Instellation	M1	$1.01^{+0.39}_{-0.26}$	$-0.37^{+0.62}_{-0.60}$	$23.17^{+22.94}_{-18.60}$	$-0.89^{+1.05}_{-0.29}$	$-2.38^{+0.48}_{-0.63}$
	M2	$0.70^{+0.29}_{-0.19}$	$0.38^{+0.74}_{-0.69}$	$23.04^{+21.87}_{-19.31}$	$-0.76^{+2.40}_{-0.35}$	$-2.43^{+0.56}_{-0.73}$
	M3	$\mathbf{0.73^{+0.33}_{-0.20}}$	$\mathbf{0.30^{+0.79}_{-0.75}}$	$\mathbf{18.20^{+18.13}_{-14.61}}$	$\mathbf{-0.67^{+2.58}_{-0.43}}$	$\mathbf{-2.27^{+0.44}_{-0.69}}$

Table 1. Median and 1σ (16th and 84th percentile) parameter values for the optimized population models in orbital period and instellation. Rows in bold are from our most complete M3 method which includes both reliability and parameter uncertainties. While we fit the log of the break parameters (i.e., $\log_{10} P_{\text{break}}$ and $\log_{10} I_{\text{break}}$), we include the corresponding linear values here for readability. Note that because Equation (5) is defined in terms of dP and dR, a value of -1 in the power law exponents (α, β_1, β_2) corresponds to a flat line in natural log occurrence (i.e., $\frac{\text{d}P}{\text{d}\ln P} = P$). Additionally, because orbital period increases with distance from a star while instellation decreases, the far-out regime (beyond the break) is governed by β_2 in orbital period but by β_1 in instellation.

ods which incorporate candidate reliability, we found that the models from both methods exhibit 1σ agreement with the inverse detection efficiency method in all parameter bins where there are planet candidates (such that Equation 4 may be evaluated). For the M1 case which does not consider reliability, we re-evaluated Equation (4) while neglecting the rel_j term (i.e., the classical definition of the inverse detection efficiency method) and found that these values agree with the M1 model predictions within 1σ for all populated bins.

The radius distribution of occurrence values (right panels of Figures 6 and 7) between the models and the inverse detection efficiency method are typically consistent within 1σ. However, we note that smaller radius bins require $\sim 1.1\sigma$ to achieve consistency, and that there is appreciable variation between methods (compared to what is seen with the period or instellation distributions). We attribute these caveats to our small sample size and correspondingly sparse coverage when splitting the sample into bins of planet radius to evaluate
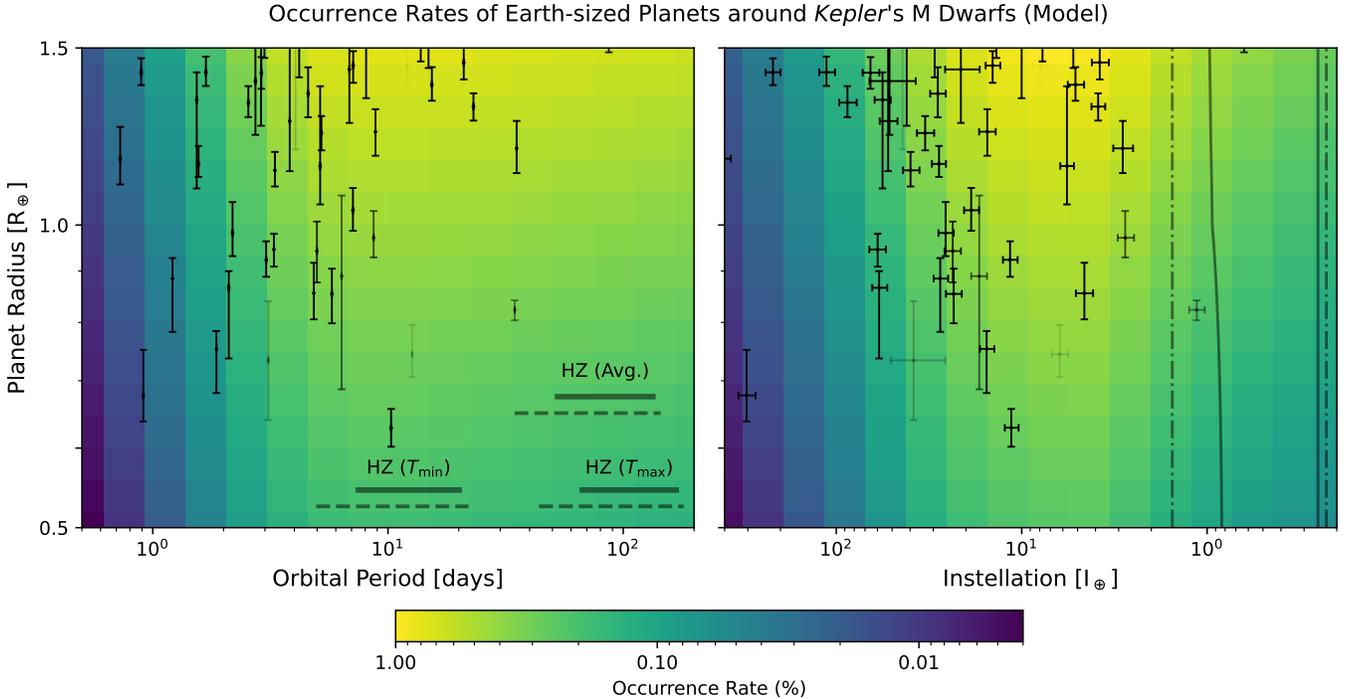
Figure 5. Modeled occurrence rates of Earth-sized planets around M dwarfs distributed in **left:** orbital period and **right:** instellation, calculated by evaluating Equation (5) with the `emcee`-output parameter vectors from method `M3` (plotted values represent median occurrence rates). Planet candidates from Figure (3) are reproduced here as points with error bars denoting their 1σ parameter uncertainties. Dashed gray lines represent the optimistic habitable zone, and solid gray lines represent the conservative habitable zone (Kopparapu et al. 2013, 2014). The three cases in orbital period represent the habitable zone bounds for the coldest ($T_{\min} \approx 2979\,\mathrm{K}$) and warmest ($T_{\max} \approx 4000\,\mathrm{K}$) stars in our sample, along with the average of the bounds for all stars in our sample.

with the inverse detection efficiency method. We further note that in the smallest planet radius bin ($\lesssim 0.6\,R_\oplus$), there is only one planet and it has a reliability score of zero, meaning that Equation (4) returns a zero occurrence rate (as shown in Figure 4).

Regarding differences between each of the model fitting methodologies, we found that (for a given dimension) model parameters are consistent at the 1σ level across all methods. There is a slight (but expected) global decrease in predicted occurrence between `M1` and `M2` characteristic of reliability incorporation, where the observed number of planets decreases due to some candidates being considered false positives or false alarms in a given inference. As in Bryson et al. (2020a), we found that input uncertainties have little effect when accounting for reliability, and thus the parameters between models fit via `M2` and `M3` are very similar. As there is no appreciable difference, we favor results from the more-complete `M3` approach incorporating both reliability and input uncertainties.

We note that the uncertainty contours on the population model results for the occurrence distribution in instellation (left panel of Figure 7) show a secondary

bump around $\sim 4\,I_\oplus$ separate from the power law's main break around $\sim 25\,I_\oplus$. In Figure (8) we highlight that the posterior distribution for the break parameter in instellation (modeled as $\log_{10} I_{\mathrm{break}}$) appears to follow a bimodal distribution, and is thus not well-represented by summary statistics that assume a unimodal (split normal) distribution. This is not true of the models in orbital period as the $\log_{10} P_{\mathrm{break}}$ distribution appears unimodal. We explore the possibility of the instellation bimodality representing physically distinct populations in Appendix (C), but ultimately conclude the feature is likely an artifact of our small sample size and sparse distribution. Nevertheless, having a model defined in instellation greatly simplifies evaluations of the habitable zone (Section 5.1) and enables comparisons across spectral types (Section 6). As such, we consider and compare occurrence estimates from both our period and instellation models where applicable throughout the remainder of this work.

5.1. *Habitable Zone Occurrence Rates*

We used the habitable zone models of Kopparapu et al. (2013) to define the optimistic (recent Venus out to early
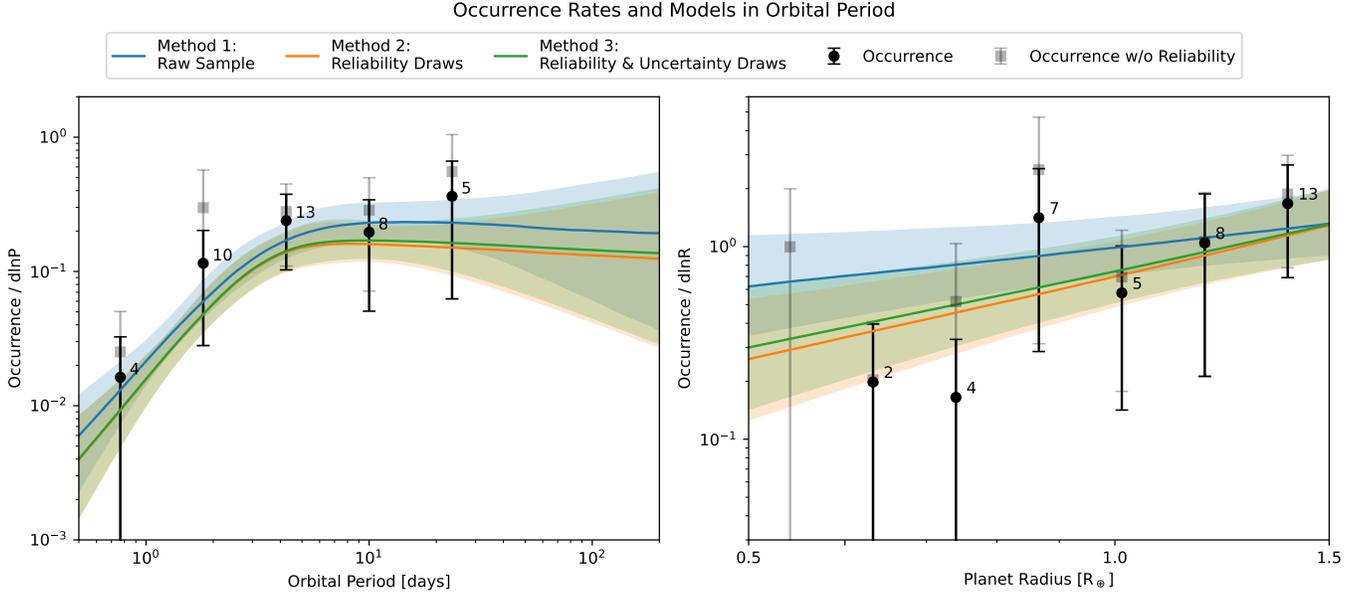
Figure 6. Best-fit population models and observed occurrence rates marginalized to show distributions in **left:** orbital period and **right:** planet radius. Individual points represent occurrence rates calculated via the inverse detection efficiency method (Eqn. 4), with and without accounting for reliability (black and gray points, respectively); numbers represent the number of candidates within each bin. For each method, the colored curve represents the median of the distribution of occurrence rates found by evaluating the population model with the distribution of parameter vectors; lighter filled regions denote the 1σ uncertainty envelopes. Colors represent the results of individual fitting methodologies: M1 (blue) considers only the median parameter values of planets within $[0.5, 200]$ days, $[0.5, 1.5]\,R_\oplus$ and treats all planets as fully reliable; *M2* (orange) is similar to M1 but repeatedly draws planets based on their reliability and concatenates the posteriors; *M3* (green) is similar to *M2* but also draws each planet's parameters from within their uncertainties, allowing for the inclusion of planets with median values outside the designated range.
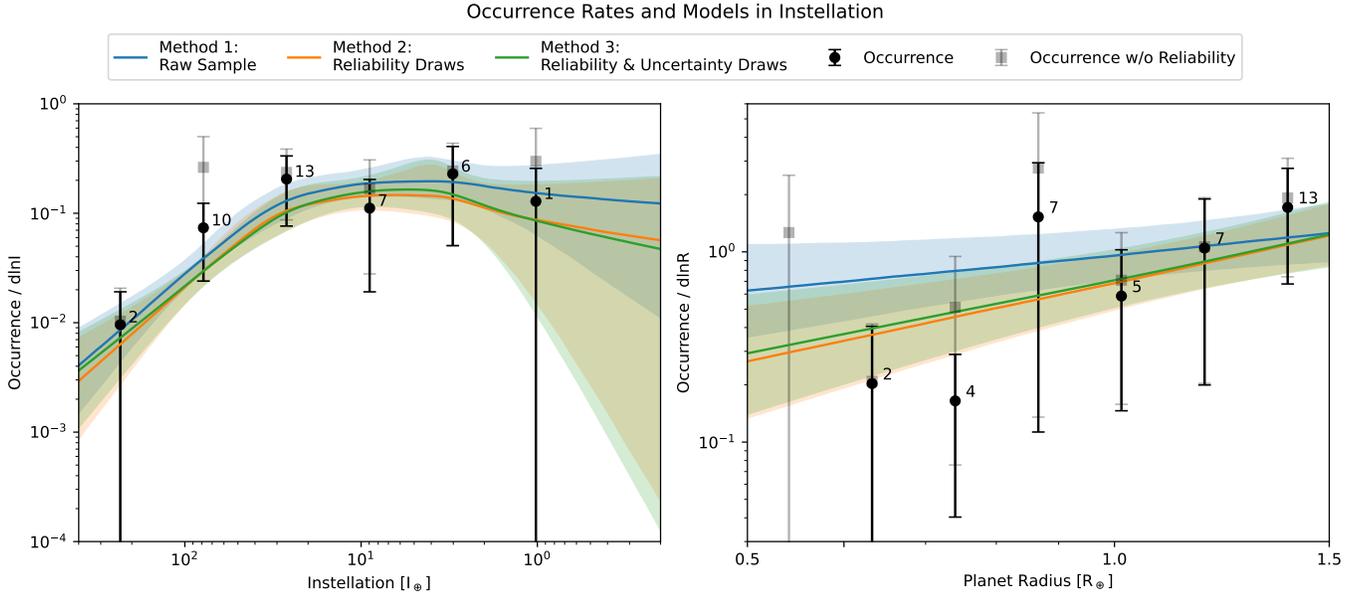
Figure 7. Similar to Figure (6) but with evaluations in installation (rather than orbital period) for the range of $[0.2, 400]\,I_\oplus$, $[0.5, 1.5]\,R_\oplus$.

Mars) and conservative (runaway greenhouse out to maximum greenhouse) boundaries. In Appendix (D), we

describe our considerations of a planet-mass-dependent modulation to the conservative inner (i.e., runaway

Posterior Distributions of Power Law Breaks
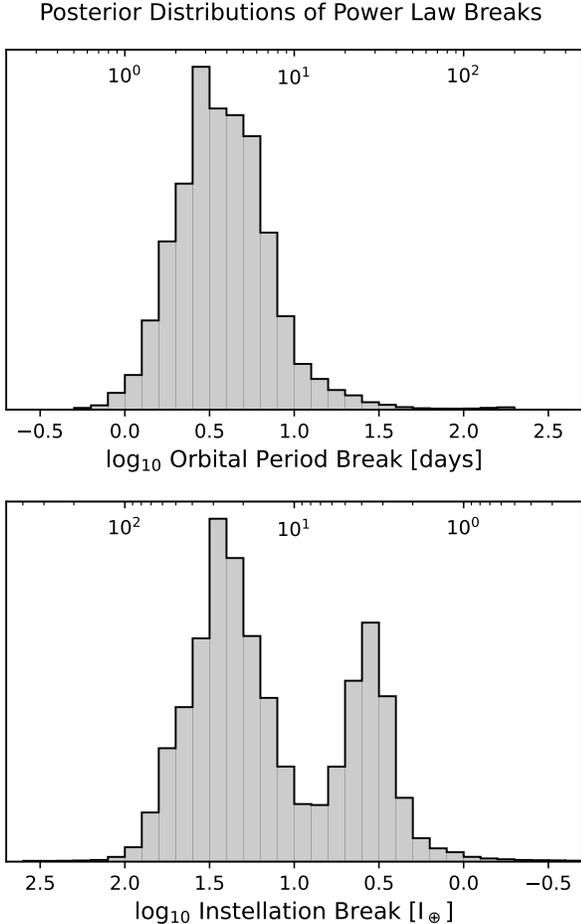
Figure 8. Posterior distributions for the $\log_{10}$ power law breaks in **top:** orbital period and **bottom:** installation from the M3 models. Tick marks on the top of each panel represent the corresponding non-logarithmic values. There is a bimodality in the installation break posteriors, meaning our model parameters are not well-constrained in this dimension; this feature is discussed further in Appendix (C).

greenhouse) edge. The Kopparapu et al. (2013) habitable zone bounds are near-constant in flux across the stellar effective temperature range of M dwarfs: the conservative inner and outer edges vary by $\sim 3\%$ and $\sim 12\%$ respectively between the warmest and coolest stars in our sample; the optimistic bounds similarly vary by $\sim 4\%$ and $\sim 14\%$. Due to this consistency, we are unlikely to benefit from a three-dimensional model that includes a stellar effective temperature dependence alongside installation and planet radius, as was used in Bryson et al. (2021) for Sun-like stars where the habitable zone varies more strongly with temperature.

However, the habitable zone bounds have a considerable range when converted to the corresponding orbital periods: the conservative innermost (outermost)

edge falls at roughly 8 (20) days for our coolest star ($T_{\min} \approx 2979\,\mathrm{K}$), but at 67 (168) days for our warmest star ($T_{\max} \approx 4000\,\mathrm{K}$). Because our stellar sample's effective temperature distribution is clustered towards hotter stars (Figure 2), the average conservative habitable zone boundaries of 52 (133) days are closer to the latter. We considered these three cases in our habitable zone calculations (for both conservative and optimistic bounds) in orbital period. While a three-dimensional model in orbital period, planet radius, and stellar effective temperature would have been useful here, fitting such a model would have required a greater number and coverage of systems than what is available with the current *Kepler* sample (see Appendix C.1).

To calculate each model's Earth-sized habitable zone planet occurrence rate ($\eta_\oplus$), we evaluated an integral of Equation (5) over the respective bounds (conservative and optimistic) for the three temperature cases in orbital period, and for the average case in installation. We considered the range of emcee-output parameter vectors, and report the median values with 1σ uncertainties in Tables (2). These values are also plotted in Figure (9) alongside previous estimates from DC15 who used the inverse detection efficiency method (without the rel_j term), and from Hsu et al. (2020) who used a forward model without per-candidate reliability. We stress that, due to the paucity of observed habitable zone planets, our results are predominantly *model-driven extrapolations*; a purely observational result derived from the reliability-weighted inverse detection efficiency method (Equation 4) is also included for comparison.

Taking the M3 orbital period model results evaluated over $[0.5, 1.5]\,\mathrm{R}_\oplus$ with the average-temperature habitable zone bounds, we found $\eta_\oplus = 13.89^{+16.70}_{-9.02}\%$ for the conservative case (and $21.75^{+23.44}_{-13.23}\%$ for the optimistic). For the M3 installation model, we found $\eta_\oplus = 8.58^{+17.94}_{-8.22}\%$ ($14.22^{+24.96}_{-12.71}\%$). These occurrence rates are consistent at the 1σ level with each other, the estimates of $15.82^{+16.60}_{-6.54}\%$ ($24.28^{+17.58}_{-8.39}\%$) for $[1.0, 1.5]\,\mathrm{R}_\oplus$ planets from DC15, and the larger prediction of $33^{+10}_{-12}\%$ for $[0.75, 1.5]\,\mathrm{R}_\oplus$ planets in the conservative habitable zone from Hsu et al. (2020).[4] Our results are further consistent with the upper limit of $\eta_\oplus < 23\%$ set by Pinamonti et al. (2022) using radial velocity data.

---

[4] Our model estimates are also consistent with literature estimates when using their unique radius ranges, included here for posterity. Evaluating our M3 models over $[1.0, 1.5]\,\mathrm{R}_\oplus$ gives $7.41^{+8.59}_{-4.78}\%$ ($11.60^{+12.00}_{-7.01}\%$) in orbital period and $7.54^{+12.74}_{-6.73}\%$ ($8.58^{+17.94}_{-8.22}\%$) in installation, all consistent with the estimates from DC15. Likewise, using $[0.75, 1.5]\,\mathrm{R}_\oplus$ gives $10.78^{+12.56}_{-6.97}\%$ ($16.88^{+17.58}_{-10.21}\%$) in orbital period and $8.58^{+17.94}_{-8.22}\%$ ($11.00^{+18.70}_{-9.83}\%$) in installation, consistent with Hsu et al. (2020).

Occurrence of Earth-sized Habitable Zone Planets around M Dwarfs
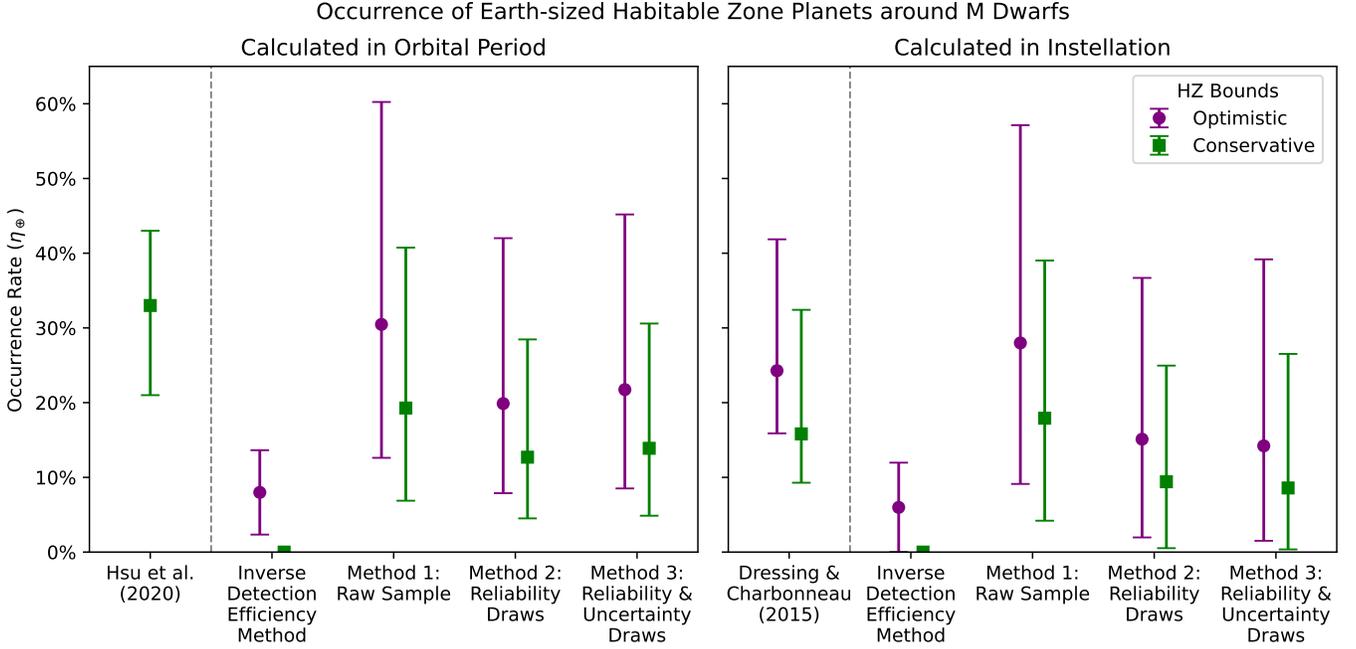
Figure 9. The occurrence rate of Earth-sized habitable zone planets around M dwarfs calculated in this work in **left:** orbital period and **right:** instellation. In either dimension, we include a comparison to previous results: the modeled results of Hsu et al. (2020) calculated in orbital period, or the results of Dressing & Charbonneau (2015) calculated in instellation with the inverse detection efficiency method. Purple and green points denote calculations for optimistic and conservative habitable zone bounds, respectively. For our plotted values in orbital period, we adopted the average habitable zone bounds for the stars in our sample; different stellar effective temperature cases are considered in Table (2).

We note that our uncertainties tend to be slightly larger than previous works: in the conservative case, the 1σ uncertainties from DC15 span $\sim 9-32\%$ and the estimate from Hsu et al. (2020) spans $\sim 21-43\%$, while our M3 orbital period estimate spans $\sim 5-31\%$ and the instellation estimate spans $\sim 0-27\%$. We attribute this increased uncertainty to the smaller number of habitable zone planet candidates and our subsequent need to rely on extrapolated models. The latter requires assumptions of a functional form and that the population continues to behave in some uniformly varying way across the extrapolated regime, which adds an inherent uncertainty not reflected in the statistical error bars of Figure (9). As such, despite comparable results to previous works, our new exploration of the *Gaia*-informed *Kepler* sample leads us to conclude that **the occurrence of Earth-sized habitable zone planets around M dwarfs is not as observationally constrained as previously thought**.

6. COMPARISONS ACROSS SPECTRAL TYPES

Previous studies have explored *Kepler*'s small planets for dependencies on stellar mass across the range of FGKM stars. Mulders et al. (2015b) studied the distribution of small $(1-4\,R_\oplus)$ planet occurrence rates with semimajor axis for different spectral types using

the planet catalog of Burke et al. (2014). They found that planet occurrence increased towards later spectral types, rising by a factor of ~ 2 from G to M stars and by a factor of ~ 3 from F to M stars. Mulders et al. (2015a) used the *Kepler* Q1-Q16 planet candidate sample (Mullally et al. 2015) to find a factor of ~ 3.5 increase from FGK to M stars for small $(1-2.8\,R_\oplus)$ planets while also noting a factor of ~ 2 decrease for larger $(>2.8\,R_\oplus)$ planets. These estimates were based on the close-in sample $(P<50\,\text{days})$, but have been used to predict $\eta_\oplus$ around M dwarfs by upscaling estimates from FGK stars (see e.g., Hardegree-Ullman et al. 2023), despite (a) the habitable zone falling at much longer orbital periods for the latter and (b) the inclusion of planets larger than $1.5\,R_\oplus$.

However, recent work suggests the factor of ~ 3.5 increase around M dwarfs may not be applicable to habitable zone occurrence rates. Employing the final DR25 sample of *Kepler* candidates along with *Gaia*-revised properties, Hsu et al. (2020) found that M dwarfs have higher occurrence rates than FGK stars (Hsu et al. 2019) at the same orbital periods, but that the two groups have comparable occurrence rates when evaluated at similar instellations. For example, they reported an M/FGK ratio of $3.1^{+5.5}_{-1.9}$ when evaluating their M and FGK models over the same period and radius range (averaged over

Occurrence of Earth-sized Habitable Zone Planets ($\eta_\oplus$, %)

	Orbital Period					
	Optimistic			Conservative		
	$T_{\min}$	Average	$T_{\max}$	$T_{\min}$	Average	$T_{\max}$
Bounds [days]	$5 - 22$	$34 - 144$	$44 - 181$	$8 - 20$	$52 - 133$	$67 - 168$
Observed Planets	14	2	0	6	0	0
IDEM	13.83 ± 3.70	8.00 ± 5.65	$-$	4.92 ± 2.01	$-$	$-$
M1	$35.11^{+11.59}_{-8.47}$	$30.47^{+29.77}_{-17.86}$	$29.36^{+34.02}_{-19.05}$	$23.47^{+8.90}_{-5.75}$	$19.28^{+21.47}_{-12.41}$	$18.70^{+24.05}_{-13.05}$
M2	$24.41^{+8.11}_{-6.13}$	$19.88^{+22.13}_{-11.99}$	$19.27^{+25.14}_{-12.55}$	$15.97^{+5.66}_{-4.16}$	$12.70^{+15.76}_{-8.19}$	$12.25^{+17.70}_{-8.42}$
M3	$25.91^{+8.93}_{-6.63}$	$21.75^{+23.44}_{-13.23}$	$21.05^{+26.59}_{-13.79}$	$17.11^{+6.36}_{-4.61}$	$13.89^{+16.70}_{-9.02}$	$13.38^{+18.79}_{-9.26}$
	Instellation					
	Optimistic			Conservative		
Bounds [$I_\oplus$]	$0.23 - 1.54$			$0.25 - 0.88^{\mathrm{a}}$		
IDEM	6.01 ± 6.01			$-$		
M1	$27.99^{+29.14}_{-18.87}$			$17.92^{+21.10}_{-13.72}$		
M2	$15.11^{+21.59}_{-13.15}$			$9.42^{+15.54}_{-8.89}$		
M3	$\mathbf{14.22^{+24.96}_{-12.71}}$			$\mathbf{8.58^{+17.94}_{-8.22}}$		

Table 2. Earth-sized habitable zone occurrence ($\eta_\oplus$, in %) estimates for M dwarfs calculated in orbital period and instellation. Because the habitable zone in orbital period varies greatly with spectral type, we provide three cases for either the optimistic or conservative bounds: the bounds for the coolest ($T_{\min} \approx 2979\,\mathrm{K}$) or warmest ($T_{\max} \approx 4000\,\mathrm{K}$) stars in our sample, and the average bounds across all stars in our sample. The habitable zone in instellation is roughly constant with spectral type, so we provide only one case evaluated for the average stellar effective temperature of our sample. [a]Our incorporation of a planet radius dependence to the runaway greenhouse edge causes this value to vary from $0.83\,I_\oplus$ at $R = 0.5\,R_\oplus$ to $0.97\,I_\oplus$ at $R = 1.5\,R_\oplus$.

$[0.5, 256]$ days, $[0.5, 4.0]\,R_\oplus$), but a ratio of $0.9^{+1.6}_{-0.2}$ while evaluating at similar instellations by scaling their FGK grid from a G2 star to match the instellations of an M2.5 star.

Petigura et al. (2022) modeled occurrence distributions in both orbital period and instellation for three bins of stellar mass with boundaries of $\{0.5, 0.7, 1.0, 1.4\}\,\mathrm{M}_\odot$ using the California-*Kepler* Survey with stellar properties derived from spectroscopic measurements (independent of *Gaia*). Referencing their Figure (13), their models of sub-Neptune ($1.7 - 4.0\,\mathrm{R}_\oplus$) occurrence rates suggest a statistically distinct factor of $\sim 3 - 4$ offset between the lowest and highest bins at the longest periods or lowest instellations, though the underlying separation-binned number of planets per star for each stellar mass bin are not distinct at the 1σ level. For super-Earths ($1 - 1.7\,\mathrm{R}_\oplus$), their models were cut off before reaching the habitable zone due to sparse detections at low instellations, and the underlying binned estimates (including some upper limits) were again not statistically distinct. Combined with Hsu et al. (2020), these results call into question whether there is sufficient evidence supporting higher occurrence of *Earth-sized* planets at habitable zone instellations around lower-mass stars.

Efforts to resolve potential stellar mass dependencies in the occurrence distribution of small planets face more difficulties when using instellation as opposed to orbital period or semimajor axis. In addition to requiring precise measurements of both stellar mass and luminosity to calculate instellation flux, the true scale factor between occurrence rates around different spectral types may only be resolvable at low instellations (where candidate detections are currently sparse). This is because the occurrence distributions resembling broken power laws appear to "break" at similar orbital periods (see e.g., Petigura et al. 2022; Bergsten et al. 2022) across FGKM stars, such that the breaks occur at different values of semimajor axes (Mulders et al. 2015b) or instellations (Petigura et al. 2022). Instellation exhibits a more prominent offset between spectral types due to a stronger stellar mass dependence when converting from orbital period,[5] meaning that the distribution for higher-mass stars breaks and subsequently "plateaus" (depending on the power law exponent) at higher instellations where the distribution for lower-mass stars may still be exhibiting a steep rise.

The resulting trend is that higher-mass stars may appear to have slightly higher occurrence rates at high in-

---

[5] From Equations (1) and (2), $a \propto M_*^{1/3} P^{2/3}$ while $I \propto L_* M_*^{-2/3} P^{-4/3}$. Since the mass-luminosity relation $L_* \propto M_*^\xi$ for main sequence dwarfs always has $\xi > 1$, the orbital period-instellation conversion has the stronger stellar mass dependence.

stellations but lower rates at low installations compared to lower-mass stars (a trend which is not observed when using semimajor axis or orbital period). Only the trend at low installations is relevant to the habitable zone, and thus only candidates at low installations can reveal a potential stellar mass dependence. However, the characteristic power law break for low mass stars may fall at installations where very few relevant candidates have been detected. From Petigura et al. (2022), plateaus should arise by $\sim 10\,I_\oplus$ for stars down to $0.5\,M_\odot$, while less-massive stars would presumably flatten out at even lower installation values. Yet *Kepler*'s limited sensitivity to Earth-sized planets orbiting M dwarfs at installations less than $10\,I_\oplus$ (Figure 3) leaves very few detections with which to characterize this regime.

Despite the aforementioned complications, installation may be the most physically relevant dimension for studies in the context of the habitable zone and/or planet formation. Incident flux from the host star likely contributes to shaping the planet formation environment within a protoplanetary disk – consider e.g., the disk's temperature profile and the location of the snow line. As cautioned in Hsu et al. (2020), we note that stellar luminosities are higher on the pre-main sequence than on the main sequence (especially true for M dwarfs), such that the incident flux a planet receives is different between the time of formation and the present day. Additionally, M dwarfs remain on the pre-main sequence at the time of planet formation, such that the relation between disk temperature, installation and stellar mass differs from pre-main sequence to main sequence. Nevertheless, installation may still offer a slightly more native tracer of the formation environment compared to semimajor axis or orbital period which include additional stellar mass dependencies. Furthermore, the location of the habitable zone for an Earth-like atmosphere is set by the incident stellar flux (Kopparapu et al. 2013), and thus the habitable zone falls at comparable installations (but different semimajor axes and orbital periods) for stars of various effective temperatures. This is true both within the M dwarf regime (see Table 2), and across the range of FGKM stars, such that installation is the most well-suited dimension for comparisons across spectral types.

6.1. *No Evidence for an Increase in $\eta_\oplus$ between M and FGK stars*

Before we present a comparison of $\eta_\oplus$ estimates between M and FGK stars, we find it useful to first discuss the current limitations of *Kepler* data regarding Earth-sized ($[0.5, 1.5]\,R_\oplus$) planets around different spectral types. To this end, we divided the Berger et al. (2020b) stellar sample and subsequent planets into M,

K, G, and F bins of stellar effective temperatures with bounds of $\{2310, 3890, 5325, 5960, 7310\}$ K following the spectral type definitions of Pecaut & Mamajek (2013); Mamajek (2022). For each spectral type, we adopted a bin width equivalent to that of the conservative habitable zone for that bin's median stellar effective temperature (Kopparapu et al. 2013). We then slid this bin over a continuum of installation values, and present the number of detected Earth-sized planets as a function of installation in the upper panel of Figure (10).

As mentioned in Section (1), the limited duration of *Kepler*'s observations meant that the survey was generally unable to provide confident detections in the habitable zone for Sun-like stars. The longest orbital period measurable within *Kepler*'s duration while still meeting the three-transit detection requirement is ~ 710 days. However, the inner edge of the conservative habitable zone exceeds this limit for $\sim 30\%$ of *Kepler*'s F-type stars, such that *Kepler* outright lacked the coverage to detect habitable zone planets around many hotter stars (see e.g., Figure 1 of Bryson et al. 2021). For example, there were zero detections of Earth-sized planets around F-type stars at installations less than $5\,I_\oplus$; more generally, all spectral types have relatively few detections at less than $10\,I_\oplus$.

Because of the paucity of detected Earth-sized planets at low installations, the *Kepler* sample is inherently ill-suited for direct estimations of occurrence rates in this regime. To illustrate this point, we used the aforementioned sliding bins and the inverse detection efficiency method (Equation 4) to create an occurrence distribution as a function of installation for each spectral type, shown in the bottom panel of Figure (10). The inverse detection efficiency method is unable to provide occurrence estimates in unpopulated bins, and thus offers the most native visualization of where occurrence estimates are currently motivated by available data.

Compared to the well-sampled close-in regime, occurrence rates at low installations are more uncertain (and trends therein more inconsistent), especially at the plateau-relevant separations ($< 10\,I_\oplus$) necessary to resolve stellar mass dependencies. While there is evidence for increased occurrence around M dwarfs (relative to FGK stars) at $10\,I_\oplus$, it is unclear how this trend persists at habitable zone installations an order of magnitude lower, due to "noise" from the generally sparse sample and the lack of relevant detections for F-type stars. As such, we argue that **the current *Kepler* sample lacks sufficient evidence to inflate $\eta_\oplus$ between FGK and M stars, as the paucity (or lack) of detections for Earth-sized habitable zone plan-**

Occurrence for [0.5, 1.5] R$_\oplus$ Planets
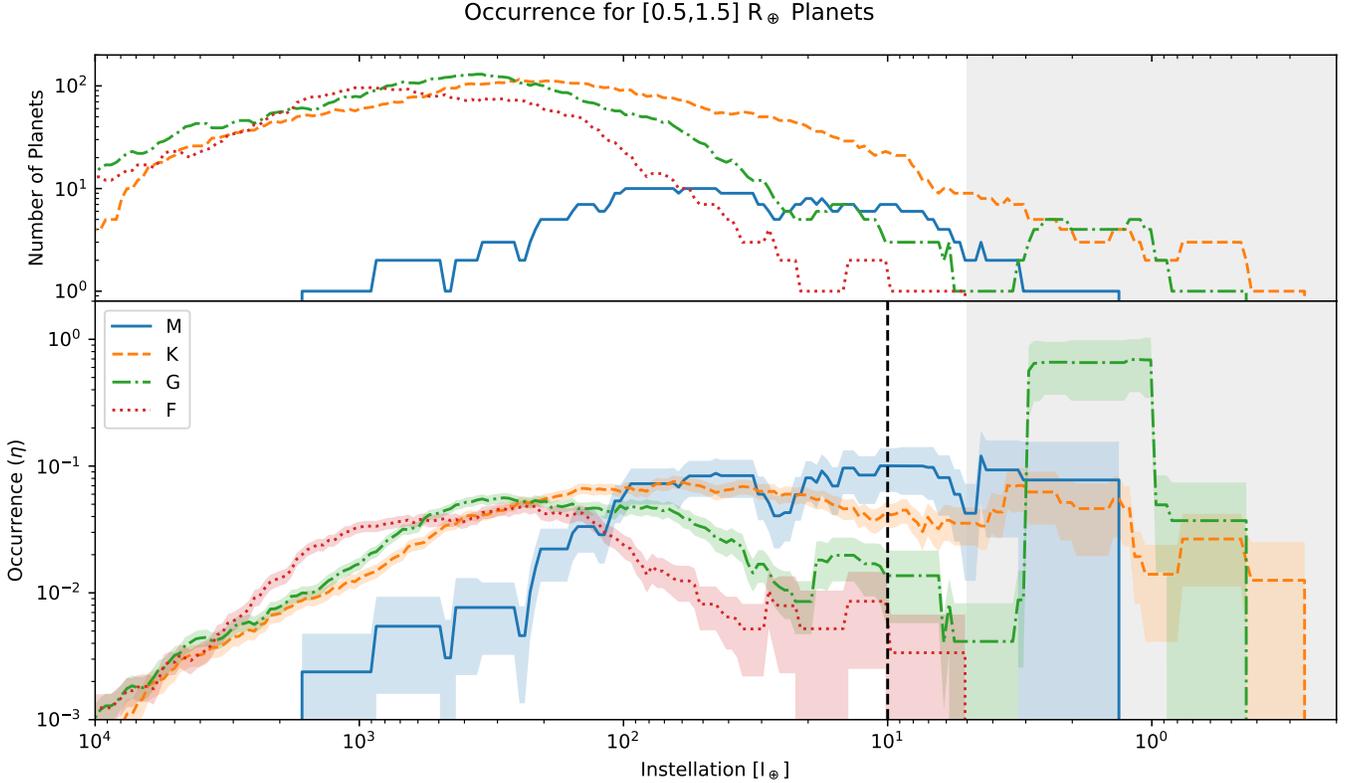
Figure 10. Because of *Kepler*'s poor sensitivity to habitable zone planets (around any star), the current sample does not offer evidence to support M dwarfs hosting more Earth-sized planets than FGK stars *at habitable zone instellations.* **Top:** The number of Earth-sized ([0.5, 1.5] R$_\oplus$) *Kepler* planet candidates orbiting M (solid blue), K (dashed orange), G (dot-dashed green) or F (dotted red) dwarf stars as a function of instellation. The shaded grey region denotes the regime where there are no detections of Earth-sized planets around F type stars, such that FGKM comparisons require models and/or extrapolations. **Bottom:** Earth-sized planet occurrence rates calculated via the reliability-weighted inverse detection efficiency method (Equation 4) over a continuum of habitable zone-width instellation bins. Vertical black dashed line represents the approximate point where the Petigura et al. (2022) occurrence distribution plateaus for [0.5, 0.7] M$_\oplus$ stars, indicating the regime where comparisons between spectral types would not be affected by differences in the break points of their power law distributions.

ets means that a purely data-motivated comparison remains elusive.

We thus turn to population models, which offer a way to provide globally informed estimates in these low-detection regimes. However, we caution that model-driven estimates can require assumptions of functional forms and how those forms behave in regions of sparse or absent data. This can obfuscate insight on what is presently motivated by available data, such as the *Kepler* sample's limitations discussed above. With these caveats stated, we now draw comparisons between our $\eta_\oplus$ estimate for M dwarfs and recent literature estimates for Sun-like stars.

Our M3 instellation-modeled occurrence rate of $\eta_\oplus = 8.58^{+17.94}_{-8.22}\%$ in this work is similar to the FGK-based value of $\eta_\oplus = 9.37^{+3.40}_{-2.48}\%$ found in Bergsten et al. (2022) which also incorporated reliability and *Gaia*-revised properties, though the former is considerably more uncertain. The latter was modelled in orbital pe-

riod rather than instellation but used the same definition of the habitable zone, and employed a different range for Earth-sized planets ([0.7, 1.5], R$_\oplus$). Normalizing both values by the $d \ln R$ of their respective radius ranges still provided consistent results of $7.81^{+16.33}_{-7.48}\%$ and $12.29^{+4.46}_{-3.25}\%$, respectively. We note that the ~ 3.5 difference from Mulders et al. (2015a) is still possible beginning at the $\sim 1.8\sigma$ level (~ 96th percentile), found by treating both radius-normalized estimates as idealized split-normal distributions and taking their ratio.

The $\eta_\oplus$ estimate of Bergsten et al. (2022) is something of a small-to-intermediate estimate compared to other recent and similarly comprehensive works (see their Section 4.2.1 for details). On the larger side, Bryson et al. (2021) also employed DR25 and *Gaia*-revised stellar properties along with treatments of completeness and reliability. They used a different parameterization from that of Bergsten et al. (2022), opting for a three-dimensional model with dependencies on in-

stellation, planet radius, and stellar effective temperature. For $[0.5, 1.5]\,R_\oplus$ planets in a conservative habitable zone around stars with $3900 < T_{\rm eff} < 6300\,{\rm K}$, their $\eta_\oplus$ estimate fell between $37^{+48}_{-21}\%$ and $60^{+90}_{-36}\%$. Some more recent estimates including both reliability and *Gaia*-revised properties include $\eta_\oplus = 11^{+7}_{-6}\%$ from Kunimoto & Matthews (2020) for $[0.75, 1.5]\,R_\oplus$ planets across $[0.99, 1.70]$ au for G-type stars, and $\eta_\oplus = 12^{+8}_{-5}\%$ from Bryson et al. (2020b) for $[0.5, 1.5]\,R_\oplus$ planets across $[237, 860]$ days orbiting GK stars. Estimates that include *Gaia*-revised properties but not reliability include $\eta_\oplus = 16^{+9}_{-6}\%$ from Hsu et al. (2019) for $[0.75, 1.5]\,R_\oplus$ across $[237, 500]$ days orbiting FGK stars, and $\eta_\oplus = 5^{+7}_{-4}\%$ from Pascucci et al. (2019) for $[0.7, 1.5]\,R_\oplus$ planets across $[330, 803]$ days orbiting G-type stars (their Model #6). Despite the range of recent FGK $\eta_\oplus$ estimates, our estimate of $\eta_\oplus = 8.58^{+17.94}_{-8.22}\%$ for M dwarfs has such large uncertainties that it is consistent with all quoted results within 1σ. To this end, even model estimates do not currently offer justification to inflate $\eta_\oplus$ between FGK and M dwarf stars.

While FGK stars are grouped together in this discussion, it is worth noting that the habitable zone boundaries for FGK stars exhibit a strong dependence on stellar effective temperature (in either instellation or orbital period). This causes the (log) width of the habitable zone to change considerably across the FGK regime (Kopparapu et al. 2013). Corrections for this may include modeling a stellar effective temperature dependence (Bryson et al. 2021) or normalizing $\eta_\oplus$ by both $d\ln P$ and $d\ln R$ (i.e., $\Gamma_\oplus$), although Bergsten et al. (2022) found no statistically significant trend across FGK stars with the latter.

6.2. *Regarding Occurrence Trends within M Dwarf sub-Spectral Types*

Previous works have suggested that small planet occurrence may exhibit a dependence on *sub*-spectral type across the broad range of M dwarf classifications, though the exact behavior is presently not well-constrained. Hardegree-Ullman et al. (2019) studied small planets around *Kepler*'s mid-M dwarfs using revised stellar properties, and found their median occurrence rates to increase from M3 to M5 spectral types (though all values were consistent at 1σ). On the theoretical side, Mulders et al. (2021) predicted that planet occurrence peaks around early ($\sim 0.5\,M_\odot$, roughly M1) M dwarfs and decreases towards later sub-spectral types. This decrease in occurrence may explain why recent efforts searching *K2*'s $\sim$M5.5-M9.5 stars (Sagear et al. 2020; Sestovic & Demory 2020) or the EDEN survey's volume-limited sample of M7-M9 stars (Dietrich et al. 2023)

have not identified *any* Earth-sized planets (the latter only studying planets with $P < 1$ day). It is worth noting that late M dwarfs are quite faint and thus difficult to observe with current sensitivity limitations, such that even the Transiting Exoplanet Survey Satellite (*TESS*, Ricker et al. 2015) may miss out on detecting tens of transiting planets around nearby late M dwarfs (Brady & Bean 2022; Dietrich et al. 2023).

Hardegree-Ullman et al. (2019) estimated an occurrence rate of $0.99^{+0.66}_{-0.50}$ for $[0.5, 1.5]\,R_\oplus$ planets within $[0.5, 10]$ days, calculated using their combined sample of 13 *Kepler* planets around seven mid-M dwarfs. We evaluated our M3 orbital period model across the same range to yield a statistically distinct estimate of $0.26^{+0.07}_{-0.06}$. The inconsistency between estimates could potentially be attributed to an overestimation caused by their simplified treatment of detection efficiency, and/or the very small sample size in Hardegree-Ullman et al. (2019) leading to a non-robust measurement.

The mid-M dwarf regime is also bright enough to be studied by *TESS*, offering additional insight beyond *Kepler*. Ment & Charbonneau (2023) used *TESS* to study a sample of mid-to-late M dwarfs (M4 to M7) with a median stellar mass of $0.17\,M_\odot$, and identified a sample of seven planets. Spanning $[0.5, 7]$ days orbital period, they found an occurrence rate of $0.13^{+0.12}_{-0.07}$ for $[0.5, 1.0]\,R_\oplus$ planets and $0.45^{+0.19}_{-0.14}$ for $[1.0, 1.5]\,R_\oplus$. Our sample (including all dwarfs below $4000\,{\rm K}$) has a median mass of $0.50 M_\odot$ such that it is dominated by earlier M dwarfs, and we found $0.09^{+0.04}_{-0.03}$ and $0.10^{+0.03}_{-0.02}$ respectively for the same period and radius ranges using our M3 models in orbital period. Despite the different spectral types probed by these studies, our results are consistent with Ment & Charbonneau (2023) in the $[0.5, 1.0]\,R_\oplus$ bin. Our estimate is considerably smaller for $[1.0, 1.5]\,R_\oplus$ planets, which could be attributed to our use of a single power law to describe the radius distribution, while Ment & Charbonneau (2023) used a normal distribution that peaks in the $[1.0, 1.5]\,R_\oplus$ bin.

However, while comparing the same radius bins evaluated over $[4, 200]\,I_\oplus$ in instellation, Ment & Charbonneau (2023) reported $0.11^{+0.10}_{-0.06}$ and $0.37^{+0.16}_{-0.12}$ while we found $0.18^{+0.08}_{-0.06}$ and $0.21^{+0.06}_{-0.05}$. Our results are consistent with theirs at the 1σ level in either radius bin, potentially suggesting that Hsu et al. (2020)'s finding of FGK/M stars having comparable occurrence at similar instellations also applies to an early/mid-to-late M dwarf comparison *for Earth-sized planets*. The size stipulation is critical: because our modeling efforts focused solely on Earth-sized planets, we cannot offer insight to Ment & Charbonneau (2023)'s comparison of $[0.5, 4]\,R_\oplus$ planet occurrence rates between their mid-to-late sam-

ple and DC15 representing early Ms,[6] nor the results of Hardegree-Ullman et al. (2019) for mid-Ms which included planets up to $2.5\,R_\oplus$. While incorporating larger planets up to $4\,R_\oplus$ would enable more detailed comparisons, this would likely require more complicated functional forms to properly address the full radius distribution or coupled dependencies, and is thus beyond the scope of this work.

Our focus on a relatively small radius regime left a limited number of detected planets which, combined with our top-heavy sample of host star temperatures (Figure 2), meant our study was unable to resolve any relationship between sub-spectral type and the occurrence of Earth-sized planets using Kepler. Details of our attempt to probe such a dependence and limitations therein are discussed further in Appendix (C.1). To address these shortcomings, we discuss the potential of future studies employing data from Kepler and additional surveys in the following subsection.

6.3. Improving M Dwarf Occurrence Rates with Additional Surveys

The K2 (Howell et al. 2014) and TESS (Ricker et al. 2015) missions will likely help to further constrain $\eta_\oplus$ around M dwarfs by providing more observations of Earth-sized habitable zone planets. Similar to the Gaia-Kepler Stellar Properties Catalog of Berger et al. (2020b), Hardegree-Ullman et al. (2020) provided a uniform catalog of Gaia-informed stellar properties for the K2 population. The K2 planet candidate sample of Zink et al. (2021) also contains 47 $[0.5, 1.5]\,R_\oplus$ planets around $T_{\rm eff} < 4000\,K$ stars, including two Earth-sized candidates: a $0.51\,R_\oplus$ planet in the conservative habitable zone, and a $1.20\,R_\oplus$ planet in the optimistic (see their Table 1). Zink et al. (2023) already combined the Kepler and K2 samples to study close-in planets around FGK stars, and similar integrated studies in the future may be well-poised to (re)address occurrence comparisons between FGK/M stars, and explore occurrence trends between early, mid, and late M dwarfs as discussed in Section (6.2).

A single TESS sector is observed for ~ 27 days, which (requiring two transits for detectability) probes the habitable zones up to M4.5 dwarf stars. The Continuous Viewing Zone at the ecliptic poles covers a smaller area of sky and thus views less stars, but it is observed for

~ 350 days which provides a long enough baseline to survey the habitable zone even for M0 stars. While later spectral types offer habitable zones at shorter orbital periods, they also exhibit increased stellar activity (Robertson et al. 2013; West et al. 2015; Astudillo-Defru et al. 2017), and the resulting increase in photometric noise may contribute to TESS's unexpectedly low yield of candidate detections around later M dwarfs (Brady & Bean 2022). Nonetheless, the study of TESS's mid-to-late M dwarfs (M4 to M7) from Ment & Charbonneau (2023) estimated a $\sim 50\%$ sensitivity to $1\,R_\oplus$ planets at 7 days (or $4\,I_\oplus$; their Figure 6), such that there may be appreciably non-zero sensitivity at the slightly larger habitable zone separations needed to probe $\eta_\oplus$ and subsequent stellar mass dependence. Future survey missions such as PLATO (Rauer et al. 2014, 2016) and the Nancy Grace Roman Space Telescope (Spergel et al. 2015; Akeson et al. 2019) may also help to constrain $\eta_\oplus$ by enriching the sample of small planets around M dwarfs.

On a more contemplative note, future studies may be able to leverage abundant planet detections to adopt a more physical definition of an "Earth-sized" planet beyond a generic $[0.5, 1.5]\,R_\oplus$ bin. For example, Hardegree-Ullman et al. (2023) adopted an instellation-dependent lower bound set by the minimum planet mass capable of retaining an Earth-like atmosphere (Zahnle & Catling 2017; Bixel & Apai 2021), and an upper bound where a planet orbiting an M dwarf is likely to have some particularly hospitable atmospheric composition (Kimura & Ikoma 2022). An alternative upper bound could be the limit beyond which small planets are no longer rocky in composition, conservatively interpreted as $1.4\,R_\oplus$ (Rogers 2015) as used in the LUVOIR (The LUVOIR Team 2019) and HabEx (Gaudi et al. 2020) mission concept studies. These may be considered in addition to coupled size-instellation boundaries like the planet mass-dependent inner edge from Kopparapu et al. (2014) used in this work (see Appendix D), and/or treatment of atmospheric factors such as clouds that may vary between Earth- and super-Earth-sized worlds (e.g., Windsor et al. 2023). If/when model uncertainties are reduced enough such that these considerations would have meaningful effects on $\eta_\oplus$ estimates, it may be worth building towards a more physically meaningful set of integral bounds for habitable zone occurrence rates.

7. SUMMARY

Since the work of Dressing & Charbonneau (2015), Gaia updates to stellar properties have changed our understanding of Kepler's M dwarfs and their sample of planet candidates, including the Earth-sized habitable

---

[6] The DC15 sample was originally believed to have a median stellar effective temperature of 3746 K (roughly M0.5; Mamajek 2022). Our Gaia-informed updates to the DC15 sample described in Section (2) suggest an actual median temperature of 4096 K (roughly K7).

zone planets used to estimate $\eta_\oplus$. Here, we presented an updated investigation of the current *Kepler* sample, fitting separate population models in orbital period and instellation with various considerations of reliability and parameter uncertainties.

- Using *Kepler* DR25, *Gaia*-updated parameters and candidate reliability, we found that the updated *Kepler* sample has few detected candidate planets in the habitable zone. As such, the inverse detection efficiency method cannot be employed as in the past to calculate $\eta_\oplus$.

- Integrating our best-fit instellation model incorporating both reliability and parameter uncertainties, we estimated an occurrence rate for Earth-sized habitable zone planets of $\eta_\oplus = 8.58^{+17.94}_{-8.22}\%$ for the conservative habitable zone (and $14.22^{+24.96}_{-12.71}\%$ for the optimistic). Our orbital period model predicted a slightly larger median value of $13.89^{+16.70}_{-9.02}\%$ ($21.75^{+23.44}_{-13.23}\%$) which is consistent within the 1σ uncertainties.

- The updated sample's paucity of Earth-sized habitable zone planet detections means that *Kepler* offers no evidence supporting an increased $\eta_\oplus$ around M dwarfs compared to FGK stars. This also applies to model estimates in the literature, as our $\eta_\oplus$ value for M dwarfs is consistent with those based on FGK stars.

We note that, in either dimension, the uncertainties on our $\eta_\oplus$ estimates are typically larger than previous works. This is largely due to the small number of *Kepler* Earth-sized candidates at larger orbital periods or lower instellations, requiring us to rely on extrapolated models whose parameters suffer similar uncertainties from the paucity of detections. The median values are also generally lower than previous works, which may be (at least partially) attributed to our consideration of candidate reliability. Compared to our model(s) in orbital period, our instellation model is less well-defined (Figure 8) and provides occurrence estimates with larger uncertainties. Yet instellation offers a more native tracer of the planet formation environment and processes defining the habitable zone, and is thus our preferred dimension for population models and comparisons across spectral types.

When evaluated in instellation, we found a lack of sufficient evidence that would support M dwarfs having more Earth-sized planets than FGK stars at habitable zone instellations. This contrasts with studies of close-in planets evaluated in orbital period and/or semimajor axes, leaving open the possibility that scaling habitable zone occurrence rates with spectral type may not be justified. We also did not find a significant difference between occurrence rates for our predominantly early M dwarf sample and those of mid-to-late M dwarfs, though we could only compare estimates for Earth-sized planets where our models are defined. Future studies with *K2* and *TESS* should be able to further probe spectral type dependencies, especially in the habitable zones of M dwarfs.

APPENDIX

A. FALSE POSITIVE DISPOSITIONS

A catalog of homogeneously characterized objects is essential to the uniform analysis required for demographic studies. *Kepler*'s DR25 satisfies this requirement by virtue of its automation, but we nevertheless acknowledge that some objects automatically classified as false positives in DR25 may have received new classifications from subsequent works. We briefly discuss the sample of relevant objects here, but stress that any updated dispositions do not come from uniform treatments applied to the entire planet sample and are thus not used in our demographic study.

The Exoplanet Archive's table of *Kepler* Certified False Positives summarizes the efforts of the *Kepler* False Positive Working Group to revisit and more rigorously classify potential false positives in the *Kepler* sample (Bryson et al. 2017). Of the five Dressing & Charbonneau (2015) planets labeled false positives in DR25, one was a certified false positive and four were "potential planets," although one star was not included in Berger et al. (2020b) and two others received updated $T_{eff} > 4000\,\mathrm{K}$. The one remaining planet, K00961.02 (Kepler-42 c), is listed as a confirmed planet in the Exoplanet Archive. In our updated investigation based on the full DR25 sample (Section 3 onward), there were nine DR25 false positives that orbited an M dwarf (via Berger et al. 2020b) and had planet parameters within our sample domain (in orbital period, installation, or both). Other than the aforementioned K00961.02, six of these were certified false positives, and the remaining two were "possible planets," although one had a reliability score of $\sim 1\%$ and is thus unlikely to be real. Again the remaining planet, K03138.02 (Kepler-1649 c), is listed as a confirmed planet in the Exoplanet Archive. Even though K03138.02 has a habitable zone installation of $0.94 \pm 0.09\,I_{\oplus}$, we reiterate that its disposition status in the literature cannot supersede its DR25 classification by virtue of the uniformity requirement necessitated for demographic study.

B. DISCUSSING NUMERICAL CHOICES RELATED TO MODEL OPTIMIZATION

All numbers of total and discarded steps in this work were chosen based on autocorrelation analysis with emcee. In general, we required the total chain length to surpass at least 50 times the autocorrelation time τ, and discard at least 2τ as burn-in. For optimization with the M2 and M3 approaches, the number of iterations was the same as those used in Bryson et al. (2021): 100 draws when including reliability, and 400 when including both reliability and input uncertainties. While we lack a quantitative justification, we found that 100 inferences suitably incorporates candidates at a rate linearly proportional to their reliability. Similarly, 400 inferences was a suitable number of iterations such that unreliable candidates could be included in enough draws to also sample the range of their uncertainties.

C. INVESTIGATING THE BIMODALITY IN INSTELLATION

Here, we investigated the cause of the installation bimodality shown in Figure (8), and discuss whether this feature can be attributed to two physically distinct populations or an intrinsic scatter from our relatively sparse sample. Figure (3) shows a relative paucity of $[1.0, 1.5]\,R_{\oplus}$ planet detections around $10\,I_{\oplus}$ separating two clusters of observed planets, roughly corresponding to the trough between the two peaks in the $\log_{10} I_{break}$ posteriors of Figure (8). This gap in detections manifests itself as a slight decrease in occurrence around $10\,I_{\oplus}$ (see estimates from the inverse detection efficiency method in Figure 7) – such that a broken power law could place the break (a transition between rising and $\sim$plateauing occurrence) on either side of this dip – but this variation is not statistically significant.

We explored a potential physical explanation for the installation bimodality. Given the unimodal distribution of $\log_{10} P_{break}$ posteriors, if the power law break occurs at the same orbital period for all M dwarfs, then this break could manifest at two different installations for two distinct enough groups of host stars (distinguished on the basis of stellar mass, luminosity and effective temperature). To test if this was the case, we split the M dwarf sample at the median stellar effective temperature ($\sim 3770\,\mathrm{K}$) to produce two bins with an equal number of stars, and repeat the M3 fitting procedure for 100 inferences in either bin. The resulting models' installation break posterior distributions are plotted in the left panel of Figure (11) alongside the full sample posteriors from Figure (8). Neither of the bins had unimodal $\log_{10} I_{break}$ distributions, and both bins shared peaks at $\log_{10} I_{break} \approx 0.6$ and 1.4, or $I_{break} \approx 4$ and $25\,I_{\oplus}$ respectively. The cooler bin's primary (stronger) peak occurs at the latter, while the warmer bin's primary peak occurs at the former.

This contradicts an expectation that, if the power law break occurs at the same orbital period for all stars, then the break should correspond to smaller installation values for less massive (and thus cooler/fainter) stars. Nonetheless, because both the warmer and cooler subsamples exhibit the same general bimodality, we found it unlikely that the effective temperature distribution of host stars is responsible for this feature in the $\log_{10} I_{break}$ posterior distribution. Lacking alternative physical explanations, we could not assign a root cause for the bimodality, and considering such

Posterior Distributions of Key Parameters for Full and Temperature-Split Samples
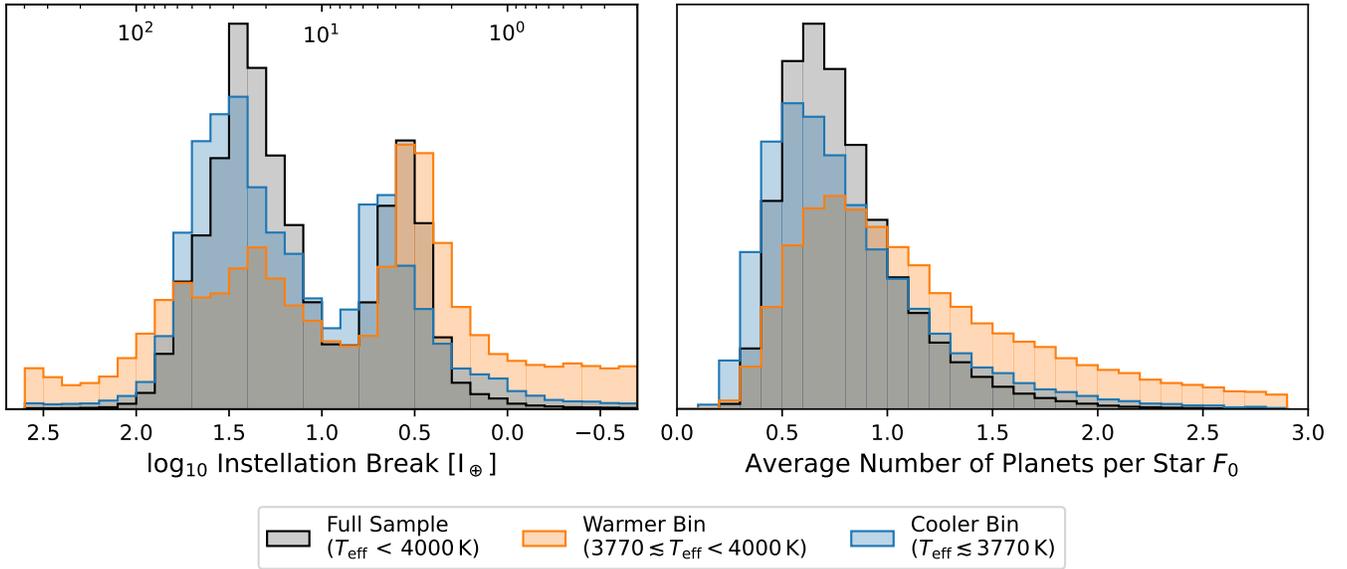
Figure 11. Posterior distributions for the **left:** $\log_{10}$ power law breaks in instellation and **right:** average number of planets per star. Grey histograms represent posteriors from the `M3` models for the full sample (same as Figure 8), while the orange and blue histograms show posteriors from models fit to the warmer and cooler subsamples split at $T_{\rm eff} \sim 3770$ K). In the left panel, tick marks on the upper edge represent the corresponding non-logarithmic values. Both temperature bins produce a bimodality in $\log_{10} I_{\rm break}$, suggesting this feature is not attributable to two host star populations distinct in temperature, though the location of the primary peaks is opposite what one would expect from scaling the same orbital period to hotter/cooler stars. In the right panel, we cannot resolve any significant difference in the average number of Earth-sized planets between warmer and cooler M dwarfs.

possibilities would (a) likely require a larger sample of planets more thoroughly sampling host star parameter space, and (b) be beyond the scope of this work. Given our present limitations, we attributed the instellation break bimodality to noise within our small sample which manifests as two possibilities for $\log_{10} I_{\rm break}$ under our functional form of choice.

C.1. *An Additional Outcome of the Temperature-Split Model*

In the right panel of Figure (11), we show the posterior distributions for the scaling parameter F_0 representing the average number of planets per star from models fit to the full, warmer, and cooler subsamples. We found a value of $F_0 = 1.05^{+0.92}_{-0.42}$ for the warmer subsample (median temperature of ~ 3894 K), and a value of $F_0 = 0.72^{+0.45}_{-0.26}$ for the cooler subsample (median temperature of ~ 3614 K), which are fully consistent at the 1σ level. We note that the accompanying uncertainties are large enough that the warmer/cooler ratio could fall anywhere from $\sim 0.5 - 4.3$ within 1σ.

As such, with these models we were not able to make any claims regarding the dependence of F_0 on sub-spectral type discussed in Section (6.2). This is most likely due to our very limited sample of relevant stars and planets. Focusing on only Earth-sized planets already limits the available sample of candidates to which one would fit a population model, even more so when splitting the sample up further based on host stellar effective temperature. While it may be true that *Kepler* lacks the candidates to resolve a finer stellar dependence for *Earth-sized* planets around M dwarfs, a less restrictive (though more involved) study considering all small planets ($< 4\,{\rm R}_\oplus$) may have enough candidates to probe such trends. However, the distribution of host star temperatures for those small planets is still top-heavy, making it difficult to explore mid and late M dwarfs with *Kepler*. As noted in Sections (6.2) and (6.3), additional surveys with improved sensitivity to such targets could provide valuable information contribution to future studies, especially those capable of integrating information from multiple surveys for greater coverage and resolution.

D. PLANETARY MASS DEPENDENCE OF HABITABLE ZONE BOUNDS

Here we describe our approach to adopting the planet-mass-dependent scaling of Kopparapu et al. (2014) to modulate the conservative inner (i.e., runaway greenhouse) edge with planet mass. We used the three example planetary masses $(0.1, 1.0$ and $5.0\,M_\oplus)$ in Kopparapu et al. (2014) to calculate the corresponding runaway greenhouse edges for each star in our sample. We then took the average bound from all stars for each example planet mass, and used these to define a one-dimensional interpolation to estimate the runaway greenhouse flux for a given planet mass. Because our models were defined in planet radius while the above is defined in planet mass, when integrating our models across the habitable zone in planet radius and instellation, we employed an empirical mass-radius relation (Chen & Kipping 2017) to translate each radius point to a corresponding mass, then used the above interpolation to compute the corresponding flux defining the inner edge of the conservative habitable zone at that radius. In Kopparapu et al. (2014) the outer conservative edge (maximum greenhouse) and the optimistic edges were not believed to change with planet mass, so these occur at constant fluxes with respect to planet size in our integrations.

REFERENCES

Akeson, R., Armus, L., Bachelet, E., et al. 2019, arXiv e-prints, arXiv:1902.05569, doi: 10.48550/arXiv.1902.05569

Astudillo-Defru, N., Delfosse, X., Bonfils, X., et al. 2017, A&A, 600, A13, doi: 10.1051/0004-6361/201527078

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

Berger, T. A., Huber, D., van Saders, J. L., et al. 2020b, AJ, 159, 280, doi: 10.3847/1538-3881/159/6/280

Bergsten, G. J., Pascucci, I., Mulders, G. D., Fernandes, R. B., & Koskinen, T. T. 2022, AJ, 164, 190, doi: 10.3847/1538-3881/ac8fea

Bixel, A., & Apai, D. 2021, AJ, 161, 228, doi: 10.3847/1538-3881/abe042

Borucki, W. J. 2016, Reports on Progress in Physics, 79, 036901, doi: 10.1088/0034-4885/79/3/036901

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977, doi: 10.1126/science.1185402

Brady, M. T., & Bean, J. L. 2022, AJ, 163, 255, doi: 10.3847/1538-3881/ac64a0

Bryson, S., Coughlin, J., Batalha, N. M., et al. 2020a, AJ, 159, 279, doi: 10.3847/1538-3881/ab8a30

Bryson, S., Coughlin, J. L., Kunimoto, M., & Mullally, S. E. 2020b, AJ, 160, 200, doi: 10.3847/1538-3881/abb316

Bryson, S., Kunimoto, M., Kopparapu, R. K., et al. 2021, AJ, 161, 36, doi: 10.3847/1538-3881/abc418

Bryson, S. T., Abdul-Masih, M., Batalha, N., et al. 2017, The Kepler Certified False Positive Table, Kepler Science Document KSCI-19093-003, id. 12. Edited by Michael R. Haas and Natalie M. Batalha

Burke, C. J., & Catanzarite, J. 2017, Planet Detection Metrics: Per-Target Detection Contours for Data Release 25, Kepler Science Document KSCI-19111-002, id. 19. Edited by Michael R. Haas and Natalie M. Batalha

Burke, C. J., Bryson, S. T., Mullally, F., et al. 2014, ApJS, 210, 19, doi: 10.1088/0067-0049/210/2/19

Burke, C. J., Christiansen, J. L., Mullally, F., et al. 2015, ApJ, 809, 8, doi: 10.1088/0004-637X/809/1/8

Chen, J., & Kipping, D. 2017, ApJ, 834, 17, doi: 10.3847/1538-4357/834/1/17

Coughlin, J. L. 2017, Planet Detection Metrics: Robovetter Completeness and Effectiveness for Data Release 25, Kepler Science Document KSCI-19114-002

Coughlin, J. L., Mullally, F., Thompson, S. E., et al. 2016, ApJS, 224, 12, doi: 10.3847/0067-0049/224/1/12

Dietrich, J., Apai, D., Schlecker, M., et al. 2023, AJ, 165, 149, doi: 10.3847/1538-3881/acba0b

Dong, S., & Zhu, Z. 2013, ApJ, 778, 53, doi: 10.1088/0004-637X/778/1/53

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

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

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1, doi: 10.1051/0004-6361/201833051

Gaudi, B. S., Seager, S., Mennesson, B., et al. 2020, arXiv e-prints, arXiv:2001.06683, doi: 10.48550/arXiv.2001.06683

Hardegree-Ullman, K. K., Apai, D., Bergsten, G. J., Pascucci, I., & López-Morales, M. 2023, AJ, 165, 267, doi: 10.3847/1538-3881/acd1ec

Hardegree-Ullman, K. K., Cushing, M. C., Muirhead, P. S., & Christiansen, J. L. 2019, AJ, 158, 75, doi: 10.3847/1538-3881/ab21d2

Hardegree-Ullman, K. K., Zink, J. K., Christiansen, J. L., et al. 2020, ApJS, 247, 28, doi: 10.3847/1538-4365/ab7230

Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, ApJS, 201, 15, doi: 10.1088/0067-0049/201/2/15

Howell, S. B., Sobeck, C., Haas, M., et al. 2014, PASP, 126, 398, doi: 10.1086/676406

Hsu, D. C., Ford, E. B., Ragozzine, D., & Ashby, K. 2019, AJ, 158, 109, doi: 10.3847/1538-3881/ab31ab

Hsu, D. C., Ford, E. B., & Terrien, R. 2020, MNRAS, 498, 2249, doi: 10.1093/mnras/staa2391

Huber, D., Bryson, S. T., Haas, M. R., et al. 2016, ApJS, 224, 2, doi: 10.3847/0067-0049/224/1/2

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90, doi: 10.1109/mcse.2007.55

Jones, E., Oliphant, T., Peterson, P., et al. 2001–, SciPy: Open source scientific tools for Python. http://www.scipy.org/

Kimura, T., & Ikoma, M. 2022, Nature Astronomy, 6, 1296, doi: 10.1038/s41550-022-01781-1

Kopparapu, R. K., Ramirez, R. M., SchottelKotte, J., et al. 2014, ApJL, 787, L29, doi: 10.1088/2041-8205/787/2/L29

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

Kunimoto, M., & Matthews, J. M. 2020, AJ, 159, 248, doi: 10.3847/1538-3881/ab88b0

Mamajek, E. E. 2022, A modern mean dwarf stellar color and effective temperature sequence. https://www.pas.rochester.edu/~emamajek/EEM_dwarf_UBVIJHK_colors_Teff.txt

Ment, K., & Charbonneau, D. 2023, AJ, 165, 265, doi: 10.3847/1538-3881/acd175

Morton, T. D., Bryson, S. T., Coughlin, J. L., et al. 2016, ApJ, 822, 86, doi: 10.3847/0004-637X/822/2/86

Mulders, G. D., Drążkowska, J., van der Marel, N., Ciesla, F. J., & Pascucci, I. 2021, ApJL, 920, L1, doi: 10.3847/2041-8213/ac2947

Mulders, G. D., Pascucci, I., & Apai, D. 2015a, ApJ, 814, 130, doi: 10.1088/0004-637X/814/2/130

—. 2015b, ApJ, 798, 112, doi: 10.1088/0004-637X/798/2/112

Mulders, G. D., Pascucci, I., Apai, D., & Ciesla, F. J. 2018, AJ, 156, 24, doi: 10.3847/1538-3881/aac5ea

Mullally, F., Coughlin, J. L., Thompson, S. E., et al. 2015, ApJS, 217, 31, doi: 10.1088/0067-0049/217/2/31

Pascucci, I., Mulders, G. D., & Lopez, E. 2019, ApJL, 883, L15, doi: 10.3847/2041-8213/ab3dac

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

Petigura, E. A., Rogers, J. G., Isaacson, H., et al. 2022, AJ, 163, 179, doi: 10.3847/1538-3881/ac51e3

Pinamonti, M., Sozzetti, A., Maldonado, J., et al. 2022, A&A, 664, A65, doi: 10.1051/0004-6361/202142828

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

Rauer, H., Catala, C., Aerts, C., et al. 2014, Experimental Astronomy, 38, 249, doi: 10.1007/s10686-014-9383-4

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003, doi: 10.1117/1.JATIS.1.1.014003

Robertson, P., Endl, M., Cochran, W. D., & Dodson-Robinson, S. E. 2013, ApJ, 764, 3, doi: 10.1088/0004-637X/764/1/3

Rogers, L. A. 2015, ApJ, 801, 41, doi: 10.1088/0004-637X/801/1/41

Sagear, S. A., Skinner, J. N., & Muirhead, P. S. 2020, AJ, 160, 19, doi: 10.3847/1538-3881/ab9224

Sestovic, M., & Demory, B.-O. 2020, A&A, 641, A170, doi: 10.1051/0004-6361/202037732

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163, doi: 10.1086/498708

Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints, arXiv:1503.03757, doi: 10.48550/arXiv.1503.03757

The LUVOIR Team. 2019, arXiv e-prints, arXiv:1912.06219, doi: 10.48550/arXiv.1912.06219

Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, ApJS, 235, 38, doi: 10.3847/1538-4365/aab4f9

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science & Engineering, 13, 22, doi: 10.1109/mcse.2011.37

West, A. A., Weisenburger, K. L., Irwin, J., et al. 2015, ApJ, 812, 3, doi: 10.1088/0004-637X/812/1/3

Windsor, J. D., Robinson, T. D., Kopparapu, R. k., et al. 2023, PSJ, 4, 94, doi: 10.3847/PSJ/acbf2d

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

Zahnle, K. J., & Catling, D. C. 2017, ApJ, 843, 122, doi: 10.3847/1538-4357/aa7846

Zink, J. K., Hardegree-Ullman, K. K., Christiansen, J. L., et al. 2021, AJ, 162, 259, doi: 10.3847/1538-3881/ac2309

—. 2023, AJ, 165, 262, doi: 10.3847/1538-3881/acd24c