The Astronomical Journal, 166:122 (19pp), 2023 September https://doi.org/10.3847/1538-3881/acebc8
© 2023. The Author(s). Published by the American Astronomical Society.
OPEN ACCESS

A Unified Treatment of Kepler Occurrence to Trace Planet Evolution. I. Methodology

Anne Dattilo[1], Natalie M. Batalha[1], and Steve Bryson[2]
[1] Department of Astronomy and Astrophysics, University of California Santa Cruz, Santa Cruz, CA 95064, USA; adattilo@ucsc.edu
[2] NASA Ames Research Center, Moffett Field, CA 94035, USA
Received 2023 April 6; revised 2023 July 22; accepted 2023 July 27; published 2023 August 22

Abstract

We present Kepler exoplanet occurrence rates for planets between 0.5 and 16 $R_\oplus$ and between 1 and 400 days. To measure occurrence, we use a nonparametric method via a kernel density estimator and use bootstrap random sampling for uncertainty estimation. We use a full characterization of completeness and reliability measurements from the Kepler Data Release 25 catalog, including detection efficiency, vetting completeness, astrophysical reliability, and false alarm reliability. We also include more accurate and homogeneous stellar radii from Gaia Data Release 2. In order to see the impact of these final Kepler properties, we revisit benchmark exoplanet occurrence rate measurements from the literature. We compare our measurements with previous studies to both validate our method and observe the dependence of these benchmarks on updated stellar and planet properties. For FGK stars, between 0.5 and 16 $R_\oplus$ and between 1 and 400 days, we find an occurrence of 1.52 ± 0.08 planets per star. We investigate the dependence of occurrence as a function of radius, orbital period, and stellar type and compare with previous studies with excellent agreement. We measure the minimum of the radius valley to be $1.78^{+0.14}_{-0.16}$ $R_\oplus$ for FGK stars and find it to move to smaller radii for cooler stars. We also present new measurements of the slope of the occurrence cliff at 3–4 $R_\oplus$, and find that the cliff becomes less steep at long orbital period. Our methodology will enable us to constrain theoretical models of planet formation and evolution in the future.

Unified Astronomy Thesaurus concepts: Exoplanets (498); Exoplanet formation (492); Exoplanet astronomy (486); Exoplanet evolution (491); Transit photometry (1709)

1. Introduction

Kepler's foremost legacy has been enabling detailed exoplanet demographic studies. Launched in 2009, Kepler was designed as a demographics mission to study the population of planets in our galaxy orbiting within 1 au of their host stars (Borucki et al. 2010). It also aimed to measure the frequency of Earth-like planets within their star's habitable zone. These small planets are difficult to detect around Sun-like stars due to both their small size and long orbital period. To detect these planets, Kepler observed a single field continuously for almost 4 yr.

In order to enable studies of planetary demographics, the Kepler Mission built a homogeneous planet catalog. A considerable amount of resources went into characterizing this catalog: both in the completeness (the fraction of planets that were correctly identified and vetted as planet candidates, hereafter PCs) and in the reliability (the fraction of PCs that are truly planets). The final uniform planet catalog, Data Release 25 (hereafter DR25), was released in 2018 and contained 4043 PCs (Thompson et al. 2018).

Over the past decade, a large amount of work has gone into measuring Kepler occurrence rates. Many early works (e.g., Borucki et al. 2011; Catanzarite & Shao 2011; Youdin 2011; Howard et al. 2012; Traub 2012) aimed to measure the frequency of small planets with the first two Kepler planet catalog releases, which contained a cumulative 1235 planets found in the first 13 months of data (Borucki et al. 2010, 2011). In the following years, multiple intermediate planet catalogs were created, and many occurrence rate studies were performed on them (e.g., Mann et al. 2012; Wright et al. 2012; Beaugé & Nesvorný 2013; Dong & Zhu 2013; Dressing & Charbonneau 2013, 2015; Fressin et al. 2013; Kopparapu 2013; Petigura et al. 2013; Swift et al. 2013; Foreman-Mackey et al. 2014; Kane et al. 2014; Morton & Swift 2014; Burke et al. 2015; Mulders et al. 2015a, 2016; Silburt et al. 2015; Bryan et al. 2016; Gaidos et al. 2016; Narang et al. 2018; Zhu et al. 2018; Hardegree-Ullman et al. 2019, among others).

These studies were improved with an accurate model of the Kepler completeness function. Understanding the Kepler completeness function via injection and recovery tests (Christiansen et al. 2013, 2015, 2016; Burke & Catanzarite 2017; Christiansen 2017) proved successful in improving our understanding of the intrinsic planet population. Burke et al. (2015) first used direct measurements of the Kepler detection efficiency to present occurrence rate calculations on the Q1–Q16 Kepler planet candidate sample from Mullally et al. (2016). These improvements revealed new features in the period–radius plane: Fulton et al. (2017) uncovered the radius valley, long theorized but not previously observationally seen. This feature is robust against various stellar samples. It can be seen in K2, second-mission Kepler, data (Hardegree-Ullman et al. 2019), implying its formation cause is not reliant on galactic position. Further, Van Eylen et al. (2018) showed that with asteroseismology the valley emptied, indicating that the more precise our stellar parameters, the more features we can pull out of the true planet population.

However, few occurrence studies used the DR25 catalog with DR25 completeness measurements, and with few exceptions, these measurements also did not incorporate any reliability measurements for individual planets. Hsu et al. (2019, hereafter H19) calculated Kepler occurrence rates with DR25

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

and updated Gaia-derived stellar properties, however without any reliability measurements. Kunimoto & Matthews (2020, hereafter KM20) used catalog reliability in their measurements but used their own pipeline to detect planets from the Q1–Q17 data (Kunimoto et al. 2020). Bryson et al. (2020a) calculated Kepler occurrence rates using a joint DR25-Gaia DR2 stellar catalog as well as DR25 reliability measurements, albeit for a focused domain in the period–radius domain.

Precise and uniform planet catalogs allowed us to measure new structures in the period–radius diagram, just as they did to reveal the radius valley in Fulton et al. (2017). The period–radius diagram is rife with features, such as the Neptune desert at low orbital period, the aforementioned radius valley bifurcating the small planet population, and the radius cliff that marks the steep drop in occurrence from the sub-Neptunes to the Neptunes. Many of these structures in the period–radius diagram have been tied to planet formation and evolution theory (e.g., Ginzburg et al. 2016; Mazeh et al. 2016; Owen & Wu 2017; Owen & Lai 2018; Kite et al. 2019). Billions of years of planet formation have driven our planets to their current places on the period–radius diagram. We can trace the history of planet formation through these end locations. By measuring planet occurrence, we can map the structure of the period–radius diagram to the physics that put each planet there. Therefore, the better we can describe the underlying structure of occurrence in the period–radius diagram, the better we can explain how planets form and evolve.

The purpose of this paper is to revisit the planet occurrence rate measurements from the past decade, now with a uniform treatment of all DR25 inputs, in order to well-measure the structure in occurrence space. We develop a methodology to calculate planet occurrence using a kernel density estimator (KDE). Our methodology incorporates DR25 completeness and reliability measurements in a nonparametric form in order to measure planet occurrence rates. Because it is of particular interest to measure structure in the period–radius diagram, the future goal of using this KDE methodology is to glean complex structure in multiple dimensions from the data, instead of fitting parametric forms to it. To validate the method, we have identified several benchmark demographic measurements from previous Kepler studies: the hot Jupiter occurrence, the radius distribution for small planets, and a parameterized period distribution.

Due to evolving planet catalogs and stellar samples, there has been a spread in the measurements for each of these benchmarks over the years. We measure these with our methodology and compare to previous studies, but we are unable to make truly direct comparisons between our measurements and those of previous studies because of different input assumptions. We combine the final DR25 parameters with Berger et al. (2020b) stellar parameters, so we can see how adding in the updated stellar values and reliability measurements impact the final results.

The paper is structured as follows. In Section 2, we discuss the stellar and planet samples we use from the Kepler survey. We then discuss how we characterize the survey bias with completeness and reliability measurements, as well as our methodology for measuring planet occurrence in Section 3. In Section 4, we compare to other works and measure features on the period–radius diagram. In Section 5, we tie these rates to physical theory. We conclude and lay out possible future work in Section 6.

2. Stellar Sample

We base our parent stellar sample on the Q1–Q17 DR25 stellar catalog, containing 198,709 stars (Mathur et al. 2017). In order to have the most precise and uniform catalog of stellar properties, we use Berger et al. (2020b) stellar properties. Berger et al. (2020b) derives a homogeneous catalog of stellar properties from isochrone fitting using Gaia parallaxes, broadband photometry, and spectroscopic metallicities where available. The Berger et al. (2020b) stellar properties cut down on the dominant sources of systematic errors in stellar radii by deriving them from isochrone-fit temperatures and luminosities, resulting in the average stellar radii error of less than 10%. Because all stellar parameters used in our model are from this catalog, we restrict the stellar sample in the DR25 catalog to those included in Berger et al. (2020b), cutting the stellar sample down to 177,661 stars.

We then make further cuts to our stellar catalog following Bryson et al. (2020a). First, we make cuts for binarity. The stellar multiplicity can bias the measured planet properties and suppress the planet occurrence for close-in binaries (Sullivan et al. 2022). The Gaia renormalized unit weight error (RUWE) is often used as a stellar multiplicity indicator. It is a goodness-of-fit metric to the astrometric fits based on Lindegren et al. (2018), who suggest a single-star cutoff at RUWE > 1.4. We apply a slightly harsher cut of 1.2 because there are few single stars below this limit (see, e.g., Bryson et al. 2020a; Berger et al. 2020b, for more discussion). After removing targets with RUWE > 1.2, there are 163,254 stars remaining. Gaia DR3 contains additional flags for binarity, specifically the `non_single_star` flag. We did not consider this flag prior to our analysis. The RUWE cutoff, however, is very effective, removing all but one of the 2319 binaries identified by the `non_single_star` multiplicity flag. The remaining flagged star has a RUWE value below the "typical" cutoff of 1.2, and therefore would not have been cut with a harsher RUWE metric. This single exception does not justify the cost of reanalysis, and it remains in our stellar sample. We also remove stars that are flagged in Berger et al. (2018) as likely binary (BIN flag = 1 or 3). Stars were flagged based on their inflated radii in the H-R diagram (161,075 remain). We leave stars flagged with BIN = 2 because the flag comes from high-resolution imaging, which was only performed for a subset of stars. We remove stars with Berger et al. (2020b) goodness-of-fit parameter <0.99, which measures the quality of isochrone fitting (160,638 remain). Our final stellar classification cut is to remove stars flagged as evolved stars according to Berger et al. (2018). This flag is determined using solar-metallicity `PARSEC` evolutionary tracks (Bressan et al. 2012) to define the terminal age main-sequence stars in the temperature-radius plane; anything above this boundary is classified as a subgiant or red giant star (and cut from our sample; 107,472 remain).

We then cull our sample of poor targets. We remove noisy targets identified in the KeplerPORTs package[3] (105,994 remain) and stars with NaN limb darkening coefficients (105,538 remain). We then cull based on duty cycle, the fraction of data cadences with valid data: we check for stars with duty cycle = NaN (105,538 remain), targets with duty cycle drop >0.3 (101,544 remain), and targets with duty cycles <0.6 (98,150 remain). We remove the stars with data span <1000 days, which ensures that each star in the sample has

[3] https://github.com/nasa/KeplerPORTs

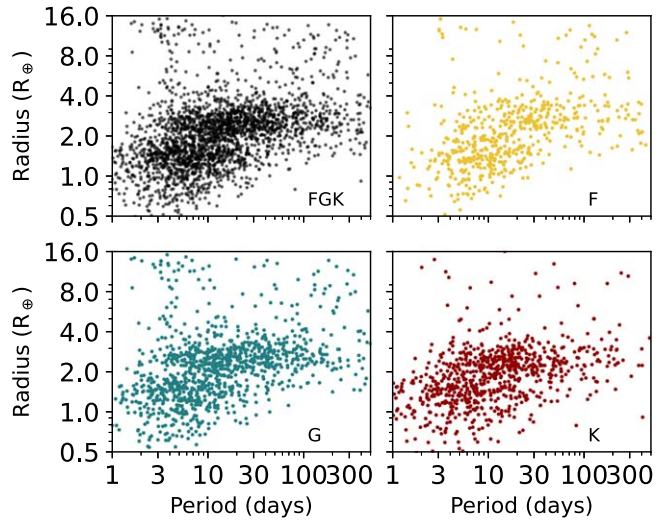

Figure 1. Period–radius diagram of the 2575 planet candidates in our planet sample, divided by stellar type.

enough data to meet the 3-transit requirement for orbital periods of 500 days (89,868 remain). The cut removes some stellar targets with short-period planets, but we require a uniform treatment for all stellar targets. Lastly, we remove the stars with the `timeoutsumry` flag $!=1$, which removes incompletely searched targets, resulting in 84,341 stars.

We split our sample by stellar type using the temperature boundaries from Pecaut & Mamajek (2013). The remaining 84,341 stars are thus split into 21,999 F stars (6000 K $\leqslant T_{\rm eff} <$ 7300 K); 41,501 G stars (5300 K $\leqslant T_{\rm eff} <$ 6000 K); 19,412 K stars (3900 K $\leqslant T_{\rm eff} <$ 5300 K). We use the sample of 82,912 FGK stars for our parent sample in the analysis.

2.1. Planet Sample

We derive our planet sample from the Q1–Q17 DR25 Kepler Object of Interest (KOI) table, restricted to KOIs with `koi_disposition` = CANDIDATE on the stars in our parent stellar sample. We recompute the planet radius, R_P, based on the ratio of planet radius to stellar radius, R_P/R_* (`koi_ror`), using the updated stellar radii from Berger et al. (2020b). Because there is an excess false-positive rate below 1 day in both the TCE population and the PC sample (see Figures 1 and 7 of Thompson et al. 2018), we cut our planet sample at 1 day.

After culling the stellar sample and cutting on planet orbital period and radius, we are left with 2575 PCs in our total FGK sample. There are 610 PCs around F stars, 1275 PCs around G stars, and 690 PCs around K stars. Figure 1 shows the period–radius distribution of the planet sample.

3. Methodology

With an idealized (100% complete and reliable) data set, our definition of planet occurrence for a defined stellar sample would be

$$\text{NPPS} = \frac{N_P}{N_*} \tag{1}$$

where the number of planets per star (NPPS) is the number of detected planets, N_P, divided by the total number of stars N_*. In practice, however, not all planets are detected by our survey, nor are all the planets that our pipeline detects truly planets. We can attempt to account for these biases by computing various performance metrics for the survey: we must account for both the *completeness* of our survey and the *reliability* of our sample.

The Kepler pipeline identifies potential transit signals in lightcurves as threshold crossing events (TCEs; Jenkins 2002). The Robovetter, an automated vetting algorithm that dispositions signals based on decision trees (Thompson et al. 2015), classifies these events as PCs or not. These signals are not always correctly categorized, however. Some planets are missed, and other dispositioned PCs are not actual planets. Completeness (or sensitivity, or recall) asks the following: out of all real planets, how many do we correctly disposition as PCs? Reliability (or precision) asks, out of all TCEs dispositioned as PCs, how many are true positives?

Using measured completeness and reliability, we update our occurrence equation:

$$\text{NPPS} = \frac{1}{N_*}\sum_i w_i \tag{2}$$

where we sum over all detections, using w_i as our weighting factor to account for the bias in our survey. The weighting factor w_i is defined as follows:

$$w_i = \frac{r_i}{\omega_i}; \tag{3}$$

ω is total summed completeness, which can be calculated at each (p_i, r_i) for each planet candidate as defined in Section 3.1, and r is reliability, calculated for each planet as defined in Section 3.2. The method expressed in Equation 3 to calculate occurrence is generally called the "Inverse Detection Efficiency Method" or IDEM (Howard et al. 2012; Petigura et al. 2013; Foreman-Mackey et al. 2014). We build upon this method using a KDE (Section 3.3).

3.1. Completeness

Our completeness model is based on the Bryson et al. (2020a) model and is comprised of two parts: detection and vetting completeness. Detection completeness is the fraction of true transiting planets that the Kepler pipeline could detect. Vetting completeness is the fraction of true transiting planets that are classified as PCs by the vetting pipeline. We expect both to be a function of orbital period and the signal-to-noise ratio of the transit lightcurve, which the Kepler pipeline measures as the multiple event statistic (MES; Jenkins 2002).

First, we calculate detection efficiency via a modified version of the KeplerPORTs code base. Burke & Catanzarite (2017) used flux-level injection tests to generate detection efficiency curves for individual target stars, which are a function of expected MES. KeplerPORTs calculates an expected MES value for a hypothetical planet at any given period and radius within the given parameter space, resulting in a per-star completeness contour $\omega_s(p, r)$ as a function of period and radius. Geometric transit probability, the probability the planet can be seen transiting its host star, is also computed at this step.

We then compute vetting completeness. Vetting completeness is the fraction of detected TCEs that are correctly dispositioned as PCs. We use the same process to calculate vetting completeness as that from Bryson et al. (2020a), where it is described in detail, and briefly summarize it here.

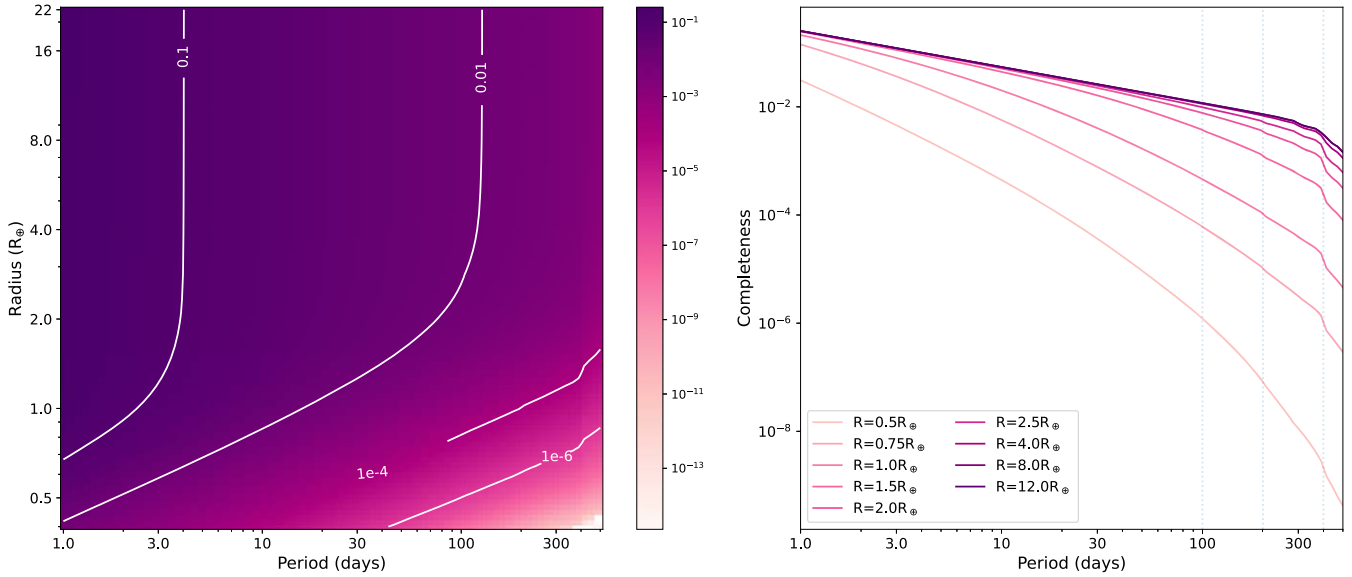

Figure 2. Left: FGK completeness function. Completeness degrades quickly along orbital period. White contours mark the 10-, 1-, 1e-2-, and 1e-4% completeness levels. Right: the completeness contour for specific radii. There is a discontinuity in occurrence at 400 days caused by discontinuities in the detection efficiency model.

Vetting completeness is computed by analyzing the fraction of recovered PCs in the injected set of planets from Christiansen (2017) in period-expected MES space. We grid period-expected MES space into cells and treat the fraction of correctly vetted PCs in each cell as a binomial rate. We fit these rates to a surface in period-expected MES space. As in Bryson et al. (2020a), this rate surface is the product of a nonrotated simplified logistic function in period and a rotated logistic in period and expected MES, fit via Markov Chain Monte Carlo (MCMC) inference using *emcee* (Foreman-Mackey et al. 2013). Since we use a different stellar sample than that from Bryson et al. (2020a; FGK versus GK), we refit the parameters. We do this for each stellar type (F, G, and K separately) but find no substantial difference between stellar types. We do find differences between our FGK fit and the GK fit from Bryson et al. (2020a), which is expected for the differences in stellar samples. See Appendix A for the specific coefficients.

Each component of the completeness function (detection efficiency, geometric transit probability, and vetting completeness) is treated as independent; for each star, they are multiplied together to calculate total completeness. We sum these functions over all stars to get an average completeness over the total stellar sample.

We do not account for how transit multiplicity impacts our detection efficiency. When searching for PCs, the highest signal-to-noise ratio signal is removed from the lightcurve by the Kepler pipeline, leaving less and less data for each subsequent search. The detection efficiency for higher-order planets decreases as a result and is a function of both planet radius and orbital period (Zink et al. 2019). Properly accounting for this effect is out of the scope of this paper. However, we would expect it to decrease our completeness contours (we are not accounting for missing planets), which would increase our occurrence rates. When comparing parametric fits to measured occurrence, the GK occurrence rates computed by Zink et al. (2019) agree with those of Burke et al. (2015), so we would expect the resulting difference in our measurements to be small.

The FGK completeness can be seen in Figure 2. It degrades as both a function of orbital period and planet radii and does not degrade smoothly. We would not expect it to be entirely smooth, though. For example, the window function is a key component to computing occurrence, which quantifies the probability that a target star meets the minimum of three transits in observational data. It is not smooth due to gaps in observations. There are also noticeable structures that appear as jumps at exactly 100, 200, and 400 days. This is a result of a correction factor for orbital period in the detection efficiency model (see Section 3.4 of Burke & Catanzarite 2017) computed in discrete bins. The discontinuities due to the window function average out over our entire stellar sample, but the orbital period discontinuities do not. Those remaining discontinuities can be seen in the right panel of Figure 2. Although small, they can add additional structure to our occurrence maps. For this study, we limit our occurrence measurements to 400 days.

3.2. Reliability

In addition to completeness, we must also characterize the potential for false positives contaminating our planet sample via reliability measurements. Using the methodology of Bryson et al. (2020a), we use a probabilistic approach to characterize whether a planet candidate is actually a false positive. We break these into two categories: (1) false positives, which are due to astrophysical signals that imitate transits, and (2) false alarms, which are due to instrumental systematics or stellar activity. The process to calculate the different parts of reliability is described in detail in Bryson et al. (2020a), and its effects are characterized in-depth in Bryson et al. (2020b).

PCs can be false positives due to astrophysical causes, such as grazing or eclipsing binaries. We use the Q1–Q17 DR25 False Positive Probabilities Table for our false-positive probabilities. These probabilities were created using the technique developed in Morton & Swift (2014) and were computed for all KOIs based largely on photometric data in Morton et al. (2016). We define the astrophysical reliability as (1 − the false-positive probability).

We characterize the false alarm reliability with an algorithm very similar to the vetting completeness. The inputs used to characterize are very different: instead of injected data, used to

find how incomplete the pipeline is, we use inverted and scrambled data (Coughlin et al. 2017). Inverted data inverts transit signals in lightcurves, while scrambled data mixes the lightcurve data order. Both of these create data that cannot contain true transits or astrophysical false positives, so any transit-like TCE is due to instrumental false alarms. As explained in Thompson et al. (2018), these inverted and scrambled false alarms have the same statistics as the dominant source of false alarms in the observed data.

The fraction of recovered signals in inverted and scrambled data is treated as a binomial rate problem in period-expected MES space, and MCMC inference is used to fit parameters to that surface. We fit the false alarm reliability with a rotated logistic function. As with vetting completeness, we fit each stellar type separately and find no substantial difference between the resulting posteriors. However, we do find a difference between the fit of Bryson et al. (2020a) and our FGK fit. See Appendix B for the specific coefficients.

The total reliability for each planet candidate is the product of the false-positive probability and the false alarm probability.

3.3. Computing Occurrence

The simplest way to measure occurrence is to count planets within some defined bin in period and radius (Equation (2)). However, this binning can leave "binning artifacts" in the measured occurrence (Foreman-Mackey et al. 2014). One can also fit power-law functions to the data to measure occurrence rates, but these fits limit the measured occurrence to that specific model, which may not be the correct choice.

We use a KDE to measure planet occurrence. KDEs allow us to measure closer to the underlying distribution by not assuming a model. We are not the first to employ KDEs in our occurrence framework. Other studies include Morton & Swift (2014), Mulders et al. (2015b), Dressing & Charbonneau (2015), Jin (2021), and Petigura et al. (2022).

We define f as the true planet occurrence and $\hat{f}$ as our measured occurrence. Traditionally, a KDE attempts to measure the underlying probability distribution of a sample using a function:

$$\hat{f}(x) = \frac{1}{N}\sum_{i=1}^{N} K(x - x_i; h) \tag{4}$$

where K is the kernel function, and h is the bandwidth. We use a 2D Gaussian kernel to represent the planet in log-uniform period–radius space, so $\hat{f}(x)$ becomes $\hat{f}(p, r)$. The bandwidth is the size of the kernel and acts as a smoothing function for the discrete data set. The bandwidth needs to be optimized so that the final $\hat{f}$ is not oversmoothed or undersmoothed. For this paper, we use an empirically optimized bandwidth of $(\sigma_p, \sigma_r) = (0.05, 0.05)$ (see Section 3.4.1). Our resulting function is as follows:

$$\hat{f}(p, r) = \frac{1}{N}\sum_{i=1}^{N} K(p - p_i; \sigma_p) K(r - r_i; \sigma_r); \tag{5}$$

where (p_i, r_i) represents each planet in our sample, and (σ_p, σ_r) denotes the bandwidth of (0.05, 0.05).

In essence, we place each planet in our sample down at its period and radius as a bivariate Gaussian with width (0.05, 0.05) and amplitude equal to its reliability value, normalized so that the area under its curve is also equal to its reliability. We calculate the variance of $\hat{f}$ by performing a bootstrap of our sample 5000 times. For each bootstrap iteration, we calculate a new KDE, $\hat{f}^*$, and average the collection for our final KDE model. The collection of $\hat{f}^*$ measures the variance about $\hat{f}$, which mimics the variance of $\hat{f}$ about f. Our final occurrence rate KDE is computed by dividing $\hat{f}$ (which is equivalent to the averaged $\hat{f}^*$) by the completeness contour described in Section 3.1.

Because the KDE is a nonparametric method, we are not limited to a set grid for our measurements. We compute the KDE over a 1000×1000 grid and then integrate to smaller bin sizes such that there are detections in each bin. This gives us the flexibility to be able to change our binning scheme to compare to other studies.

Throughout the paper, we refer to the average number of planets per star over a set domain as NPPS, which is computed as follows:

$$\text{NPPS} = F = \int_P \int_R \frac{\hat{f}(p, r)}{\omega}\, dr\, dp. \tag{6}$$

3.4. Method Validation

There are two questions we seek to address on the limitations of this methodology. Because the KDE kernel acts as a smoothing parameter, we need to understand what scale of structure we are able to measure. We also need to understand how we are limited by KDE edge effects and our survey completeness.

To address both of these, as well as to justify the choice of kernel size, we create a toy model. We generate a synthetic population with a known occurrence rate across our parameter space. Each (log-uniform) grid cell is sampled by a Poisson function, where the Poisson rate is the occurrence rate within that cell. This gives us the number of planets per cell. Each planet is then assigned a period and radius value from a uniform distribution bounded by the cell edges.

To create the *detected* synthetic population, we must determine if each planet is "observed." Whether a planet is observed or not can be thought of as a Bernoulli trial–it is either observed (a success) or not (a failure). We set an observational completeness across the entire parameter space. The completeness value at each planet's period and radius acts as the probability of success for each "trial." The planets successfully "observed" create the detected synthetic population, ready for our KDE pipeline. We use a reliability value of 1 for every planet. Once the KDE is created, it is divided by the completeness to get the measured occurrence. To validate the method, we compare input and measured occurrence. We input an occurrence of 1.0 planet per star within each grid cell. Between 1–400 days and 0.5–16 $R_\oplus$, we measure an average occurrence of 1.04 ± 0.06 . The occurrence outside this range, between 0.4–0.5 $R_\oplus$ and 16–22 $R_\oplus$, is systematically lower due to KDE edge effects (discussed in Section 3.4.2). We examine the level of structure we can measure in Section 3.4.1 and low-completeness boundaries and KDE edge effects in Section 3.4.2.

3.4.1. Testing Structure

To test the amount of structure we can truly measure, we inject occurrence patterns and see what details we can recover. We do this with four different tests with varying Gaussian occurrence width and separation, shown in the first row of Figure 3:

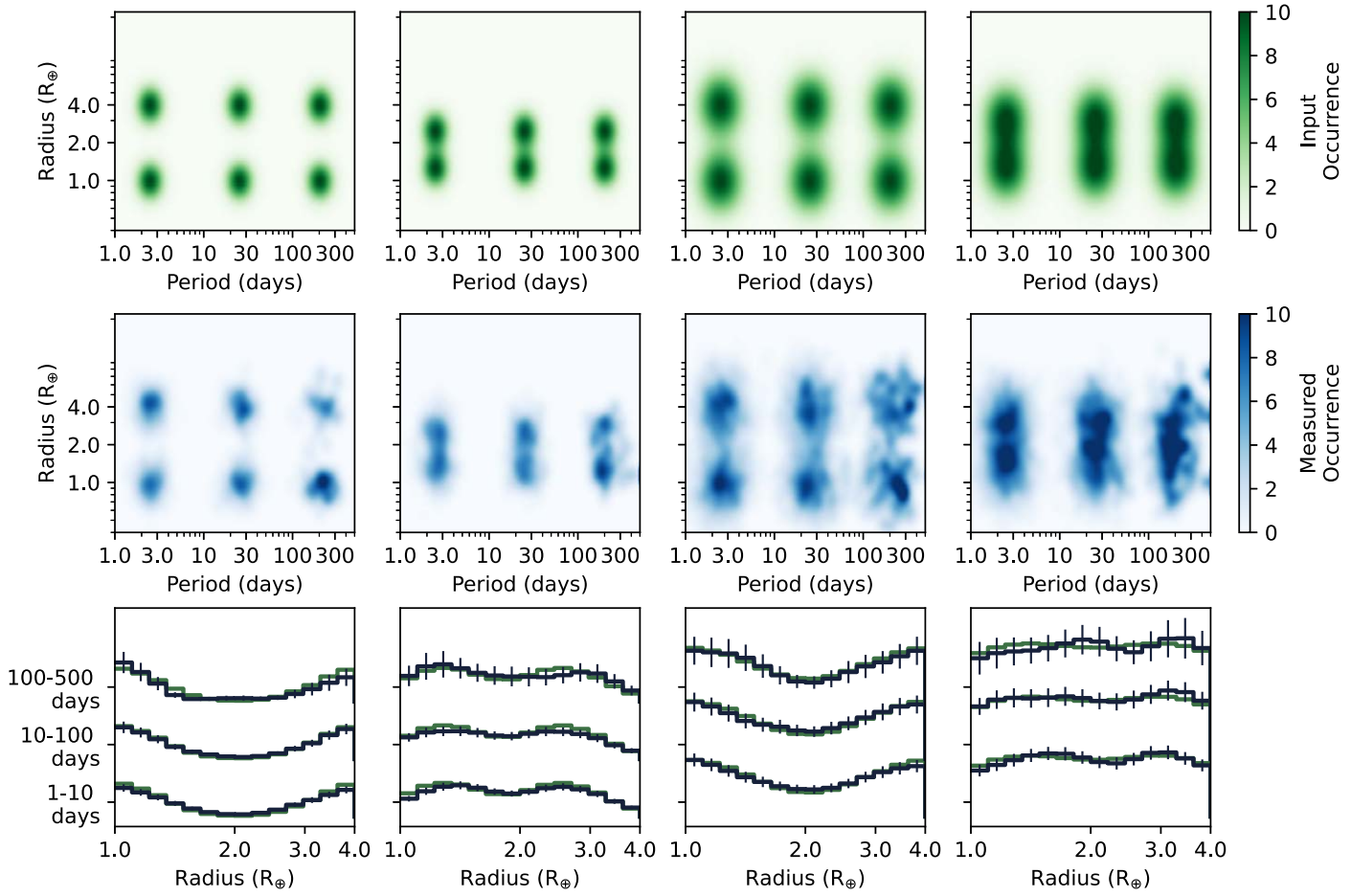

Figure 3. Testing what type of structure we can measure with a kernel size of $(\sigma_x, \sigma_y) = (0.05, 0.05)$. From top to bottom, the three parts to test our ability to measure structure: the input occurrence (top, blue); the measured occurrence (middle, green); and the marginalized radius distributions for each pair of Gaussian occurrence, with the input occurrence in green and the measured occurrence in blue, with 1σ uncertainties. We test four different scenarios as described in the text. For each scenario, we use a linear completeness function to degrade the input occurrence, with high completeness to the left (low period) and low completeness on the right (high period). We are able to recover the input occurrence within 1σ of our measurements.

1. "Small/far" has Gaussian occurrence width of (period width, radius width) = (0.1, 0.1) and separated in radius by log(radius) = 0.5 log $R_\oplus$.
2. "Small/near" has Gaussian occurrence width of (0.1, 0.1) and separated in radius by log(radius) = 0.25 log $R_\oplus$.
3. "Large/far" has Gaussian occurrence width of (0.17, 0.17) and separated in radius by log(radius) = 0.5 log $R_\oplus$.
4. "Large/near" has Gaussian occurrence width of (0.17, 0.17) and separated in radius by log(radius) = 0.25 log $R_\oplus$.

For each test, we use a proxy completeness that decreases linearly with increasing orbital period. We use three pairs of Gaussian occurrence for each test to investigate the effect of decreasing completeness and run the injected occurrence through the entire methodology. The second row of Figure 3 shows our measured occurrence. In each of the four tests, the low-completeness occurrence at long orbital period is noise-dominated.

We marginalize each Gaussian occurrence pair and look at their radius distribution, shown in the third row of Figure 3. We are able to recover the structure (maximum of the two Gaussian occurrence peaks and the minimum between them) for all but one case: the low-completeness pair for the large/near case. The measured occurrence is noise-dominated, and the measured peak is 0.51 ± 0.34 $R_\oplus$ from the input location. We are able to measure the structure on the expected scale of the radius valley for completeness values greater than 1%.

We empirically determine the kernel size by using the small/near injected kernels with three different kernel sizes, shown in Figure 4. The smallest kernel size $(\sigma_x, \sigma_y) = (0.05, 0.01)$ (first column) recovers the injected structure well, with each pair measuring the peak or minimum within 1σ of the true value. However, there is a large amount of fine structure in the marginalized occurrence because the σ_y is smaller than our bin size. The middle column, with $(\sigma_x, \sigma_y) = (0.05, 0.05)$, likewise recovers the input structure within 1σ of the input values for all levels of completeness. The right column, with $(\sigma_x, \sigma_y) = (0.1, 0.1)$, oversmooths the occurrence distribution and does not measure a valley at all.

For the smallest kernel size, the measured peaks are in agreement with the input occurrence, but the measured distributions are noisy. The largest kernel size oversmooths the distribution, resulting in a lack of any structures. We use a kernel size of $(\sigma_x, \sigma_y) = (0.05, 0.05)$ for the occurrence measurements in Section 4. The choice of kernel does not affect our total occurrence measurement; the total occurrence measured in each case is within 5% of the input occurrence. Jin (2021) also found this to be true; they used a KDE to measure relative planet occurrence and likewise found that kernel size

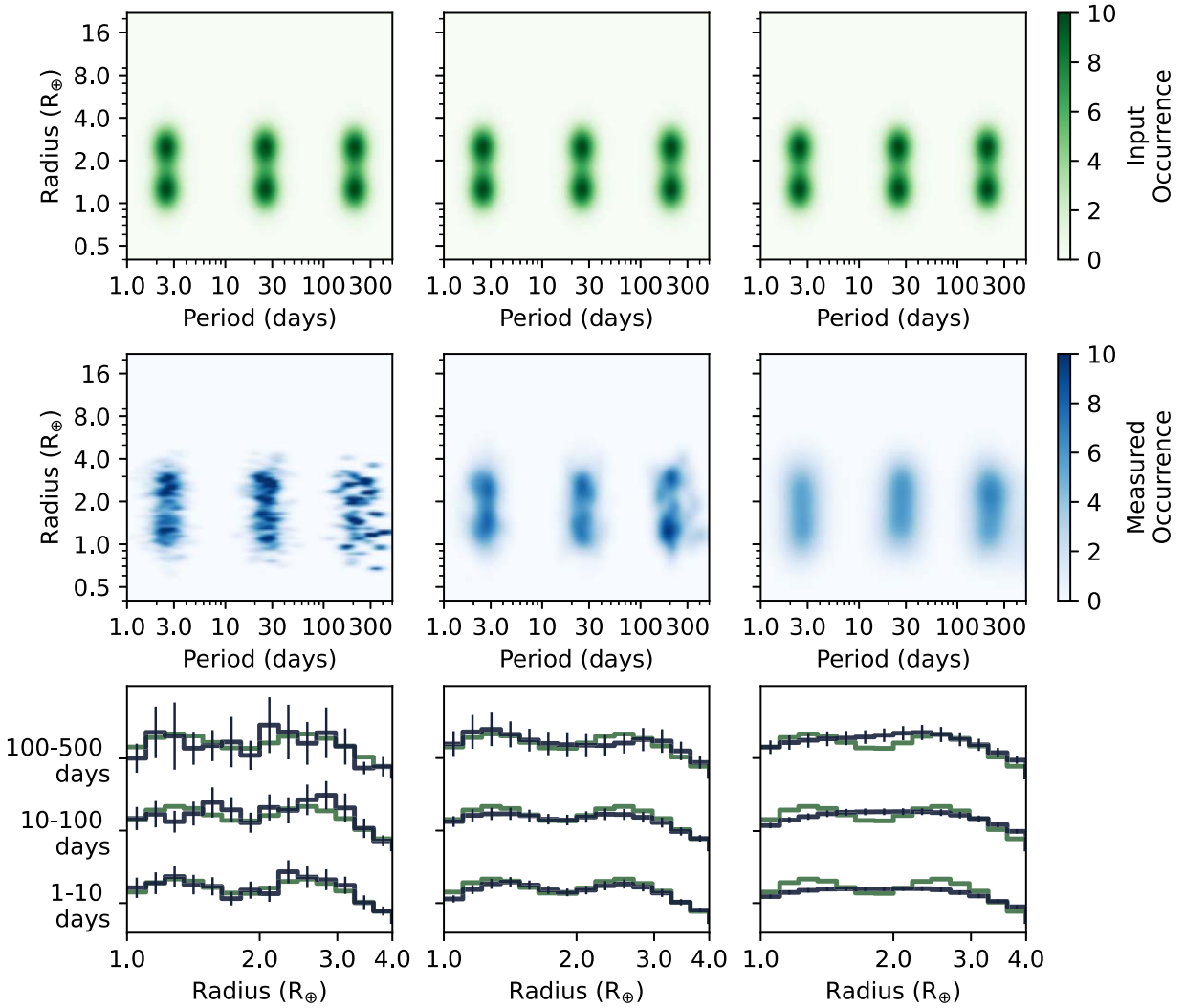

Figure 4. Comparison of different kernel sizes. Each column, from left to right: $(\sigma_x, \sigma_y) = (0.05, 0.01)$; (0.05, 0.05); (0.1, 0.1). The top row is the true input occurrence. The middle row is the occurrence returned from our methodology. The bottom row compares the marginalized occurrences with 1σ uncertainties for the measured occurrence. The 1–10 days marginalization for the largest kernel size has uncertainties smaller than the line width.

acted only as a smoothing factor and that overall occurrence was robust to choice of kernel size. The difference in smoothing, however, is significant, which can limit the structure we are able to measure.

3.4.2. Looking at Low Completeness and Edge Effects

There are also limits on *where* we are able to measure occurrence with a KDE. We are limited by KDE edge effects and our completeness function.

The KDE is computed for a specific parameter space. Within 1 kernel-width of the boundary, there is a systematic under-reporting of occurrence because the KDE acts as if there is no population on the other side of the boundary as opposed to it being undefined. Our completeness function also limits where we may measure occurrence: the Kepler completeness function steeply degrades at small radius and long orbital period, creating a region where we expect to detect less than one planet.

We can test both of these limits by applying a uniform occurrence across the entire parameter space and seeing what we can recover by assuming (1) a uniform, 100% completeness to test edge effects, or (2) the true Kepler completeness to see where the planet recovery per bin falls to zero.

We test this by uniformly populating our parameter space at 10 planets per star per grid cell and using a 100% completeness function (so there is no degradation of the occurrence). We use the largest tested kernel size of $h = (0.1, 0.1)$ as this will have the largest edge effect. The sample is then bootstrapped 5000 times and averaged for our final measured occurrence. The first column of Figure 5 shows the measured occurrence (center panel, in blue) and the residuals (bottom panel, in black) of this test.

Edge effects. The density of the KDE will always be lower at the edges; there is no contribution of neighboring Gaussians outside of the specified parameter space. How large of an effect this is depends on the width of the Gaussian kernel. The size of the underdensity region is 1 kernel-width.

We specifically calculate our KDE from 1–500 days and 0.4–22$R_\oplus$ and only calculate occurrence on the subdomain of 1–400 days and 0.5–16$R_\oplus$ in order to mitigate these edge effects. This completely removes any edge effects in radius. In the period, we only integrate to 400 days in part to avoid the edge effects; however, we are unfortunately limited in our approach at the lower-period boundary.

We require occurrence measurements down to 1 day to compare with previous studies (as is the goal of this paper). To

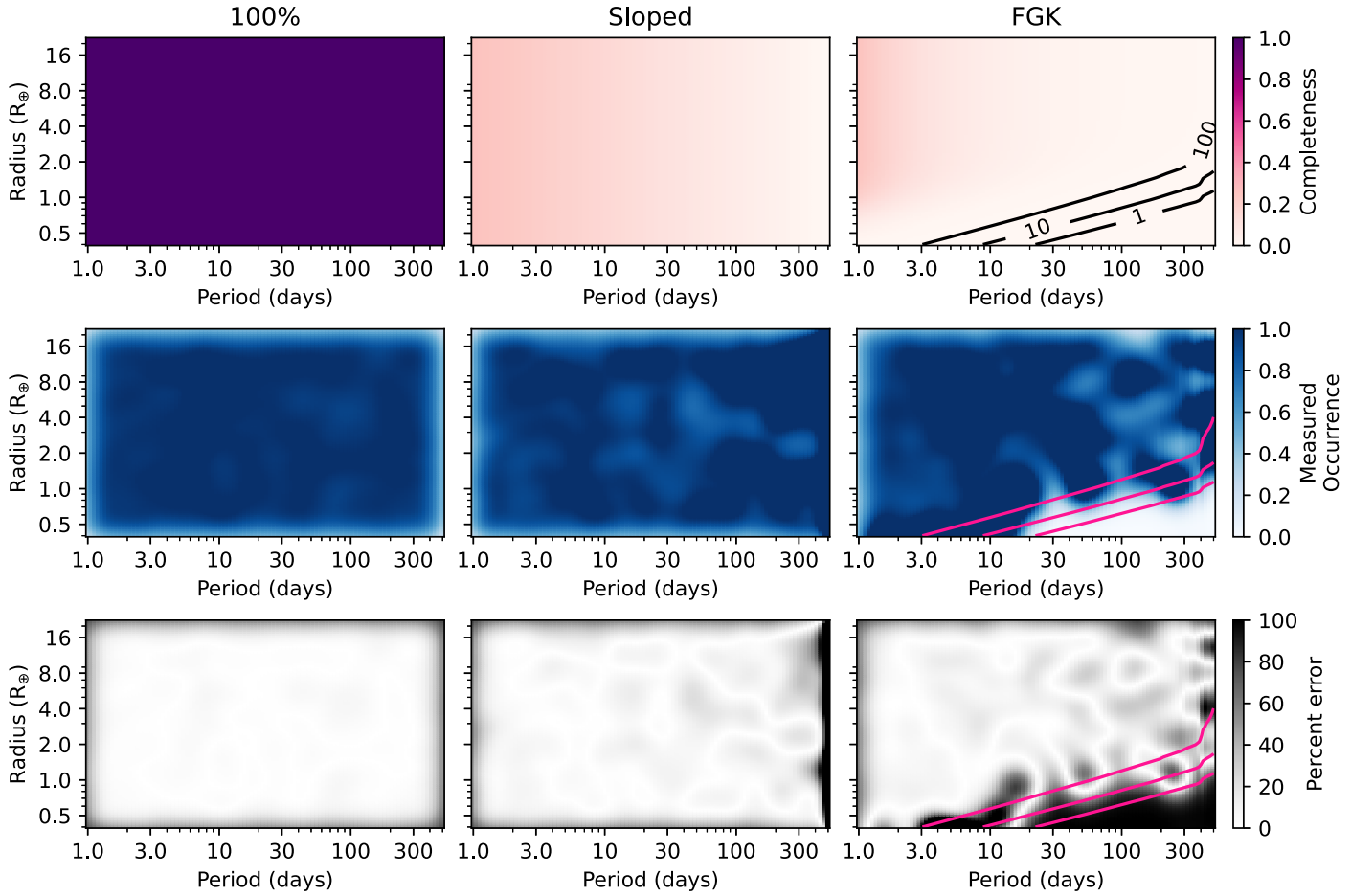

Figure 5. Toy model for three different completeness contours with uniform input occurrence. Left column: 100% completeness. Middle column: sloped completeness contours used in Figures 3, 4. Right column: Kepler FGK occurrence. The center row is the measured occurrence from our KDE methodology. The bottom row is the percent error. The leftmost column shows the impacts of edge effects on the resulting KDE. In the rightmost column, we use the middle contour line as our FGK low-completeness boundary.

avoid edge effects at the low-period boundary, we would need to extend our computation domain to 0.9 days for our completeness, reliability, and planet density KDE. We do not extend our KDE domain to below 1 day because we do not want to include the spurious detections below one day (discussed in Section 2.1). Including these planets would extend our domain, therefore avoiding KDE edge effects, but including those spurious PCs would impact our occurrence rate measurements more than the edge effects.

KDE edge effects are a limitation of the methodology in general. A different way to mitigate these effects would be to use "reflected boundaries," where it is assumed that the KDE is constant on the other side of the boundary (Silverman 1986). We do not believe this would be an appropriate assumption for low-period planets due to the small number of planet detections in this regime.

Further, the region where these edge effects are impactful is only from 1–1.12 days for our kernel size. These edge effects are the most extreme when there is a high KDE density right at the boundary. For low orbital period, between 1 and 1.12 days, the low density affects the occurrence measurement in Figure 5 by about 27%. 1–1.12 days is small relative to the bin size we use throughout the paper. How much this affects the reported occurrence values scales with the bin width; for the binning scheme presented in Figure 7, this amounts to ∼4% of the reported occurrence for the lowest-period bin, 1.0–2.11 days.

Low completeness. The low-completeness limit can also be investigated by setting a uniform occurrence and using the real FGK completeness function.

We measure this in the third column of Figure 5. The three contours represent the completeness level of $100/N_{\rm stars}$, $10/N_{\rm stars}$, and $1/N_{\rm stars}$ where $N_{\rm stars}$ is the total number of stars in our sample. A conservative threshold of low completeness can be set at the $100/N_{\rm stars}$ level; there is complete recovery above this contour. While there is not complete recovery at the $10/N_{\rm stars}$ level, there is still some recovery, and we set this as our low-completeness boundary. Below this contour, there is very little recovery, and none below the $1/N_{\rm stars}$ level. We only report upper limits on the occurrence below the $10/N_{\rm stars}$ level.

4. Occurrence

We have described a nonparametric technique for computing occurrence rates. The long-term goal of this methodology is to map features in multidimensional parameter space to theoretical models of planet formation and evolution. The observed planet KDE and its occurrence counterpart can be seen in Figure 6. There is a wealth of features to be seen in the occurrence map. There appears to be a limit in period to the small planet distribution, tracing out the lower boundary to the Neptune desert. The radius valley can clearly be seen just below 2 $R_\oplus$. Above the large population of sub-Neptunes,

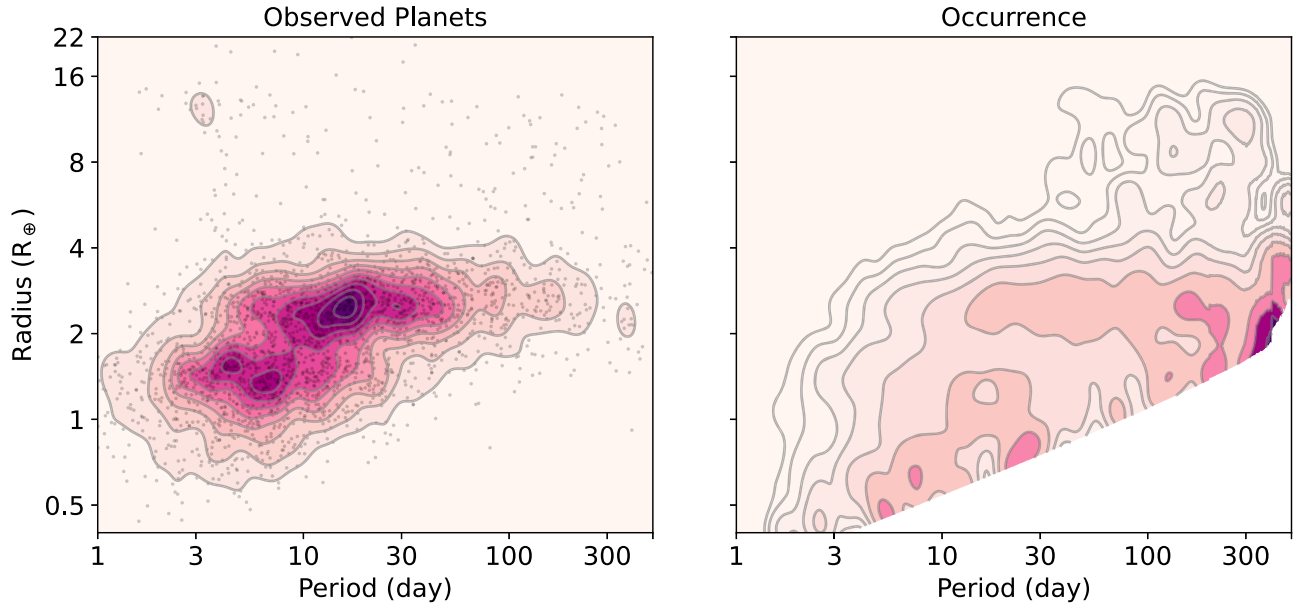

Figure 6. KDE maps of the observed FGK planet population (left) and the completeness-corrected intrinsic occurrence (right).

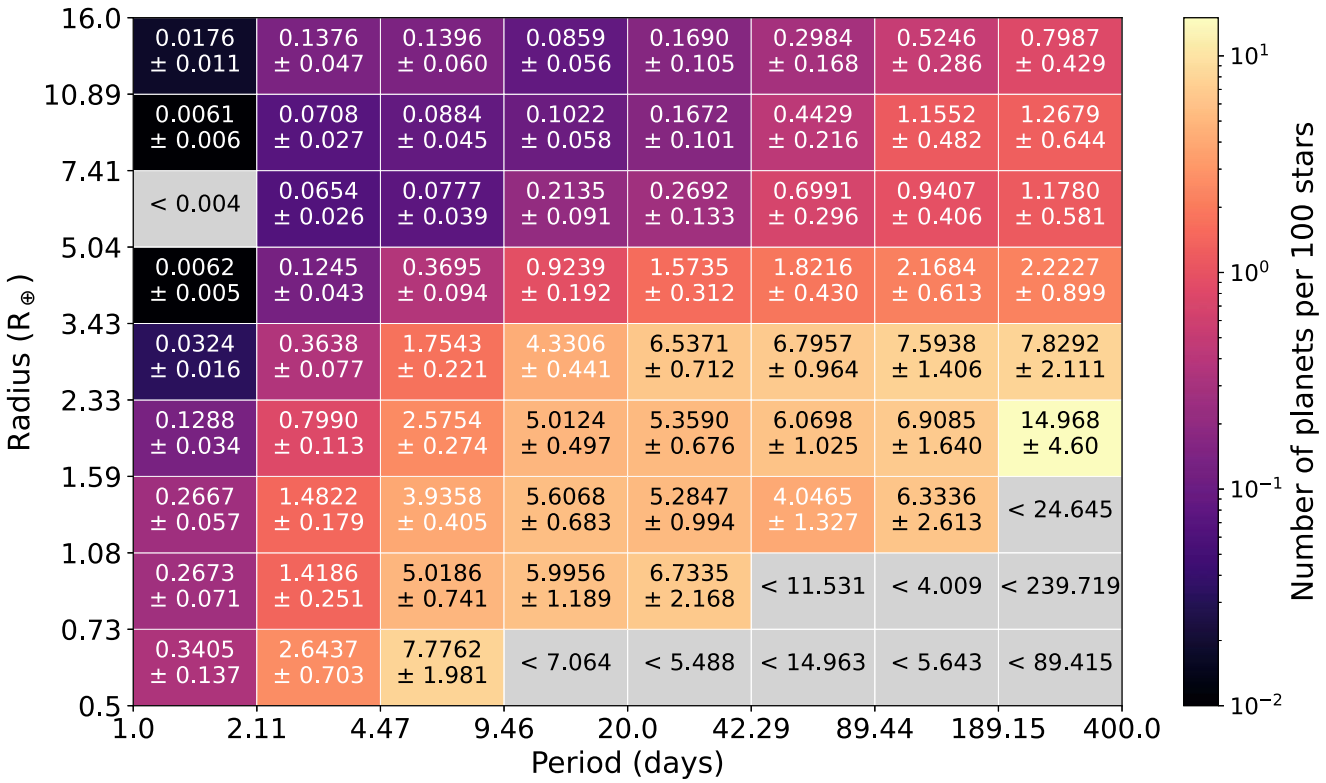

Figure 7. FGK occurrence rate estimates. The number of planets per 100 stars (shown) is comparable to a percentage. Gray cells are either those with no planet detections or below the low occurrence boundary (see Section 3.1), with only the upper limit shown.

there is a decline in occurrence. Further, the shape of this drop-off in occurrence is not constant across different orbital periods. Not all of these structures can be well-described with analytic functions, so a nonparametric approach is key to be able to describe these structures. In order to tie these structures in with theory, we must first validate our methodology by comparing benchmark measurements against literature values. To do this, we bin the parameter space, integrate occurrence over specific domains, and collapse the 2D map into 1D distributions for comparisons. In Section 4.1, we compute integrated rates; in Section 4.2, we look at the 1D radius distribution; in Section 4.3, we look at the 1D period distribution; in Section 4.4, we look at these as a function of stellar type. In Section 4.5, we explore the occurrence cliff.

4.1. Integrated Rates

Figure 7 shows our occurrence rate estimates and 1σ uncertainties for the entire period and radius range of 1–400 days and 0.5–16 $R_\oplus$. The gray bins represent either bins with

no planet detections in our chosen binning scheme or bins that fall below the completeness boundary established in Section 3.4; they contain upper limits for the occurrence. We use the log-uniform bins $P = [1.0, 2.11, 4.47, 9.46, 20.0, 42.29, 89.44, 189.15, 400.0]$ days in period, and $R_P = [0.5, 0.73, 1.08, 1.59, 2.33, 3.43, 5.04, 7.41, 10.89, 16.0]$ $R_\oplus$ in radius to minimize the number of empty bins in our sample.

Integrating over bins with detections above the low-completeness boundary (colored bins in Figure 7), we measure 1.52 ± 0.08 planets per star. Petigura et al. (2018, hereafter P18) report an average value of 1.10 NPPS (albeit for a smaller parameter space; from Figure 6 of their paper), KM20 report $1.06^{+0.09}_{-0.07}$ NPPS, and H19 measure a lower limit of $2.37^{+0.93}_{-0.67}$ NPPS (for periods out to 500 days; derived from Figure 2 of their paper) on total FGK occurrence.

While the differences between these measurements are significant, they are not comparable quantities. The integrated occurrence rate does not serve as a reliable benchmark because integrating over such a large domain encompasses regions of parameter space with zero detections and/or exceptionally low completeness. The degree to which this impacts the integrated occurrence depends not only on the input assumptions but also on the details of sample selection, which varies between studies. The benchmarks discussed in subsequent sections, however, are based on marginalized distributions that better facilitate comparisons.

A more reliable benchmark is the hot Jupiter domain, where completeness is high. The Kepler hot Jupiter occurrence rate has stayed almost constant over the past decade. Between 8–16$R_\oplus$ and over a period of 1–10 days, we measure an occurrence rate of $\eta = 0.46\% \pm 0.1\%$. Our measurement is in agreement with the literature values, which range from 0.4%–0.6% (Howard et al. 2012; Fressin et al. 2013; Mulders et al. 2015c; Petigura et al. P18; KM20). However, this rate is roughly half of the rate measured from radial velocity surveys, which measure between 0.8% and 1.2% (Mayor et al. 2011; Wright et al. 2012; Wittenmyer et al. 2020). The pile-up could be a result of metal-rich stars, while the Kepler sample has systematically lower-metallicity stars than radial velocity (RV) samples (Wright et al. 2012; Dawson & Murray-Clay 2013). Studies on the dependence of giant planet occurrence on host star metallicity have found a strong correlation between the two (P18). However, this is still debated as Guo et al. (2017) found that metallicity alone cannot explain the difference in rates. It could possibly be explained by differences in false-positive rates (Wang et al. 2015; Guo et al. 2017). The Kepler results have been shown to be within 1σ of the large uncertainty associated with the RV data (Fernandes et al. 2019). Recently, Beleznay & Kunimoto (2022) measured the hot Jupiter occurrence rate for AFG stars with TESS and found a rate of $0.39\% \pm 0.06\%$ for F and G stars, which are within 1σ of our result.

4.2. The Small Planet Radius Distribution

The radius distribution of small planets ($<4R_\oplus$) is rife with structure. The large population of observed planets can be split into the dense super-Earths and the puffier sub-Neptunes, separated by a region of low occurrence called the radius valley. The bimodal population was first observationally seen with the California-Kepler Survey (CKS; Petigura et al. 2017) in Fulton et al. (2017), following years of theoretical predictions (e.g., Lopez & Fortney 2013; Lee & Chiang 2016; Owen & Wu 2017; Gupta & Schlichting 2020; Rogers & Owen 2021). The exact location and slope of the radius valley have been the focus of many studies, as it could inform which physical process is the primary driver of planet formation (see Table 6 of Ho & Van Eylen 2023 for a summary of values).

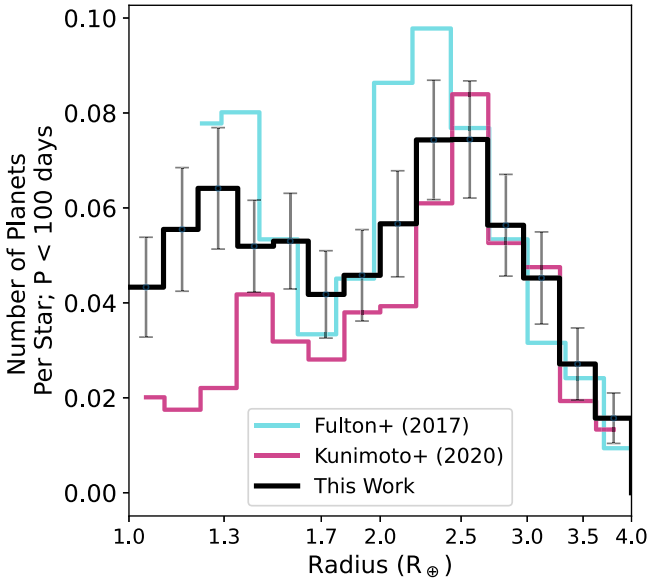

Figure 8. Radius distribution for small planet with $P < 100$ days. Blue is data taken from Fulton et al. (2017), and pink is data taken from KM20. We find the minimum of the radius valley to be $1.78^{+0.14}_{-0.16}$ $R_\oplus$.

We compare the radius distribution within 100 days with selected previous studies in Figure 8. Only considering occurrence within 100 days not only agrees with the limit to most atmospheric mass loss processes but allows us to integrate over a range where we are not hindered by low completeness (see Section 3.4 for bounds). Fulton et al. (2017; blue in Figure 8) measured this structure with the CKS sample, a precisely characterized subsample of the Kepler data set, including 36,075 stars and 900 planets. Fulton et al. (2017) used an independent planet detection pipeline and associated injection/recovery experiment to correct for completeness; we compare it here to benchmark the evolution of radius valley measurements. KM20 (pink in Figure 8) used their own independent pipeline to detect and vet planets and calculated their own reliability and completeness. They incorporate Gaia DR2-derived stellar radii to update planet radii. Their catalog of 96,280 FGK stars and 2829 planets is similar to ours but not the same. Overall, we find good agreement for all three distributions. We compare specific features of the distribution next.

There are three distinct structures we measure in the radius distribution: the minimum of the radius valley, the peak of the super-Earths (which we define as the portion of the distribution to the left of the valley), and the peak of the sub-Neptunes (which we define as the portion of the distribution to the right of the valley). We measure them as follows: for each iteration of our bootstrap simulation, we integrate the occurrence within 100 days binned to 50 log-uniform bins between 1 and 4 $R_\oplus$. Then, we find the minimum occurrence bin between 1.4 and 2.2 $R_\oplus$ and set that radius, $R_{\rm min}$, as the minimum of the valley. Finally, we rebin to the final binning scheme ($R_P = [1.0, 1.1, 1.21, 1.35, 1.49, 1.64, 1.81, 2.0, 2.21, 2.44, 2.69, 2.97, 3.28, 362, 4.0]$ $R_\oplus$); we use more bins than in the 2D occurrence

maps both to resolve the radius valley and to compare with KM20. We fit both planet populations with a Gaussian distribution, from [1.0, $R_{\rm min}$] to the left and [$R_{\rm min}$, 4.0] to the right with `scipy curve_fit`, giving us the location of the peaks of the sub-Neptunes and the super-Earths. For our FGK population, we find the peak of the sub-Neptunes to be at $2.51^{+0.06}_{-0.08}$ $R_\oplus$, the minimum of the radius valley to be at $1.78^{+0.14}_{-0.16}$ $R_\oplus$, and the super-Earth peak to be at $1.37^{+0.19}_{-0.44}$ $R_\oplus$. These measurements are within 1σ agreement to the peak and minimum measurements reported in previous studies. Fulton et al. (2017) finds the super-Earth peak to be 1.3 $R_\oplus$ and the sub-Neptune peak to be 2.4 $R_\oplus$ while KM20 finds the super-Earth peak to be 1.4 $R_\oplus$, the minimum of the valley to be 1.7 $R_\oplus$, and the sub-Neptune peak to be 2.6 $R_\oplus$. Martinez et al. (2019) derive precise, spectroscopically derived radii of 1633 planets and obtain radius distribution measurements of 1.47 ± 0.05 $R_\oplus$, 1.89 ± 0.07 $R_\oplus$, and $2.72 \pm 0.10 R_\oplus$, which are in agreement with our measurements within 1σ.

It is not obvious that the super-Earth occurrence below the radius valley turns over toward smaller radii (that there is a true super-Earth "peak," in other words). While Fulton et al. (2017) reported a bimodal distribution with two distinct peaks, KM20 found a flattening of the population below the valley, and H19 found an increasing occurrence of planets below the valley down to 0.5 $R_\oplus$. We fit the super-Earths ($P = [1, 100]$ days, and $R = [1, 1.78]$ $R_\oplus$) with both a Gaussian distribution and linear function to differentiate between these scenarios. To tell which model is better, we compute the Akaike Information Criterion (AIC; Akaike 1974). The (minimum, mean) AIC for the Gaussian fit is (−33.04, −26.22) while the (minimum, mean) AIC linear fit is (−30.48, −27.59). The difference in AIC minimums is less than 3, which constitutes no substantial evidence that either model is preferred, so we also compare their Bayesian Information Criterion (BIC; Schwarz 1978). The (minimum, mean) BIC for the Gaussian fit is (−36.85, −30.03) while the (minimum, mean) BIC for the linear fit is (−33.09, −30.27). $\Delta {\rm BIC}_{\rm min} = 3.76$, which means that the Gaussian fit is slightly favored but still does not constitute substantial evidence of a preferred model. We do not have precise enough occurrence measurements below the valley to say if there is a super-Earth peak or if occurrence flattens out below the valley.

While we find no substantive evidence that the population within 100 days turns over at 1 $R_\oplus$, we can look at the population below 1 $R_\oplus$ to see if there is a peak in occurrence for the terrestrials, which we define here as the planets with $R = [0.7–1.0]$ $R_\oplus$. We limit our orbital period range to 30 days for this analysis, in order to stay within our low-completeness boundary. Overall, occurrence is higher at small radius, with large error bars. We compute the same fits as before, a Gaussian distribution and a linear fit, for the smaller orbital period range ($P = [1, 30]$ days, and $R = [0.7, 1.75]$ $R_\oplus$). $\Delta AIC_{\rm min} = 1.85$, and $\Delta {\rm BIC}_{\rm min} = 0.66$ when comparing the Gaussian and linear fits, so we are unable to say which model is preferred. We find less evidence of a peak in occurrence for the terrestrials than for the super-Earths. Substantial uncertainties on occurrence (∼23% relative errors compared to <20% for the 100 day occurrence) are likely the driver of the uncertainty in the fits. In comparison, H19 measure occurrence down to 0.5 days and found increasing occurrence down to small radii. Even below 30 days, we do not measure the substantial population of terrestrials that H19 does;

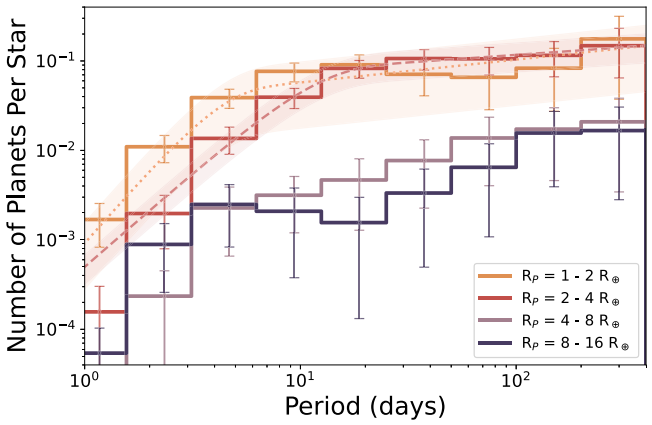

Figure 9. The period distribution of small planets, with fits to Equation 7 for the small planets. The terrestrial planets, 1–2$R_\oplus$, are plotted in orange, and the sub-Neptune planets, 2–4$R_\oplus$ are plotted in pink. Both small planet populations show a rise in occurrence until a break period, where occurrence flattens out. The larger planets, in bins 4–8$R_\oplus$ (purple) and 8–16$R_\oplus$ (navy), show no break point and rise in occurrence across all orbital periods.

our occurrence for the bin P = [1., 32.] days, and $R_P = [0.75, 1.]$ $R_\oplus$ is $F = 0.139 \pm 0.03$, where theirs is $F = 0.186 \pm 0.04$.

4.3. Period Distribution

We marginalize our FGK occurrence rates into four different radius bins to investigate the period distribution. The four radius bins are 1–2 $R_\oplus$ (which we refer to as super-Earths in this section), 2–4 $R_\oplus$ (which we refer to as sub-Neptunes in this section), 4–8 $R_\oplus$, and 8–16 $R_\oplus$. We use these specific radius bins to directly compare with previous studies, instead of using the radius valley minimum measured in Section 4.2 as the dividing radius between super-Earths and sub-Neptunes.

The period distribution can be seen in Figure 9.[4] Overall, the small planets, 1–2 $R_\oplus$ and 2–4 $R_\oplus$, show a sharp increase in occurrence until a break period, where occurrence flattens out. The large planets do not show evidence of a period break and slowly rise in occurrence with increasing orbital periods.

We fit the small plant distributions with a broken power law as in Howard et al. (2012):

$$\frac{df}{d\log P} = CP^{\beta}(1 - \exp^{(-P/P_0)^{\gamma}}). \quad (7)$$

We fit this distribution with the `scipy curve_fit` package, via the same method as the radius distribution fit. The broken power law simplifies into a piecewise function of the following form:

$$\frac{df}{d\log P} \propto \begin{cases} P^{\alpha} & \text{if } P \ll P_0, \text{ where } \alpha = \beta + \gamma \\ P^{\beta} & \text{if } P \gg P_0 \end{cases}. \quad (8)$$

The steep rise in occurrence, with slope α, traces out the edge of the Neptune desert at short orbital period. The sharp rise in occurrence implies there is a limit to how close planets can reside to their host star. The processes that result in this limit could be the disk inner edge inhibiting planet formation, the disk edge trapping

[4] Figure 9 shows the integrated occurrence within each histogram bin, F, or NPPS, to be consistent with all other plots in the paper. Equation (7) is given for the occurrence *rate*, or $\frac{df}{d\log P}$. To convert between the two, we multiply by the bin size, or $\log P = 0.25$. Comparisons can be made to Figure 7 of P18, which shows the period distribution as a rate, while Figure 6 of Howard et al. (2012) and Figure 12 of KM20 show the integrated rates, as we do here.

Table 1
Period Distribution Fit

R_P ($R_\oplus$)	C	β	γ	P_0 (days)
1-2	$0.13^{+0.05}_{-0.03}$	$0.26^{+0.09}_{-0.08}$	$2.5^{+0.20}_{-0.12}$	$4.10^{+0.58}_{-0.66}$
2-4	$0.21^{+0.5}_{-0.04}$	$0.17^{+0.04}_{-0.05}$	$1.94^{+0.14}_{-0.12}$	$11.12^{+1.14}_{-0.93}$

Note. Median and 68.3% confidence intervals for the fits to the parametric period distribution from Equation (7).

migrating planets, or tidal interactions removing planets. Past a break point P_0, the occurrence flattens out. The specific period P_0 where this transition occurs could also relate to limits of where planetary migration or formation is. Differences between the 1–2 $R_\oplus$ and 2–4 $R_\oplus$ break locations can indicate how those formation processes depend on planetary size.

All of the fit parameters and 68.3% confidence intervals for super-Earths (1–2 $R_\oplus$) and sub-Neptunes (2–4 $R_\oplus$) can be found in Table 1. For planets 1–2 $R_\oplus$, occurrence rises with a slope of $\alpha = 2.76^{+0.29}_{-0.20}$, to a transition period of $P_0 = 4.10^{+0.58}_{-0.66}$ days. The slope is much steeper, and the transition period is much smaller than those measured in KM20 ($\alpha = 1.9^{+0.1}_{-0.1}$, $P_0 = 5.9^{+0.5}_{-0.5}$) or P18 ($\alpha = 2.4^{+0.4}_{-0.3}$, $P_0 = 6.5^{+1.6}_{-1.2}$). The difference might be explained by P18 measuring only 1–1.7 $R_\oplus$ for their fits. We notably measure a much higher occurrence of super-Earths than KM20 (as in the radius distribution in Figure 8), which appears here as a steeper slope and earlier period break.

For the sub-Neptunes, 2–4 $R_\oplus$, we measure $\alpha = 2.11^{+0.18}_{-0.17}$, and $P_0 = 11.12^{+1.14}_{-0.93}$. These are consistent with $\alpha = 2.2^{+0.1}_{-0.1}$, and $P_0 = 13.3^{+1.4}_{-1.5}$ from KM20; and $\alpha = 2.3^{+0.2}_{-0.2}$, and $P_0 = 11.1^{+1.7}_{-1.5}$ from P18.

We measure an increase in occurrence out to long orbital period for both super-Earths and sub-Neptunes, with $\beta = 0.26^{+0.09}_{-0.08}$ for 1–2 $R_\oplus$ planets. A positive slope is in conflict with the previous studies that measure a decrease in the occurrence of small planets: KM20 measures the slope to be decreasing with $\beta = -0.5^{+0.1}_{-0.1}$ for 1–2 $R_\oplus$, and P18 measures $\beta = -0.3^{+0.2}_{-0.2}$ for $1-1.7$ $R_\oplus$. Even Dong & Zhu (2013), which measures a relatively flat slope at $\beta = -0.10 \pm 0.12$ for 1–2 $R_\oplus$, is outside of our uncertainties.

For 2–4 $R_\oplus$ planets, we also measure a slope higher than those from previous studies, with $\beta = 0.17^{+0.04}_{-0.05}$. Both KM20 and P18 measure a nearly flat slope of $\beta = -0.1^{+0.1}_{-0.1}$. Dong & Zhu (2013) does measure a rise in slope for this radius range, but it is still outside of our uncertainties at $\beta = 0.11 \pm 0.05$.

We attribute this rise in slope to the higher occurrence of planets we find in the last period bin, 189.15–400 days. For both radius ranges, this period bin is the main driver of the positive β slope. P18 does not measure past 300 days and so does not see this increase. KM20 systematically measures lower planet occurrence than our measurements below 2 $R_\oplus$. H19 does not fit any distributions to their occurrence rate estimates, but they do measure an increased occurrence for their last period bin, 256–500 days. If we exclude the bin, we measure slopes of $\beta = -0.086^{+0.01}_{-0.01}$ for the 1–2 $R_\oplus$ bin, and $\beta = 0.089^{+0.02}_{-0.01}$ for the 2–3 $R_\oplus$ bin. Without the last bin, super-Earth slope is negative and in line with those from Dong & Zhu (2013), P18. The sub-Neptune slope is still positive, but without the last bin is comparable to that from Dong & Zhu (2013). Further, if we were to take into account the effects of transit multiplicity into account, we would see an even more increased positive slope relative to the other measurements.

The larger planet populations, in the 4–8$R_\oplus$ and $8-16R_\oplus$ bins, do not show a break in occurrence but are instead a steady rise across all orbital periods. We do not see evidence for a 3 day pile-up in the hot Jupiters, as found in radial velocity surveys (Mayor et al. 2011; Wright et al. 2012; Weiss & Marcy 2014; Wittenmyer et al. 2020). Studies from Kepler have likewise not seen this overabundance (Howard et al. 2012; Fressin et al. 2013; KM20).

4.4. Stellar Type Dependence

We separate the parent FGK population into F, G, and K stars based on the effective temperature boundaries from Pecaut & Mamajek (2013). Figure 10 shows the occurrence for F stars (6000 K $\leqslant T_{\rm eff} <$ 7300 K), Figure 11 for G stars (5300 K $\leqslant T_{\rm eff} <$ 6000 K), and Figure 12 for K stars (3900 K $\leqslant T_{\rm eff} <$ 5300 K).

The occurrence between 1–400 days and 0.5–16 $R_\oplus$ is 1.05 ± 0.09 planets per F star; 1.26 ± 0.09 planets per G star; and 1.61 ± 0.13 planets per K star. As in Section 4.1, these numbers are the lower limits to occurrence and are the sum of filled cells in their respective figures; if combined and weighted appropriately, they do not add up to the given FGK occurrence. The total occurrence calculated above the strictest occurrence boundary but including the upper limits for otherwise empty cells (e.g., the Neptune desert) is 1.01 ± 0.09 planets per F star, 0.97 ± 0.06 planets per G star, and 1.47 ± 0.12 planets per K star.[5] We measure a significant increase in planet occurrence toward cooler stars.

The relationship between increased planet occurrence at lower stellar effective temperature has been well documented in occurrence rate studies (Howard et al. 2012; Mulders et al. 2015b; KM20; Yang et al. 2020). In each study, an increased occurrence for later-type stars is measured. The physical cause of the increased occurrence is not well understood. Mulders et al. (2015b) found that the increasing occurrence rates for later stellar types (M dwarf compared to FGK) cannot be simply explained by redistributing the same amount of heavy element mass of larger (sub-Neptune-sized) planets in systems around earlier-type systems into many smaller planets.

KM20 is the closest comparable study to ours. They report a much stronger trend of increasing occurrence with cooler stars: they report $0.89^{+0.23}_{-0.16}$ planets per F star; $1.67^{+0.21}_{-0.16}$ planets per G star; and $2.56^{+0.29}_{-0.24}$ planets per K star. These differences in integrated occurrence are subject to the same issues that prevent a quantitative comparison as discussed in Section 4.1.

4.4.1. Radius Valley as a Function of Stellar Type

We measure the features of the radius distribution via the same method as we did for the FGK sample for each stellar type. This analysis supposes there is an underlying two-peak structure to the occurrence. Because F, G, and K total occurrence differs, this structure is not as obvious by-eye as it is for the FGK occurrence. The choice of *two* peaks may also be suspect, as discussed at the end of Section 4.2.

As seen in Figure 13, there is a substantial shift in the minimum of the valley between stellar types: the minimum of the valley shifts to smaller radii with lower stellar effective

[5] Following, the total FGK occurrence calculated above the strictest completeness boundary (the K star boundary) but including the upper limits for otherwise empty cells is 1.09 ± 0.04. These measurements are then within firm agreement with the previously cited studies.

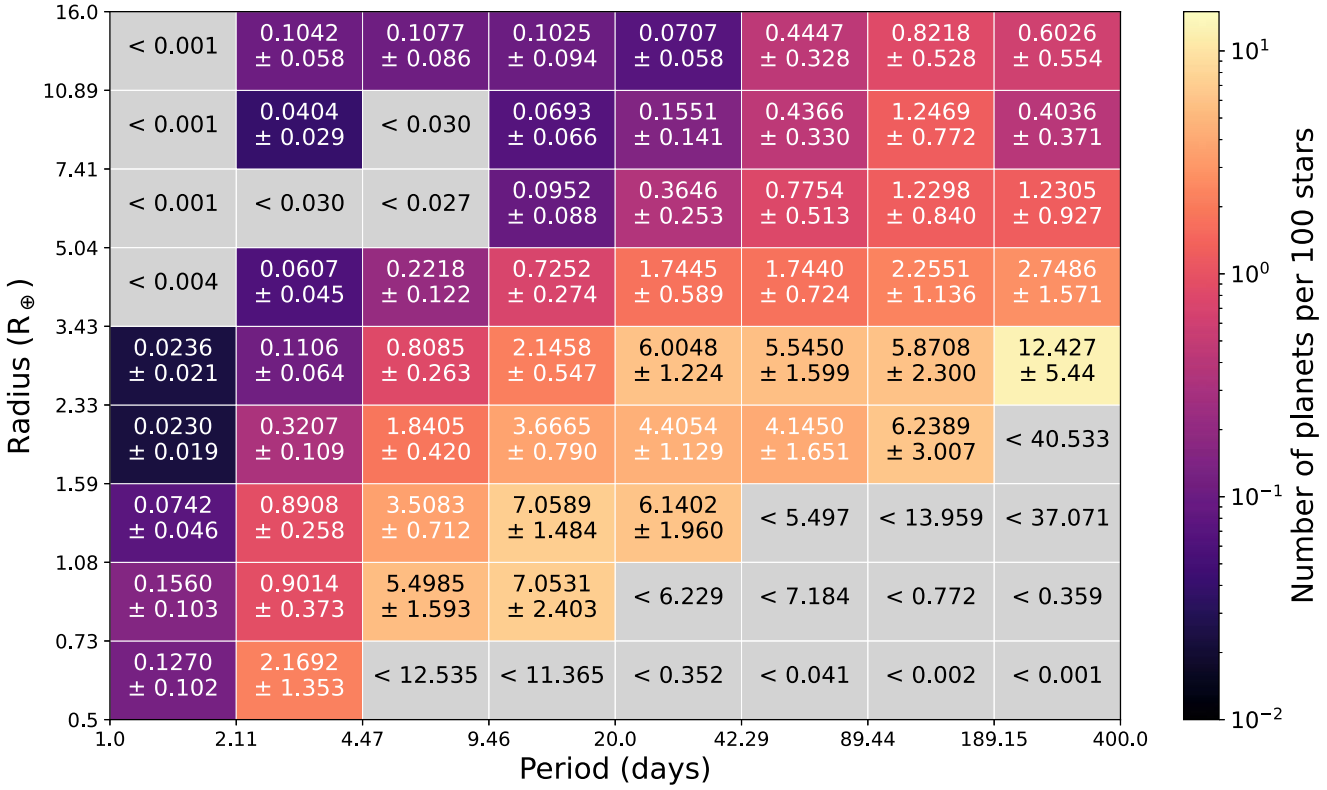

Figure 10. Same as Figure 7, calculated only for F-type stars.

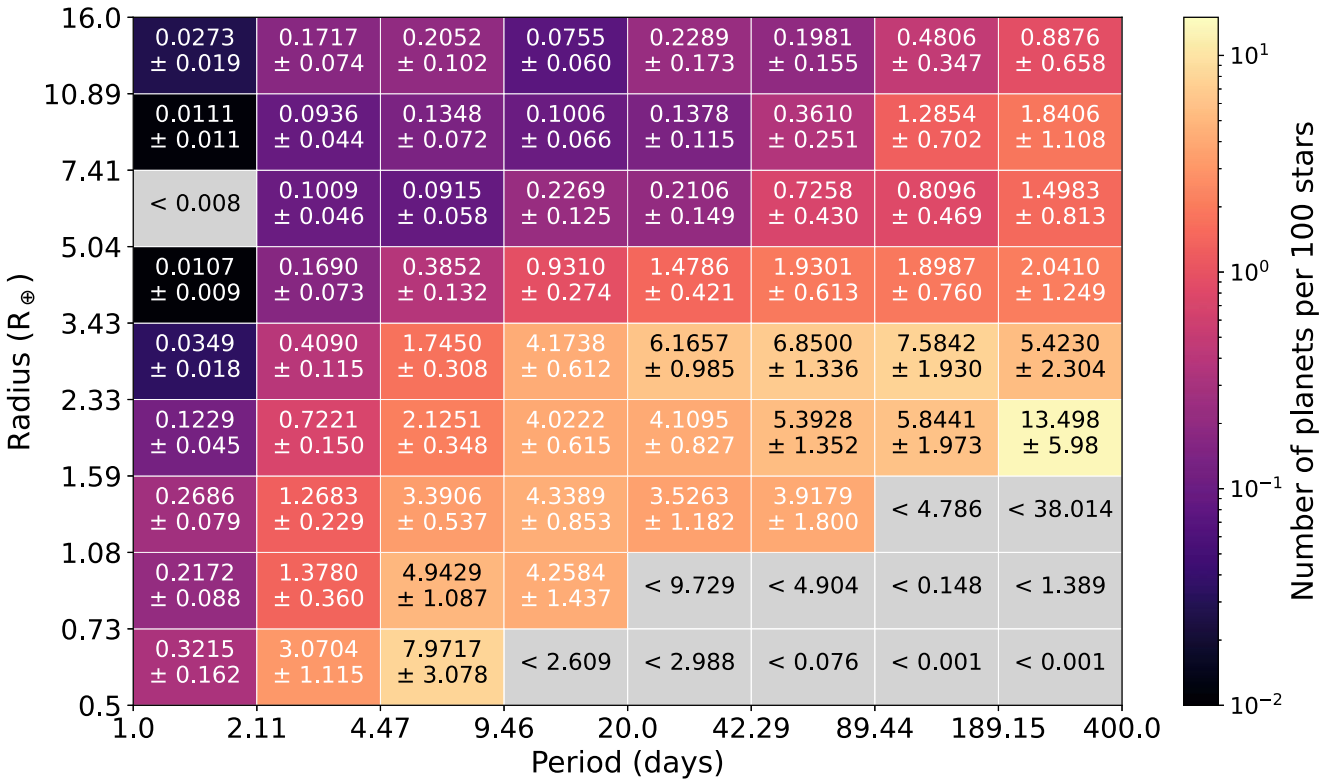

Figure 11. Same as Figure 7, calculated only for G-type stars.

temperature. We measure the minimum of the valley to be $R_{\rm min} = 1.92^{+0.26}_{-0.18}\ R_\oplus$ for F stars, $R_{\rm min} = 1.80^{+0.16}_{-0.05}\ R_\oplus$ for G stars, and $R_{\rm min} = 1.62^{+0.02}_{-0.04}\ R_\oplus$ for K stars. As with the FGK sample, we measure the location of the super-Earth and sub-Neptune peaks individually for each stellar type. The sub-Neptune peak follows the minimum of the radius valley and

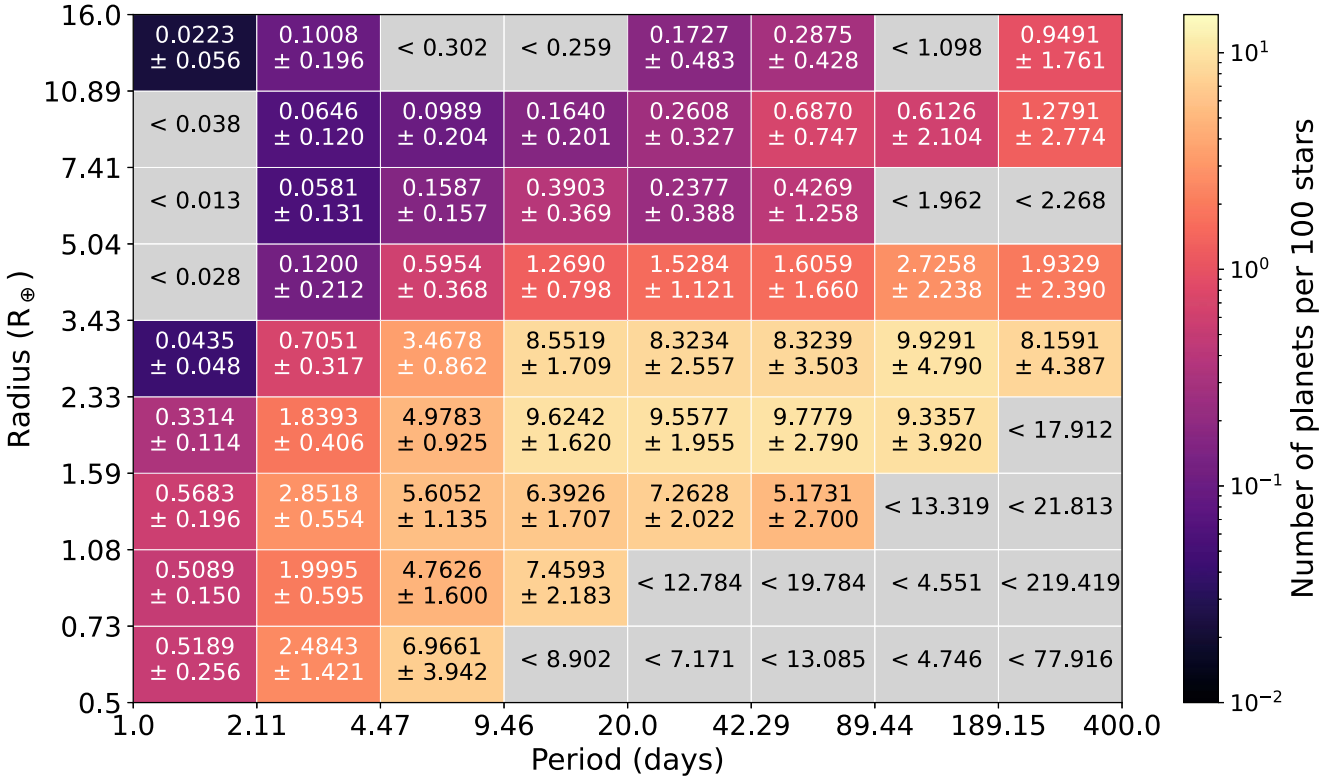

Figure 12. Same as Figure 7, calculated only for K-type stars.

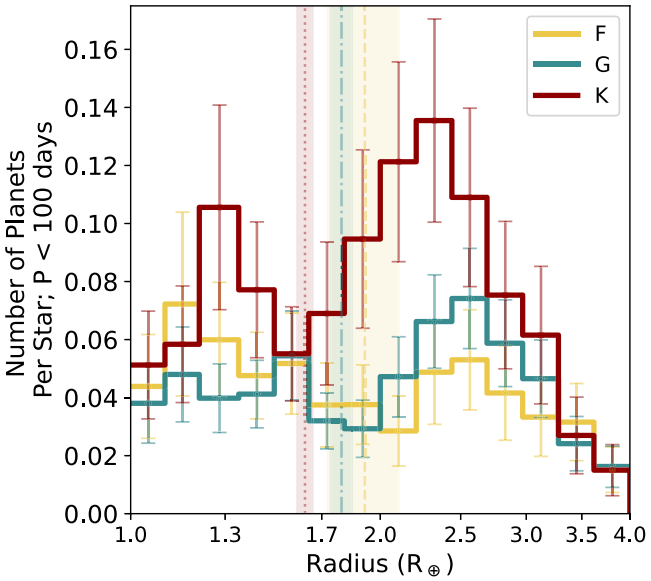

Figure 13. Population of small planets within 100 days (same as in Figure 8, now with valley minimums marked with vertical lines), now separated by stellar type. The minimum of the radius valley moves to smaller radius for cooler stars: $R_{\rm min} = 1.92^{+0.26}_{-0.18}\ R_\oplus$ for F stars (red; dotted), $R_{\rm min} = 1.80^{+0.16}_{-0.05}\ R_\oplus$ for G stars (teal; dotted–dashed), and $R_{\rm min} = 1.62^{+0.02}_{-0.04}\ R_\oplus$ for K stars (yellow; dashed), compared to $R_{\rm min} = 1.78^{+0.14}_{-0.16}\ R_\oplus$ for all FGK stars.

moves toward smaller radii for cooler stellar types. Individual measurements are reported in Table 2.

We can also compare the stellar type dependence of the radius valley to those from previous studies if we take stellar effective temperature as a proxy for stellar mass. For main-sequence stars, mass and effective temperature correlate. We

Table 2
Radius Valley Fits

	Super-Earth Peak	Minimum	Sub-Neptune Peak	Ratio
FGK	$1.37^{+0.19}_{-0.44}$	$1.78^{+0.14}_{-0.16}$	$2.51^{+0.06}_{-0.08}$	$2.24^{+0.78}_{-0.64}$
F	$1.28^{+0.32}_{-0.76}$	$1.92^{+0.26}_{-0.18}$	$2.58^{+0.19}_{-0.59}$	$1.37^{+0.83}_{-0.52}$
G	$1.41^{+0.24}_{-0.59}$	$1.80^{+0.16}_{-0.05}$	$2.61^{+0.07}_{-0.08}$	$2.50^{+0.64}_{-0.71}$
K	$1.39^{+0.25}_{-0.12}$	$1.62^{+0.02}_{-0.04}$	$2.34^{+0.08}_{-0.09}$	$3.41^{+0.77}_{-0.68}$

Note. Measured locations of radius distribution features for different stellar types. Each location given in $R_\oplus$.

can derive the relationship via the Stefan–Boltzmann law and the classical mass–luminosity relationship $L = 4\pi R^2 \sigma T_{\rm eff}^4$, and $\frac{L}{L_\odot} = \frac{M}{M_\odot}^{\alpha_C}$, where classically $\alpha_C = 3.5$. Eker et al. (2018) uses a variable mass–luminosity relationship for different mass ranges, with α_C ranging from 2.028–5.743. For all proposed values of α_C, the relationship has a positive correlation and classically is $M \propto T^{\frac{8}{7}}$.

The stellar mass dependence of the radius valley has been investigated in Wu (2019), Berger et al. (2020b), Petigura et al. (2022), and Ho & Van Eylen (2023) by quantifying the slope, α, of the valley in the stellar mass versus planet radius plane, with values ranging from $\alpha = 0.23$–0.35. Wu (2019) investigated the mass dependence of small planets (both super-Earths and sub-Neptunes together) on host star mass. The analysis of 1841 confirmed Kepler planets (using Gaia-derived stellar mass and planet radii) finds a positive correlation for both super-Earth and sub-Neptune population's radius with stellar mass, at a slope of $\alpha = 0.23$–0.35, indicating that there is a linear relationship between planetary core mass and host star mass. Berger et al. (2020a) find agreement and measure the slope of the radius valley in radius–stellar mass space to be

$\alpha = 0.26^{+0.21}_{-0.16}$. Petigura et al. (2022) looked at the radius valley as a function of stellar mass in the CKS survey. They do not find a constant slope for the radius valley in radius–stellar mass space, instead finding that the super-Earth population stays constant over stellar mass and that the sub-Neptune population grows to larger radii with stellar mass. The positive correlation between planet size and stellar mass for the sub-Neptunes implies that larger stars produce larger planet cores, agreeing with Wu (2019). Ho & Van Eylen (2023) use Kepler short cadence observations to find a slope of $\alpha = 0.231^{+0.053}_{-0.06}$. It should be noted that these analyses are in the *observed* density of planets, and not in the occurrence of planets. Observed versus intrinsic populations could explain why we see a shift in the lower edge of the radius valley where they see that the super-Earth population is constant.

We do not directly measure the slope of the valley as a function of stellar mass–planet radius, but can make an approximation by looking at the average stellar mass within each stellar type bin as a function of the measured minimum of the radius valley integrated from 1–30 days, plotted in Figure 14. Our measurements are in good agreement with the previous studies. We further compare this relationship with theoretical models in Section 5.1.

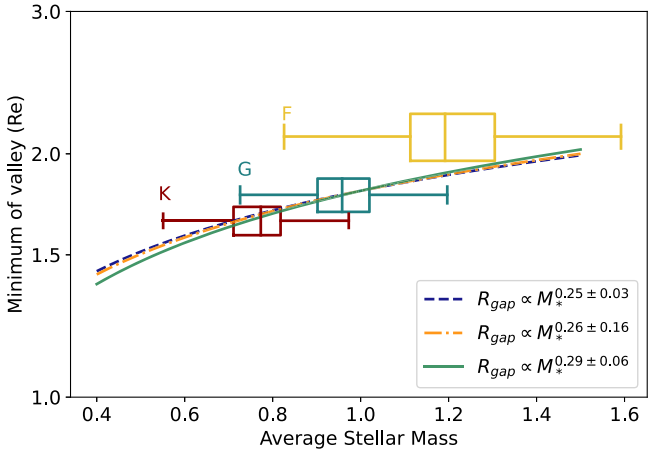

Figure 14. Stellar mass-dependence of the radius valley. We do not compute occurrence as a function of stellar mass but plot box plots of the mass distribution for each stellar type. The center line of the box plot represents the median stellar mass, with first and second quartiles being marked by the edges of the box. The whiskers extend to the minimum and maximum stellar mass values. The height of the box is the uncertainty in the measured minimum of the radius valley integrated from 1–30 days. Observed fits to the slope of the radius valley in the planet radius–stellar mass plane are plotted; our measurements are in agreement.

4.5. The Occurrence Cliff

We pay special attention to the occurrence cliff–the sharp decline in occurrence between 3 and 4 $R_\oplus$. The feature is one of the most readily apparent features in the planet distribution. However, it is the focal point of very few theoretical studies. Evolutionary models focusing on the radius valley often yield an occurrence cliff qualitatively similar to that observed by Kepler, yet the physical processes and tuning parameters required to yield a steep cliff are not described in detail.

Before correcting for completeness, the cliff appears to run along a constant radius contour. However, the slope of the occurrence cliff is not the same at every orbital period: the cliff extends to larger radii for longer orbital periods; i.e., the slope of the occurrence cliff becomes flatter with increasing orbital period. This phenomenon has been noticed as early as Dong & Zhu (2013), who noticed the "fast rise of the Neptunes" in early Kepler data.

We inspect this effect in Figure 15 by looking at the radius distribution. Instead of looking only at occurrence below 100 days, we break the radius distribution into three period bins: low (10–100 days), mid (100–300 days), and high (300–500 days).[6] We measure the slope of the cliff between 2.5 and 6 $R_\oplus$ for each bin and find that the occurrence cliff becomes gentler (with a slope closer to zero) for the longer period bins.

We measure the slope of each period bin with the following equation[7]:

$$\text{NPPS} = F = m_C \log R_\oplus + b \tag{9}$$

where m_c is the slope of occurrence cliff, and b is an intercept point. Similar to the distributions fit to the radius gap, we fit each iteration of our bootstrap simulation to measure the slope of each period bin. For our FGK sample, we measure mean slopes of $m_C = -0.41^{+0.01}_{-0.01}$, for $P = 10$–100 days; $m_C = -0.21^{+0.02}_{-0.02}$, for $P = 100$–300 days; and $m_C = -0.12^{+0.03}_{-0.21}$, for $P = 300$–500 days. The right panel of Figure 15 shows the measured slopes for each period bin. The slope of the cliff gets progressively flatter for longer orbital periods. There are no similar measurements in the literature, so we do not compare here.

4.5.1. Occurrence Cliff as a Function of Stellar Type

Similar to the radius valley, we can also look at how the occurrence cliff changes over stellar type. As in Section 4.5, we break the radius distribution into three period bins: low (10–100 days), mid (100–300 days), and high (300–500 days). We measure the slope of each period bin for each stellar type with the same process and linear function as for the FGK sample. Table 3 reports the measured mean slopes and 68.3% confidence intervals, as well as the number of planets within each bin. Figure 16 shows the measured cliff slopes for each stellar type. Each stellar type follows the same trend of the shortest period bin having the steepest slope, with the slope becoming flatter with each progressive period bin. The slopes become more similar with a longer orbital period for each stellar type. However, the rate of change of the slope appears to be more gradual for F stars compared to G and K stars. We exclude the cliff slope measurements for K occurrence for the longest period bin because there are only 2 planets that reside within it.

5. Discussion

Our KDE methodology gives us the ability to measure structure in multiple dimensions that we cannot measure with parametric forms. It is flexible for comparisons with previous studies and incorporates the best-of-our knowledge input parameters in a uniform manner. There are limitations to this methodology, though. We cannot extrapolate occurrence rates

[6] The longest period bin goes out to 500 days (instead of 400 as all other occurrence measurements) to cover the largest area possible to compare with the other period bins. The impacts of the completeness degradation discussed in Section 3.1 affect smaller radii more strongly than this regime.

[7] Note that we are finding the slope of the integrated occurrence here, not the occurrence rate as in Equation (7).

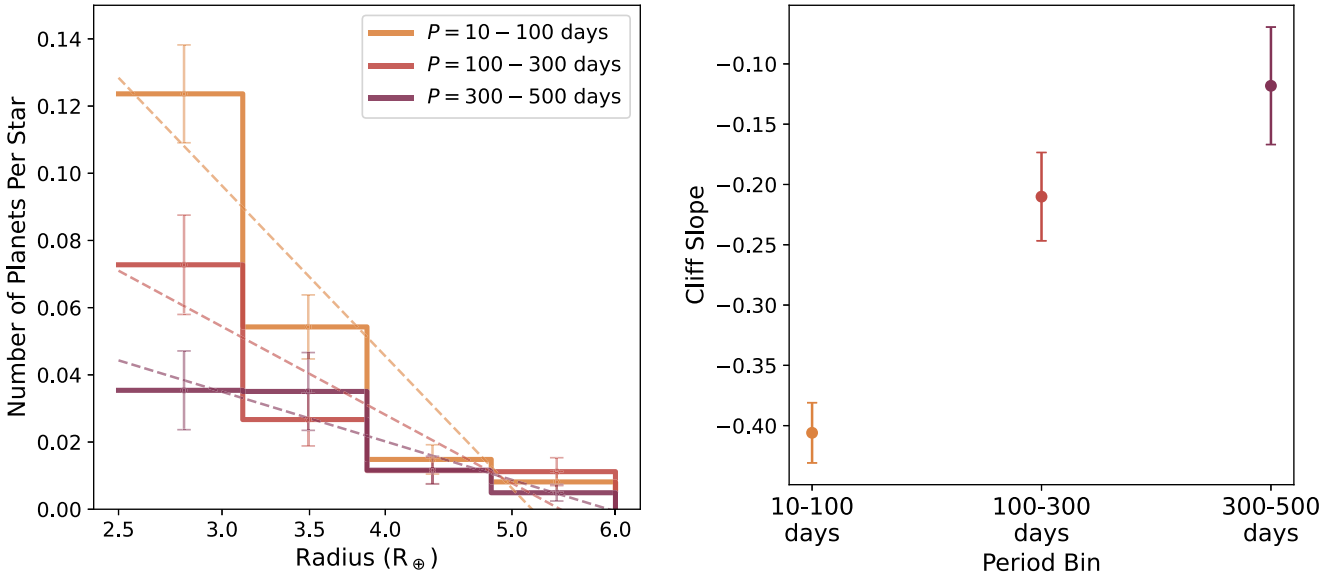

Figure 15. Left: radius distribution for planets between 2.5 and 6 $R_{\oplus}$ separated into three different period bins, low (10–100 days, orange); mid (100–300 days, pink); high (300–500 days, purple). The occurrence of short-period planets is steeper than the occurrence of planets at long orbital period. Right: distribution of occurrence cliff slopes for three different period bins, measured for 1000 Monte Carlo runs. The shortest period bin, 10–100 days (orange), has the steepest slope of $m_C = = -0.406^{+0.012}_{0.012}$; the middle period bin, 100–300 days (pink), has a slope of $m_C = -0.210^{+0.018}_{-0.017}$; and the longest period bin, 300–500 days (purple), has the flattest slope at $m_C = -0.121^{+0.026}_{-0.21}2$.

Table 3
Occurrence Cliff Slope Fits

	10–100 days	100–300 days	300–500 days
FGK	$-0.406^{+0.01}_{-0.01}$ (447)	$-0.210^{+0.02}_{-0.02}$ (75)	$-0.121^{+0.03}_{-0.02}$ (16)
F	$-0.304^{+0.02}_{-0.02}$ (122)	$-0.238^{+0.04}_{-0.04}$ (24)	$-0.194^{+0.05}_{-0.04}$ (5)
G	$-0.411^{+0.02}_{-0.02}$ (218)	$-0.177^{+0.02}_{-0.02}$ (32)	$-0.093^{+0.04}_{-0.03}$ (9)
K	$-0.559^{+0.03}_{-0.03}$ (107)	$-0.268^{+0.05}_{-0.04}$ (19)	na (2)

Note. Measured cliff slopes for each period bin and stellar type. The parenthetical numbers beside each slope measurement are the number of planets inside that bin.

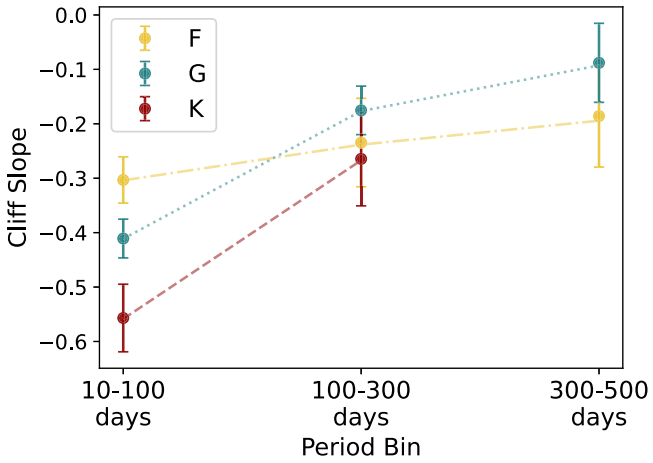

Figure 16. Slope of the occurrence cliff for each stellar type. For each stellar type, the slope of the occurrence cliff becomes flatter for longer orbital period. The K- (red) and G-star (teal) occurrence have similar changes in cliff slope; the F-star (yellow) occurrence is more similar from low to high orbital period. We exclude the last bin for K stars because it contains only 2 planet detections.

to areas of parameter space with low or no planet detections, such as the habitable zone.

It is rarely possible to reproduce independent experiments exactly, but we have tested a variety of benchmark measurements across a broad range of parameter space and find excellent agreement with previous studies. Our measurements for total occurrence, hot Jupiter occurrence, and the radius distribution are all within 1σ of previous studies.

We find the most tension with regards to small radius and/or long period planets, such as the turn over of the super-Earth population or the slope of their period distribution. This is to be expected: there are few planet detections in these regions, and both completeness and reliability are at their lowest. These planets require very careful treatment to resolve disagreement.

5.1. What Causes the Shift in the Radius Valley?

When looking at the small planet radius distribution separated by stellar type, there is a distinct shift in the minimum of the radius valley to smaller radii for cooler stars. This shift could be caused by formation or evolutionary processes.

The first possible explanation is that hotter stars can strip the volatile envelopes of larger cores. According to atmospheric mass loss models, the bottom edge of the valley corresponds to bare rocky cores, and the top edge corresponds to cores with ∼4% H–He atmospheres (Owen & Wu 2017). If the X-ray and ultraviolet (XUV) flux from the star is strong enough, larger cores are completely stripped, pushing the lower edge of the valley to higher radii. There is a slight dependence of integrated lifetime XUV with stellar mass (McDonald et al. 2019), so hotter stars could indeed photoevaporate planets more easily than cooler stars.

Core-powered mass loss also causes planets to lose their primordial envelopes, primarily through luminosity from the cooling rocky cores Ginzburg et al. (2018). A mechanism solely caused by the planet, independent of the stellar host environment, could not explain the shift in the radius valley that we see. However, core-powered mass loss has a dependence on the bolometric flux from the host star

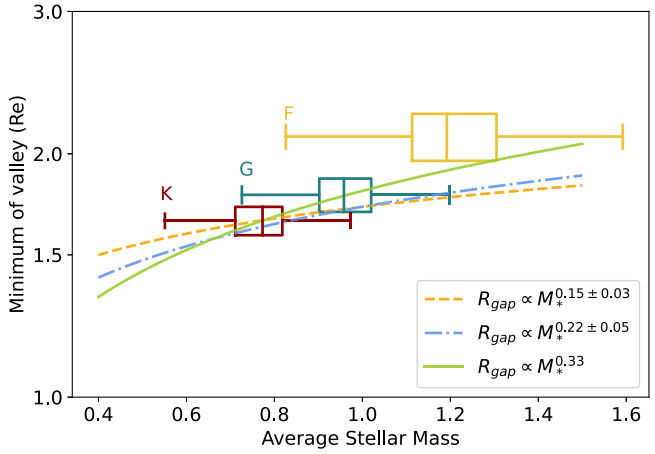

Figure 17. Dependence of the radius valley as a function of stellar mass. We do not measure occurrence as a function of stellar mass; here, we investigate this dependence by plotting a box plot to represent the distribution of stellar masses for each stellar type. The center line for each box plot represents the median stellar mass value. The left and right edges of the box show one quartile above and below the median, while the whiskers show the minimum and maximum mass values. The width of the box is the uncertainty in the radius valley minimum measurement integrated from 1–30 days. We show power-law fits to the radius valley from Lee et al. (2022): the bottom-heavy core mass model ($R_{\rm gap} \propto M_*^{0.22 \pm 0.05}$) in blue and the top-heavy core mass model ($R_{\rm gap} \propto M_*^{0.15 \pm 0.03}$) in orange, as well as the power law derived from core-powered mass-loss models ($R_{\rm gap} \propto M_*^{0.33}$ Gupta & Schlichting 2020). Our results are most consistent with the steeper slope of the core-powered mass-loss model.

(via equilibrium temperature, which sets cooling and mass-loss rates), and thus cannot be ruled out.

In order to disentangle these mechanisms, we need to look at the distribution of planets as a function of planet radius, insolation flux, and stellar mass simultaneously. Some recent work (e.g., Rogers & Owen 2021; Berger et al. 2023) has begun to study this, but we will defer studying it in occurrence space to a future paper.

There could also be a formation cause for the shift in the valley. Lee et al. (2022) theorized that with more realistic disk physics, the radius valley can be created without any atmospheric mass loss processes by gas accretion alone. There is a maximal isothermal envelope mass each planet is able to accrete according to how massive the planet core is and its disk location. Small cores, 2–3$M_\oplus$, are unable to accrete any atmosphere and remain bare, while massive cores >10 $M_\oplus$ undergo runaway accretion before being halted by disk dispersal. The isothermal-to-cooling transition carves out a gap at $\sim$1.2 $R_\oplus$ in the radius distribution, replicating the observed radius valley.

This phenomenon should replicate the trend we see with a shifted radius valley with both stellar temperature and mass. We take the average stellar mass within each stellar type for our sample and plot it against our measured minimum of the radius valley from 1–30 days, seen in Figure 17. We can compare this to the power laws that Lee et al. (2022) established for bottom-heavy (blue) and top-heavy (orange) core-mass distributions, as seen in Figure 12 from Lee et al. (2022). This theoretical fit for the gap agrees with those found in Petigura et al. (2022), Berger et al. (2020a). As with the comparison with the observational fits in Figure 14, our measurements show good agreement with the models. However, a steeper slope like the core-powered mass-loss model (Gupta & Schlichting 2020, green) is preferred.

5.2. Change in the Slope of the Occurrence Cliff

The same thing that shifts the location of the radius valley can shape the occurrence cliff, as it is often an outcome of models that replicate the radius valley. The change in cliff slope across this wide span in period could be tied to different physical processes. If we assume the main drivers of the radius valley–photoevaporation and core-powered mass loss–primarily work within 100 days, the difference in slopes for the low- and mid-period bins (10–100 days versus 100–300 days) can be attributed to the drop-off in efficacy in those physical processes. As a planet becomes less irradiated and cooler at longer orbital period, it can potentially retain more of its envelope and sit at a larger radius. As a population, a wider range of radii causes the cliff slope to be gentler.

If the drivers of the radius valley mainly act within 100 days, they cannot explain the continued trend in gentler cliff slopes we see beyond 300 days. Instead, this more closely matches the prediction from Kite et al. (2019), where they theorized that the cause of the occurrence cliff is the "fugacity crisis," where the radii of sub-Neptune-sized planets are limited by the sequestration of hydrogen in the planet's magma ocean. This mechanism is less efficient, on average, for cooler planet temperatures. The implication is that the planets at longer orbital periods tend to maintain larger radii, thereby reducing the slope of the occurrence cliff.

The changing slope of the cliff could also trace a changing core-mass distribution for small planets. As the orbital period (and therefore semimajor axis) increases, the availability of materials in the disk increases, which could lead to increased planet core sizes.

6. Summary and Conclusion

In this paper, we do the following:

1. We present a KDE methodology for measuring Kepler occurrence rates. We incorporate various measurement uncertainties, Gaia DR2 stellar radii, and reliability measurements into the calculations.
2. We validate the method with a toy model and quantify the level of structure we can measure, KDE edge effects, and a low-completeness boundary where we are unable to measure occurrence below.
3. We measure the period and radius distribution of the Kepler FGK sample and compare to previous studies. We find a total FGK occurrence of 1.52 ± 0.08 NPPS between 1–400 days and 0.5–16 $R_\oplus$.
4. We fit the period distribution to a parametric distribution from Howard et al. (2012) and find that our fit agrees with previous measurements.
5. We measure the shape of the small planet radius distribution. We find the minimum of the radius valley to be $1.78^{+0.14}_{-0.16}$ $R_\oplus$.
6. We fit the period distribution to a parametric distribution from Howard et al. (2012) and find that our fit agrees with previous measurements.
7. We measure the occurrence dependent on stellar type and confirm that the occurrence increases toward cooler stars. We find the total occurrences ($P = 1$–400 days; $R = 0.5$–16 $R_\oplus$) of 1.05 ± 0.09 NPPS for F stars, 1.26 ± 0.09 NPPS for G stars, and 1.61 ± 0.13 NPPS for K stars.

8. We find that the minimum of the radius valley moves to smaller radii for cooler stars, from $1.92^{+0.26}_{-0.18}$ $R_\oplus$ for F stars, to $1.80^{+0.16}_{-0.05}$ $R_\oplus$ for G stars, to $1.62^{+0.02}_{-0.04}$ $R_\oplus$ for K stars.
9. We measure the slope of the occurrence cliff and find it to be less steep at longer orbital period. There is also a dependence on stellar type; the total change between short and long orbital period bins for cliff slope becomes larger for cooler stars.

Our methodology enables us to map the occurrence structure to physical models of planet formation and evolution theory because we can measure any type of underlying structure in planet occurrence. Throughout Section 4, we pointed to measurements that could be directly tied to theories. The next step is to use this framework to compare these measurements with populations derived from models of planet formation and evolution.

The features we look at in this paper, namely the radius valley and occurrence cliff, will be more informative in insolation rather than orbital period, to mitigate the effects of varying stellar biases. We reserve this work for a future paper.

Utilizing multidimensional occurrence measurements will allow us to constrain input parameters for planet formation and evolution. In the future, we will use population synthesis models in a Bayesian framework to evolve planets through physical models and measure their occurrence to see if they match our actual measurements. This will enable us not only to constrain free parameters in the theoretical models but also to compare theoretical models to each other.

Acknowledgments

The authors would like to thank the anonymous referee whose comments helped improve the paper. A.D. would like to thank Michelle Kunimoto, Danley Hsu, Kendall Sullivan, Ruth Murray-Clay, James Rogers, Joey Murphy, Jonathan Fortney, and Adam Kraus for helpful conversations in the completion of this manuscript. AD gratefully acknowledges support from the Heising-Simons Foundation through grant 2021-3197. We acknowledge use of the lux supercomputer at University of California Santa Cruz, funded by NSF MRI grant AST 1828315. This work used the Extreme Science and Engineering Discovery Environment (XSEDE) *Expanse* at the San Diego Supercomputer Center through allocation TG-PHY210033 (Towns et al. 2014).

Software: Additional packages used include *Astropy* (Astropy Collaboration et al. 2013), *numpy* (van der Walt et al. 2011), *emcee* (Foreman-Mackey et al. 2013), *pandas* (McKinney 2010), *scipy* (Virtanen et al. 2020), *matplotlib* (Hunter 2007).

Appendix A
Vetting Completeness

Following Bryson et al. (2020a), we grid period-expected MES space and treat the fraction of recovered PCs in each cell as a binomial rate. We fit the rate ρ with a product of a nonrotated simplified logisitic function in period p times a rotated logistic in p and expected MES m. This gives us 6 parameters to fit with our Bayseian inference: $\theta = [x_0, y_0, k_x, k_y, \phi, A]$; where

$$
\begin{aligned}
x &= \frac{(p - p_{\rm min})}{p_{\rm max} - p_{\rm min}};\\
y &= \frac{(m - m_{\rm min})}{m_{\rm max} - m_{\rm min}};\\
y_{\rm rot} &= (y - 0.5)\cdot\cos\phi - (x - 0.5)\cdot\sin\phi;\\
\rho &= A\ Y(x, x_0, -k_x, 1)\times Y(y_{\rm rot} + 0.5, y_0, k_y, 1).
\end{aligned}
$$

Our MCMC computation used 100 walkers and ran for 10,000 steps. We ensured convergence by checking that the number of steps were greater than 50× the estimate of the integrated autocorrelation time. The maximum parameter was $\tau = 92$.

For our FGK stellar sample, the MCMC median posteriors and 84th and 16th percentiles are as follows:

$x_0 = 1.040^{+0.035}_{-0.031}$, $y_0 = 0.153^{+0.006}_{-0.007}$,
$k_x = 5.055^{+0.501}_{-0.489}$, $k_y = 18.877^{+1.320}_{-1.283}$,
$\phi = 4.019^{+0.911}_{-0.875}$, and $A = 0.964^{+0.005}_{-0.005}$.

Our separate F, G, and K inferences are all within 1σ uncertainties for the FGK fits, so we do not report them here.

Appendix B
Reliability

B.1. False Alarm Effectiveness

We follow the same process as vetting completeness, following Bryson et al. (2020a), to characterize the false alarm effectiveness, instead fitting the binomial rate ρ with a simple rotated logistic function for $\Theta = [x_0, k_x, \phi, A]$:

$$
\begin{aligned}
x &= \frac{(p - p_{\rm min})}{p_{\rm max} - p_{\rm min}};\\
y &= \frac{(m - m_{\rm min})}{m_{\rm max} - m_{\rm min}};\\
x_{\rm rot} &= (x - 0.5)\cdot\cos\phi - (y - 0.5)\cdot\sin\phi;\\
\rho &= A\ Y(x_{\rm rot} + 0.5, x_0, k_y, 1).
\end{aligned}
$$

Our MCMC computation used 100 walkers and ran for 10,000 steps. We ensured convergence by checking that the number of steps were greater than 50× the estimate of the integrated autocorrelation time. The maximum parameter was $\tau = 197$. For our FGK stellar sample, the MCMC median posteriors and 84th and 16th percentiles are as follows:

$x_0 = 1.079^{+0.034}_{-0.025}$, $k_x = 51.6292^{+4.585-}_{-15.666}$,
$A = 0.996^{+0.001}_{-0.001}$, and $\phi = 91.151^{+0.694}_{-0.529}$.

Our separate F, G, and K inferences are all within 1σ uncertainties for the FGK fits, so we do not report them here.

B.2. Observed False-positive Rate

We follow the same process as the false alarm effectiveness, following Bryson et al. (2020a), to characterize the observed

false-positive rate, instead fitting the binomial rate ρ with a simple rotated logistic function for $\Theta = [x_0, k_x, \phi, A]$:

$$x = \frac{(p - p_{\rm min})}{p_{\rm max} - p_{\rm min}};$$
$$y = \frac{(m - m_{\rm min})}{m_{\rm max} - m_{\rm min}};$$
$$x_{\rm rot} = (x - 0.5) \cdot \cos\phi - (y - 0.5) \cdot \sin\phi;$$
$$\rho = A\ Y(x_{\rm rot} + 0.5, x_0, k_y, 1).$$

Our MCMC computation used 100 walkers and ran for 10,000 steps. We ensured convergence by checking that the number of steps were greater than 50× the estimate of the integrated autocorrelation time. The maximum parameter was $\tau = 154$.

For our FGK stellar sample, the MCMC median posteriors and 84th and 16th percentiles are as follows:
$x_0 = 0.639^{+0.019}_{-0.019}$, $k_x = 9.320^{+0.411}_{-0.377}$,
$A = 0.991^{+0.004}_{-0.005}$, and $\phi = -157.457^{+2.039}_{-2.012}$.

Our separate F, G, and K inferences are all within 1σ uncertainties for the FGK fits, so we do not report them here.

ORCID iDs

Anne Dattilo https://orcid.org/0000-0002-1092-2995
Natalie M. Batalha https://orcid.org/0000-0002-7030-9519
Steve Bryson https://orcid.org/0000-0003-0081-1797

References

Akaike, H. 1974, ITAC, 19, 716
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Beaugé, C., & Nesvorný, D. 2013, ApJ, 763, 12
Beleznay, M., & Kunimoto, M. 2022, MNRAS, 516, 75
Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, ApJ, 866, 99
Berger, T. A., Huber, D., Gaidos, E., van Saders, J. L., & Weiss, L. M. 2020a, AJ, 160, 108
Berger, T. A., Huber, D., van Saders, J. L., et al. 2020b, AJ, 159, 280
Berger, T. A., Schlieder, J. E., Huber, D., & Barclay, T. 2023, arXiv:2302.00009
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Sci, 327, 977
Borucki, W. J., Koch, D. G., Basri, G., et al. 2011, ApJ, 736, 19
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Bryan, M. L., Knutson, H. A., Howard, A. W., et al. 2016, ApJ, 821, 89
Bryson, S., Coughlin, J., Batalha, N. M., et al. 2020a, AJ, 159, 279
Bryson, S., Coughlin, J. L., Kunimoto, M., & Mullally, S. E. 2020b, AJ, 160, 200
Burke, C. J., & Catanzarite, J. 2017, Kepler Science Document, KSCI-19111-002
Burke, C. J., Christiansen, J. L., Mullally, F., et al. 2015, ApJ, 809, 8
Catanzarite, J., & Shao, M. 2011, ApJ, 738, 151
Christiansen, J. L. 2017, Kepler Science Document, KSCI-19110-001
Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2013, ApJS, 207, 35
Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2015, ApJ, 810, 95
Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2016, ApJ, 828, 99
Coughlin, J. 2017, Kepler Science Document, KSCI-19114-002
Dawson, R. I., & Murray-Clay, R. A. 2013, ApJL, 767, L24
Dong, S., & Zhu, Z. 2013, ApJ, 778, 53
Dressing, C. D., & Charbonneau, D. 2013, ApJ, 767, 95
Dressing, C. D., & Charbonneau, D. 2015, ApJ, 807, 45
Eker, Z., Bakış, V., Bilir, S., et al. 2018, MNRAS, 479, 5491
Fernandes, R. B., Mulders, G. D., Pascucci, I., Mordasini, C., & Emsenhuber, A. 2019, ApJ, 874, 81
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Foreman-Mackey, D., Hogg, D. W., & Morton, T. D. 2014, ApJ, 795, 64
Fressin, F., Torres, G., Charbonneau, D., et al. 2013, ApJ, 766, 81
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
Gaidos, E., Mann, A. W., Kraus, A. L., & Ireland, M. 2016, MNRAS, 457, 2877
Ginzburg, S., Schlichting, H. E., & Sari, R. 2016, ApJ, 825, 29
Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, MNRAS, 476, 759
Guo, X., Johnson, J. A., Mann, A. W., et al. 2017, ApJ, 838, 25
Gupta, A., & Schlichting, H. E. 2020, MNRAS, 493, 792
Hardegree-Ullman, K. K., Cushing, M. C., Muirhead, P. S., & Christiansen, J. L. 2019, AJ, 158, 75
Ho, C. S. K., & Van Eylen, V. 2023, MNRAS, 519, 4056
Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, ApJS, 201, 15
Hsu, D. C., Ford, E. B., Ragozzine, D., & Ashby, K. 2019, AJ, 158, 109
Hunter, J. D. 2007, CSE, 9, 90
Jenkins, J. M. 2002, ApJ, 575, 493
Jin, S. 2021, MNRAS, 502, 5302
Kane, S. R., Kopparapu, R. K., & Domagal-Goldman, S. D. 2014, ApJL, 794, L5
Kite, E. S., Fegley, Bruce, J., Schaefer, L., & Ford, E. B. 2019, ApJL, 887, L33
Kopparapu, R. K. 2013, ApJL, 767, L8
Kunimoto, M., & Matthews, J. M. 2020, AJ, 159, 248
Kunimoto, M., Matthews, J. M., & Ngo, H. 2020, AJ, 159, 124
Lee, E. J., & Chiang, E. 2016, ApJ, 817, 90
Lee, E. J., Karalis, A., & Thorngren, D. P. 2022, ApJ, 941, 186
Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2
Lopez, E. D., & Fortney, J. J. 2013, ApJ, 776, 2
Mann, A. W., Gaidos, E., Lépine, S., & Hilton, E. J. 2012, ApJ, 753, 90
Martinez, C. F., Cunha, K., Ghezzi, L., & Smith, V. V. 2019, ApJ, 875, 29
Mathur, S., Huber, D., Batalha, N. M., et al. 2017, ApJS, 229, 30
Mayor, M., Marmier, M., Lovis, C., et al. 2011, arXiv:1109.2497
Mazeh, T., Holczer, T., & Faigler, S. 2016, A&A, 589, A75
McDonald, G. D., Kreidberg, L., & Lopez, E. 2019, ApJ, 876, 22
McKinney, W. 2010, in Proc. of the 9th Python in Science Conf., ed. S. van der Walt & J. Millman (Austin, TX: SciPy), 56
Morton, T. D., Bryson, S. T., Coughlin, J. L., et al. 2016, ApJ, 822, 86
Morton, T. D., & Swift, J. 2014, ApJ, 791, 10
Mulders, G. D., Pascucci, I., & Apai, D. 2015a, ApJ, 798, 112
Mulders, G. D., Pascucci, I., & Apai, D. 2015b, ApJ, 814, 130
Mulders, G. D., Pascucci, I., & Apai, D. 2015c, ApJ, 798, 112
Mulders, G. D., Pascucci, I., Apai, D., Frasca, A., & Molenda-Żakowicz, J. 2016, AJ, 152, 187
Mullally, F., Coughlin, J. L., Thompson, S. E., et al. 2016, PASP, 128, 074502
Narang, M., Manoj, P., Furlan, E., et al. 2018, AJ, 156, 221
Owen, J. E., & Lai, D. 2018, MNRAS, 479, 5012
Owen, J. E., & Wu, Y. 2017, ApJ, 847, 29
Pecaut, M. J., & Mamajek, E. E. 2013, ApJS, 208, 9
Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, PNAS, 110, 19273
Petigura, E. A., Howard, A. W., Marcy, G. W., et al. 2017, AJ, 154, 107
Petigura, E. A., Marcy, G. W., Winn, J. N., et al. 2018, AJ, 155, 89
Petigura, E. A., Rogers, J. G., Isaacson, H., et al. 2022, AJ, 163, 179
Rogers, J. G., & Owen, J. E. 2021, MNRAS, 503, 1526
Schwarz, G. 1978, AnSta, 6, 461
Silburt, A., Gaidos, E., & Wu, Y. 2015, ApJ, 799, 180
Silverman, B. W. 1986, Density Estimation for Statistics and Data Analysis (London: Chapman and Hall)
Sullivan, K., Kraus, A. L., & Mann, A. W. 2022, ApJ, 935, 141
Swift, J. J., Johnson, J. A., Morton, T. D., et al. 2013, ApJ, 764, 105
Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, ApJS, 235, 38
Thompson, S. E., Mullally, F., Coughlin, J., et al. 2015, ApJ, 812, 46
Towns, J., Cockerill, T., Dahan, M., et al. 2014, CSE, 16, 62
Traub, W. A. 2012, ApJ, 745, 20
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2018, MNRAS, 479, 4786
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Wang, J., Fischer, D. A., Horch, E. P., & Huang, X. 2015, ApJ, 799, 229
Weiss, L. M., & Marcy, G. W. 2014, ApJL, 783, L6
Wittenmyer, R. A., Wang, S., Horner, J., et al. 2020, MNRAS, 492, 377
Wright, J. T., Marcy, G. W., Howard, A. W., et al. 2012, ApJ, 753, 160
Wu, Y. 2019, ApJ, 874, 91
Yang, J.-Y., Xie, J.-W., & Zhou, J.-L. 2020, AJ, 159, 164
Youdin, A. N. 2011, ApJ, 742, 38
Zhu, W., Petrovich, C., Wu, Y., Dong, S., & Xie, J. 2018, ApJ, 860, 101
Zink, J. K., Christiansen, J. L., & Hansen, B. M. S. 2019, MNRAS, 483, 4479