Modeling The True Underlying Mass-Radius-Period Distribution of Kepler Exoplanets

Steven M. Blodgett[1] - Darin Ragozzine[1] - Daniel K. Jones[1]

[1] Brigham Young University

smb9564@byu.edu

Introduction

Three very important dimensions of exoplanet demographics are mass, radius, and period (MRP). These three parameters largely define a planet's composition, structure, and even habitability.

One challenge to producing a MRP distribution has been finding a large, homogeneous sample that includes information on all 3 parameters.

We use a new, homogeneous catalog of Kepler multiplanet systems with high-quality mass estimates to model the global MRP distribution of close-in planets around FGK stars.

Our model is a non-parametric grid; it's a 3D histogram with no parameters imposed on the MRP distributions. We solve for the occurrence rate of planets per star per a particular bin ("voxel") of MRP space : $\mathcal{R}_{MRP}$.

Data

We use the new extensive Kepler Multis Dynamical Catalog (KMDC) (Jones et al., in prep.).

KMDC facts:

- composed of >90% of all Kepler multiplanet systems
- 661 multiplanet systems, 1665 planets, 1000 posterior draws per planet
- homogeneously derived using photodynamical modeling (essentially the best possible mechanism for measuring masses in Kepler)
- major limitation: the majority of planets have no meaningful mass measurement (Kepler could not detect planet-planet interactions)
- ≳100 planets with <25% error in mass
- KMDC density priors of 0.01-30 g/cm^3, (our model imposes 10 g/cm^3 as our upper limit)
- coming soon to the NASA Exoplanet Archive!

We also include single-planet systems taken from the last Kepler data release, DR25. These planets (1428 in total) are given a uniform density prior from 0.01-10 g/cm^3, representing a generous range of possible exoplanet densities.

Animations of mass/radius/period slices of the full 3-dimensional MRP occurrence-rate models.

Model & Fitting

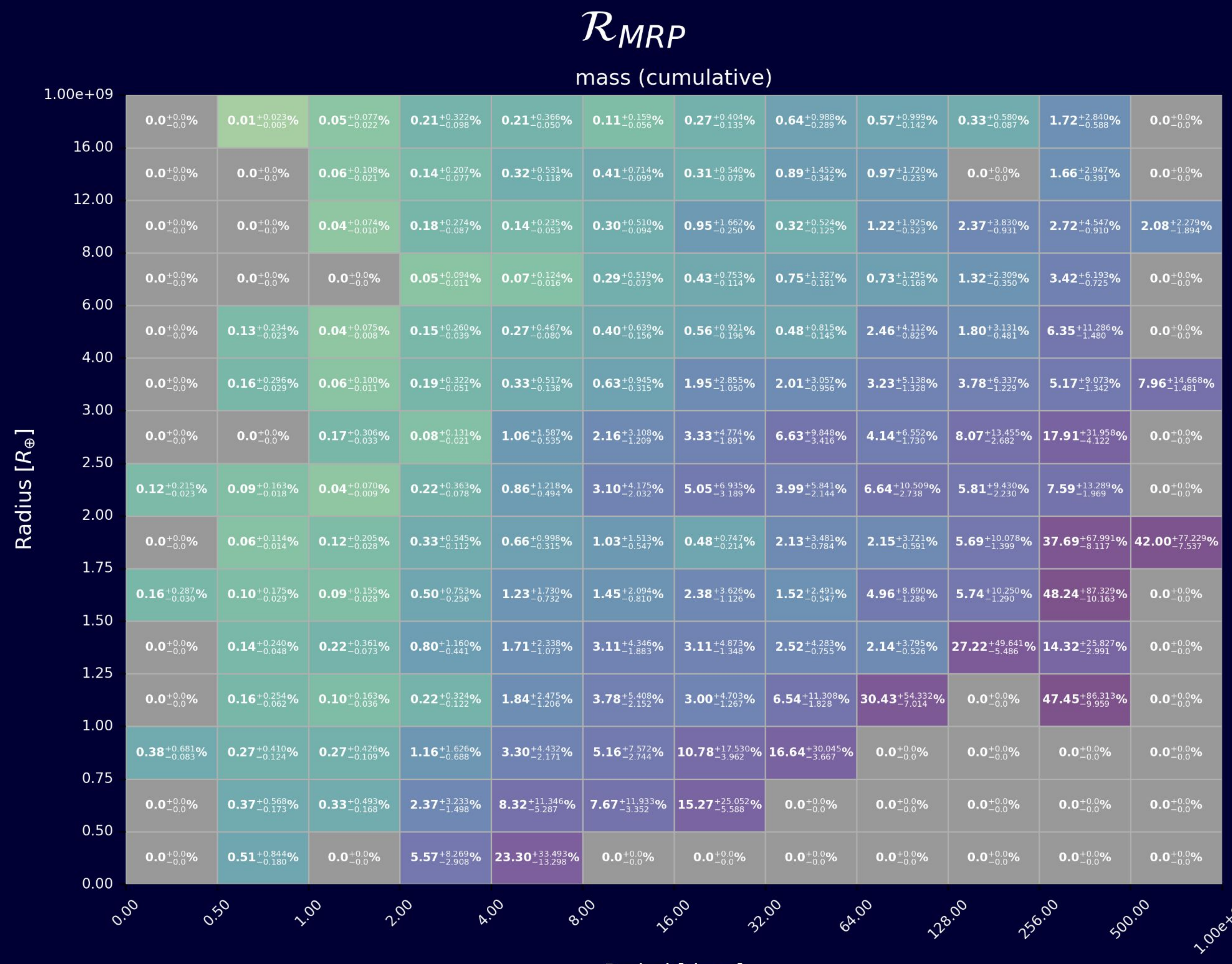
We are seeking $\mathcal{R}_{MRP}$, the occurrence rate of planets within a particular mass-radius-period range.

We use the completeness inferred from Hsu et al. 2019, which performed a detailed modeling of Kepler's ability to detect planets (including geometric, detection, and vetting efficiency) per radius-period voxel. We assume this completeness is independent of mass (introducing a small systematic uncertainty).

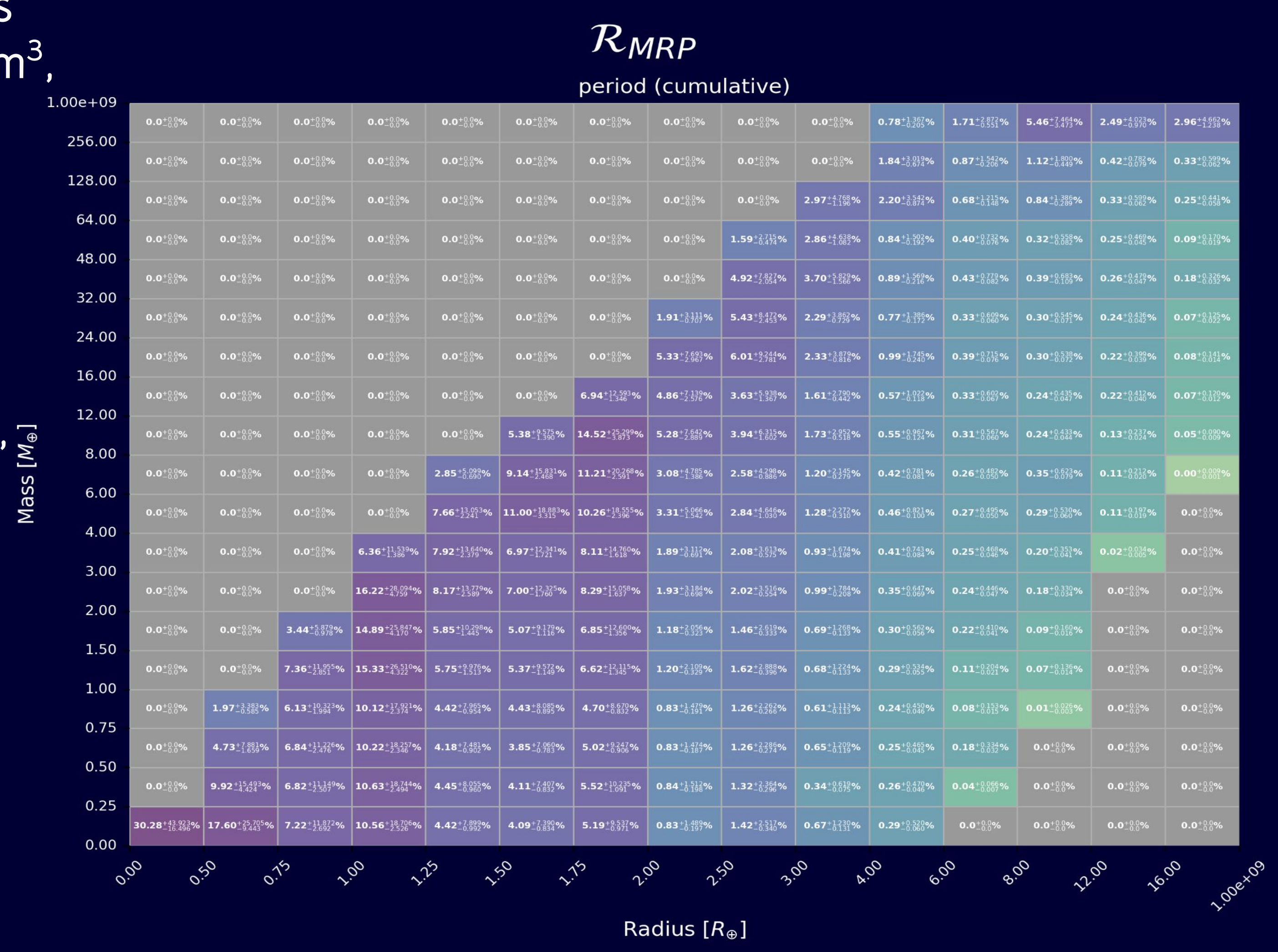
For each individual voxel, we can then compare the model prediction of the number of detected planets ($\mathcal{M}_{MRP}$ = $\mathcal{R}_{MRP}$ × **completeness**) to the observed number inferred from the KMDC posteriors ($\mathcal{D}_{MRP}$) using the following Poisson log-likelihood:

$$\log(\mathcal{L}_{MRP}) = \log(\mathbf{Pois}(\mathcal{D}_{MRP} \mid \mathcal{M}_{MRP}))$$

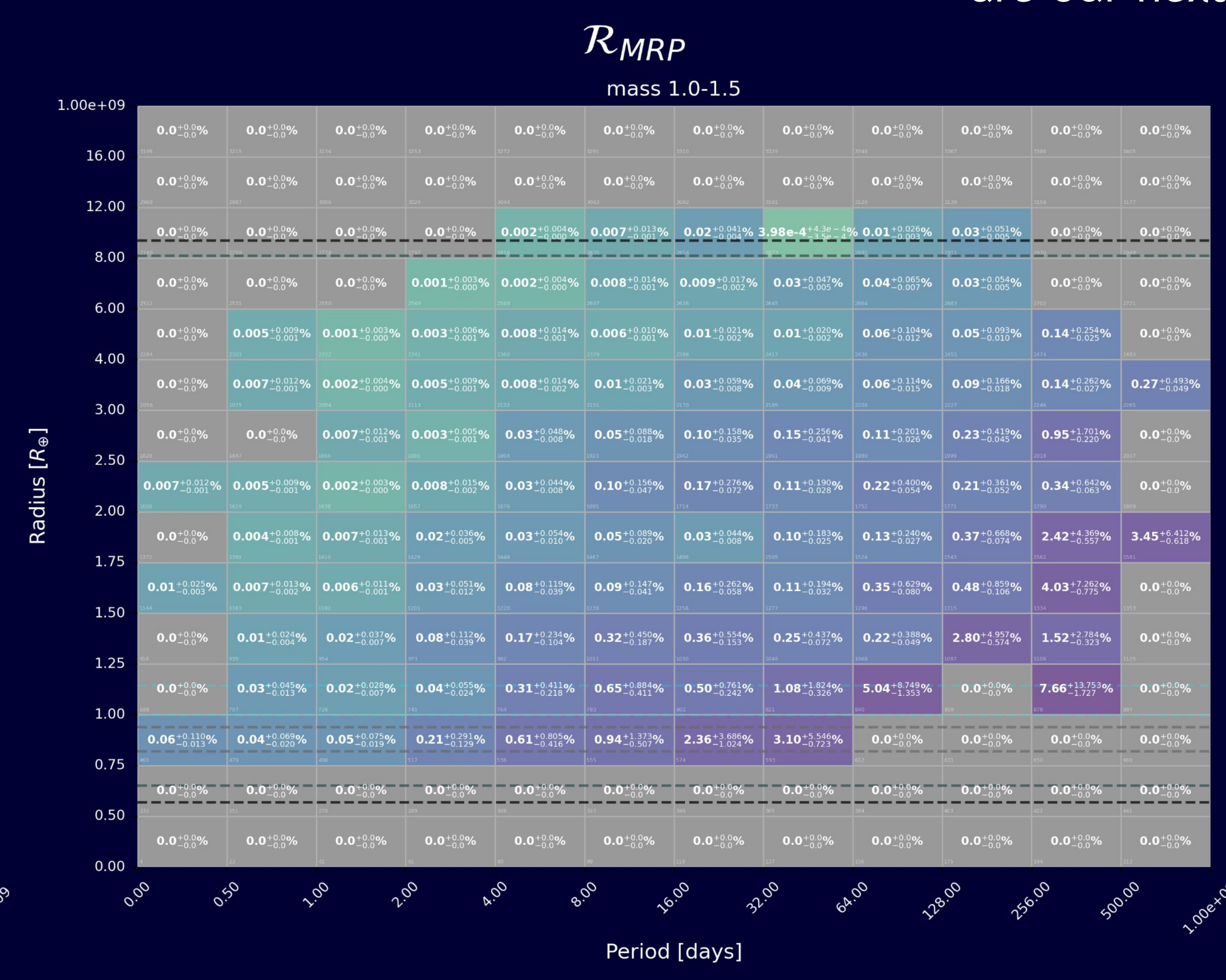
We treat each voxel independently and use `emcee` to sample the posterior for $\mathcal{R}_{MRP}$ using this likelihood. We report the median and uncertainties based on the 16th to 84th percentiles for each voxel.



Modeled occurrence rates of planets per star per radius-period bins summed over all masses. As expected, planet frequency increases with period and decreases with radius.



Mass-radius occurrence rates summed over period. The non-parametric model provides little mass information, despite the detailed photodynamical modeling in the KMDC.



The 1-1.5 M$_⊕$ slice of the radius-period distribution. Horizontal dashed lines show various densities relevant to the priors.

Results

Consistent with previous studies, the incidence rate of planets per star increases as the planet becomes smaller and more distant from the star.

Using a broad definition of Earth-like planets (1.0-1.5 R$_⊕$, 256-500 days, 0.75 -2.0 M$_⊕$), we find that the percentage of FGK stars with a rocky Earth ("eta-Earth-rocky" or $\eta_{⊕,rocky}$) is:

$$\eta_{⊕,rocky} = 12.92^{+23.54}_{-2.499}\ \%$$

The lack of constrained mass data and small number statistics within individual voxels seriously hamper non-parametric modeling efforts. Without parameterization, the many unconstrained systems overwhelm the information of the constrained systems—the differences between uniform density and PhoDyMM-derived density models are slight (<1σ) at best.

A more parameterized, constraining technique is thus necessary to utilize the KMDC's mass information.

Future Work

A hierarchical Bayesian model (HBM) will maximize the information from the best-constrained systems while incorporating clues from known power-law distributions and planetary formation theory (as in Neil & Rogers 2020). This approach also allows distinct subpopulations (giants, cores, terrestrials, etc.) to be modeled.

Adding further dimensions to this HBM could also increase its explanatory power. Eccentricity and argument of periapsis (both present in the KMDC) are our next steps.

References

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

Foreman-Mackey, D., Hogg, D. W., & Morton, T. D. 2014,ApJ, 795, 64, doi: 10.1088/0004-637X/795/1/64

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

Hsu, D. C., Ford, E. B., Ragozzine, D., & Morehead, R. C. 2018, AJ, 155, 205, doi: 10.3847/1538-3881/aab9a8

Kane, M., Ragozzine, D., Flowers, X., et al. 2019, AJ, 157, 171, doi: 10.3847/1538-3881/ab0d91

Jones, D., Ragozzine, D., Fabrycky, D. C. in preparation

Lissauer, J. J., Rowe, J. F., Jontof-Hutter, D., et al. 2024, PSJ, 5, 152, doi: 10.3847/PSJ/ad0e6e

Neil, A. R., & Rogers, L. A. 2020, ApJ, 891, 12, doi: 10.3847/1538-4357/ab6a92