McGill Astrophysical Journal

Discussing the Transit Planet M-R Relation and Observational Bias at Small Radii

K. Sienko, M. Walkington, M. Askew

McGill University Department of Physics

April 17th, 2023

Abstract

Transit light curves are one of the most common methods for exoplanet discovery. However, due to its nature, this method has an observational bias against planets of small radii due to a shallower dip in the flux of its host star during transit. After plotting the M-R relation of transit-discovered exoplanets and quantifying it through MCMC fitting, the relationship was found to be particularly consistent with a power law for planets of radii >1.6 $M_{\oplus}$, which follows other literature theories. More specifically, its logarithmic plot was found to obey the linear relation $y = 1.63x + 0.72$. The observation bias was then examined through simulated transit light curves of planets of relatively small radii by taking into account the limiting magnitude of the telescope. The conclusion drawn from these simulations was that the bias could affect the M-R relation at smaller radii <1.6 $M_{\oplus}$.

1 Introduction

The study of exoplanets has revolutionized our understanding of the universe and the potential for life beyond our own solar system. With the discovery of thousands of exoplanets, one of the most important tasks is to accurately measure their physical properties, such as their mass and radius. One relation that has been thoroughly studied is the M-R relation of lower radii exoplanets. By empirical observations, the relationship between the planet's mass and radius has been described by a power law, where the mass increases with radius like $M \propto R^{1.81}$ [1]. This relation has helped infer the mass and radius of many exoplanets that haven't been measured using other techniques. However, the measurement of these properties is not always straightforward, and observation bias can significantly impact the accuracy of these measurements.

One source of observation bias occurs in the detection of small planets using transit light curves, which is one of the most common methods for detecting exoplanets. Notable surveys where exoplanets were detected by transit include Kepler and TESS (Transiting Exoplanet Survey Satellite). Transit light curves occur when a planet passes in front of its host star, causing a small dip in the star's brightness that can be detected by telescopes. A varying planet-radii-to-stellar-radii ratio can change the size of the dip in flux (see Appendix A), thus, from these transit light curves, the radius of a planet can be measured. The mass of a planet is obtained independently through other methods such as the radial velocity method. A smaller planet-radius-to-stellar-radius ratio produces a shallower transit light curve, therefore the detection of small planets using the transit method can be challenging and may be obscured by noise or other astrophysical effects.

The prevention of detection of smaller planets due to noise may lead to bias for the M-R relationship at smaller scales. Inaccurate measurements can also lead to incorrect assumptions about the planet's properties and limit our ability to draw meaningful conclusions about the nature of exoplanets.

In this paper, we perform an MCMC fitting to planets that were detected by transit from the Kepler and TESS surveys and present an analysis of the observation bias that could potentially arise due to noise for small planets using simulated transit light curves.

2 Methodology

2.1 Data Acquisition

The NASA Exoplanet Archive [2], a comprehensive database containing information for over 5300 exoplanets, was utilized in this study. Several cuts were applied to narrow down the data set and facilitate analysis. Initially, planets orbiting non-main sequence stars or stars that were exceedingly bright (those not of class F, G, K, or M) were excluded, along with binary systems. Furthermore, only exoplanets detected using the transit method by the Kepler, Kepler 2 (K2), and Transiting Exoplanet Survey Satellite (TESS) missions were selected. Exoplanets whose masses were estimated using the M-R relationship were also removed to avoid bias for the fits. Additionally, restrictions were implemented on the planet's radii and mass. Planets with $R > 8.0R_{\oplus}$ were removed, as this study is focused on the M-R relationship of small planets, and studies show that large planets have little to no dependence between radius and mass [1]. Planets with $M > 70M_{\oplus}$ were also removed. Finally, data without well-defined errors on planet radius, mass, and stellar radius were excluded. In total, a sample of 169 exoplanets fitting these conditions was used in this analysis.

2.2 MCMC Fitting

From our resulting database of exoplanets, the assumption that the exoplanet M-R relationship fits a power law was made. For simplicity, the logarithm of both the planetary radius and mass was taken with the goal of obtaining a linear fit. A linear model provides a simple and intuitive way to describe the relationship between two variables such as mass and radius, allowing for easy interpretation and comparison of results. While other modelling approaches may be used depending on the specific research question, the use of a linear model is often a good starting point for analyzing the radius/mass relation of exoplanets. To fit this data, the emcee Python package was used which is an ensemble sampler for the Markov chain Monte Carlo (MCMC) method.

An MCMC analysis was performed to study the posterior $p(a, b|\{y_i\}) \propto p(\{y_i\}|a, b)p(a, b)$ for the two parameters of a linear model $y_i = ax_i + b$. With the assumption of a uniform

prior on a and b as well as gaussian error bars, our likelihood takes the form of a multivariate gaussian distribution that assumes a covariance of zero as the methods of determining the planets' radii and masses are independent of each other:

$$p(\{y_i\}|a,b) = \prod_{i=1}^{n} \frac{\exp\left[-\frac{[y_i - ax_i - b]^2}{2\sigma_{y_i}^2} - \frac{[x_i - \bar{x}]^2}{2\sigma_{x_i}^2}\right]}{\sqrt{2\pi\sigma_{y_i}\sigma_{x_i}}}. \quad (1)$$

For the MCMC fitting, 200 chains of 5000 iterations were run for the exoplanet radius vs exoplanet mass model. By inspection using trace plots, the first 100 iterations were discarded to eliminate the "burn-in" phase in order to allow the chain to reach its equilibrium distribution. The chain was also thinned out by taking only one out of every 15 steps. The resulting posteriors of a and b were visualized using the corner Python package, which made a 2D contour plot. Detailed justifications for the choices made in the analysis can be found in the accompanying Jupyter notebook titled "ExoplanetsLab.ipynb".

2.3 Transit Light Curve Simulations

The limitations of the plotted data can be examined by simulating the transit curves for hypothetical small planets with radii less than that of the smallest detected planet of each host star. Using the Python package Bad-Ass Transit Model cAlculatioN (BATMAN) [3], transit curves with varying radius ratios of $R_{\text{planet}}/R_\star$ were simulated.

The limiting magnitude of a telescope, defined as the faintest apparent magnitude that the telescope would be able to detect, was compared to the apparent magnitude of the percentage dip in these transit curves. For the Kepler/K2 missions (which both use the Kepler space telescope), while most detected stars are between apparent magnitudes of 14-16, the telescope has a limiting magnitude of around 22 [4]. For TESS, the limiting magnitude for deep field missions is about 18 [5]. These values correspond to the "noise level" on the corresponding transit curve. Using this noise level, and comparing it to the percentage of flux blocked by the transit, the minimum radii for a planet to be detected at a specific host star by a particular telescope can be determined. Then, the existence of an observational bias can be discussed. In other words, if the dip in the transit curve caused by a planet of small radius is smaller

than the noise level (corresponding to the limiting magnitude of the telescope), then the dip is indistinguishable from noise and thus the existence of the planet remains elusive.

The relative flux of the stars to solar flux within our sample was found by first using the luminosity-stellar mass relation for main sequence stars:

$$\frac{L_\star}{L_\odot} \approx (\frac{M_\star}{M_\odot})^{3.5}. \tag{2}$$

From here we calculate the relative flux of the star to solar flux as,

$$\frac{F_\star}{F_\odot} = \frac{L_\star / L_\odot}{4\pi d_\star^2}, \tag{3}$$

where $d_\star$ is its distance from Earth. From the values of $\frac{F_\star}{F_\odot}$, the apparent magnitude of the star, labeled $m_{\text{app}_\star}$, was calculated using the following equation:

$$m_{\text{app}_\star} = -2.5 * log_{10}(\frac{F_\star}{F_\odot}) + m_{\text{app}_\odot}, \tag{4}$$

where $m_{\text{app}_\star}$ is the resulting apparent magnitude and $m_{\text{app}_\odot} = -26.74$ is the apparent magnitude of the Sun.

With the apparent magnitude of the star and the limiting magnitude of the telescope, denoted as m_{lim} the amplitude of noise as a relative flux F_{noise} was calculated:

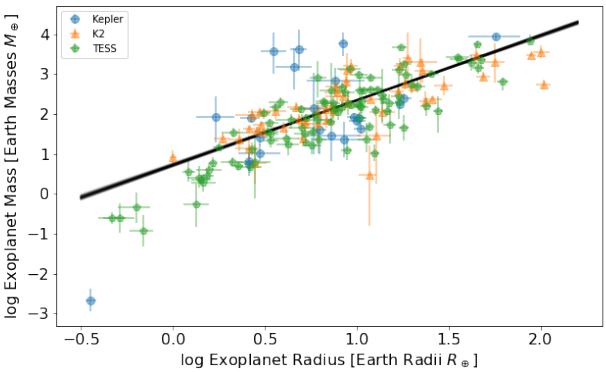
$$F_{\text{noise}} = 10^{\frac{m_{\text{app}_\star} - m_{\text{lim}}}{2.5}}. \tag{5}$$

The noise level amplitude as a relative flux can be plotted alongside the transit light curve as a horizontal line, and the radius of the planet can be varied to quantify what threshold of R_{planet} is detectable from a certain star with a certain telescope.

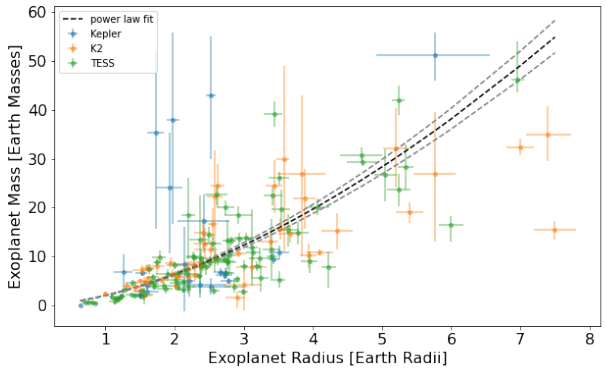
3 Results

3.1 Model Fit

Figure 1a shows the best-fit linear model for the logarithmic M-R relation of exoplanets. Table 1 summarizes the posteriors for the a and b parameters used to model the fit, and Figure 2 visualizes them. Our final result displays a best-fit line of the form $y = 1.63x + 0.72$



(a) Linear regression model for logarithmic exoplanet M-R relation. The partial transparency of the best-fit lines shows where the prediction is very tight (intense color) and where the prediction is quite uncertain (faint color).



(b) Power law fit for exoplanet M-R relation, including uncertainties in the parameters the power law is of the form $y = Ce^a$ where $C = e^b$ and a and b are the parameters found from the MCMC fit.

Figure 1: M-R relationship plotted on different scales.

for our log-log M-R relation. Using these values, we can visualize the corresponding power law model for the M-R relation, as shown in Figure 1b.

Table 1: The two resulting parameter values from the linear model posteriors, along with the values of the 16th and 84th percentiles to illustrate the 68% credibility region.

Parameter	16%	Median	84%
a	1.61	1.63	1.65
b	0.70	0.72	0.74

3.2 Transit Curve

Simulated transit curves can be made for each telescope (each of which has a particular limiting magnitude) and each star (each of which has a particular mass and radii $M_\star$ and $R_\star$) for which we can vary the radius of a planet to find where the threshold of detectable planet radii lies.

As a demonstration, Figure 3 is the simulated transit curve for the planet KOI-142 b, which orbits its host star KOI-142, and was detected by the Kepler space telescope. With a limiting magnitude of 22, the noise level amplitude of this telescope observing KOI-142 (a

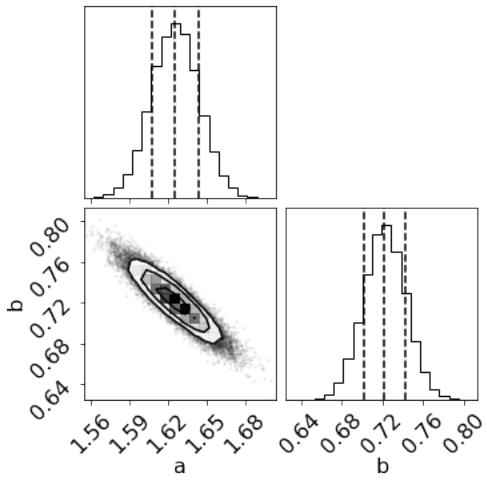


Figure 2: Posterior distributions from the two-parameter M-R relation. Contours show the 68% and 95% credibility region, while the dashed lines represent the 68% credibility region with the 16th, 50th, and 84th quantiles.

star of apparent magnitude 12.74) is ≈ 0.9998. The maximum dip of planet KOI-142 b is ≈ 0.9988, which is a difference of 0.001 percent flux from the noise amplitude (making this planet easily detectable by the telescope). Had the planet been smaller, at $\approx 40\%$ of its true radius, its transit curve would be in line with the noise amplitude, making it virtually undetectable.

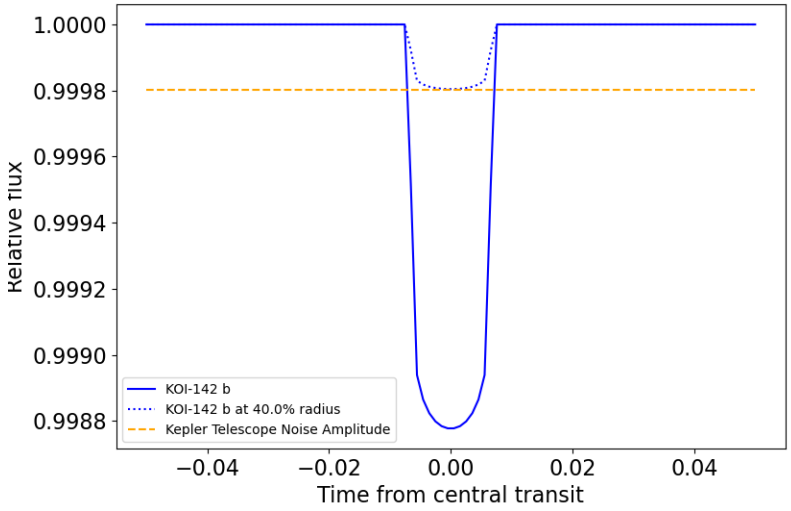


Figure 3: Transit light curves for KOI-142 b (solid blue) and a planet with 40% the radius of KOI-142 (dashed blue), and the noise level amplitude for the Kepler telescope (dashed orange).

4 Discussion

4.1 M-R Relation

Consistent with our initial assumption that planet mass and radius can be related by a power law for radii <8 $R_\oplus$, the MCMC linear fitting in Figure 1a suggests that the logarithm of the mass and radius is well fit by a line, as most of the error bars of the data include the representative fit lines. However, our a parameter, responsible for the exponent in the power law of 1.63 ± 0.02, is not consistent with the exponent value from the regime $M \propto R^{1.81}$. We can see from Figure 1 that planets with radii <1.6 $R_\oplus$ start to deviate from the best-fit line. This could indicate that before 1.6 $R_\oplus$, the M-R relation follows a different power law or the empirical assumption that a power law describes this relationship is not entirely correct. However, it is also highly possible that since for this MCMC fitting we only considered two parameters, we may need more parameters dependent on other planetary parameters to accurately fit the planet mass and radius data.

Since the number of planets with radii <1.6 $R_\oplus$ is outnumbered by planets that have greater radii, the MCMC is not affected by the lower radii planets and is well fit to the planets with greater radii.

4.2 Missing Planets

The fact that there are a lower number of planets with radii <1.6 $R_\oplus$ that appear to deviate from the line of best fit may indicate that we do have some bias on our fit. From the simulated transit curve in Figure 3, we see for example that for a planet that consists of the same planetary properties as KOI-142 b but with only 40% the radius, the planet would be undetectable as its transit light curve is comparable with the noise level amplitude. With a radius of about 3.438 for KOI-142, 40% of that is $\approx$ 1.375. This value falls below 1.6 $R_\oplus$, where the deviation begins to occur. With noise corrections that could lead to this hypothetical planet's detection, one could also find the planet's mass and then see where it lies in the M-R relation. Missing planets with radii <1.6 $R_\oplus$ could shed light on possible discrepancies from the empirically found power-law regime.

4.3 Limitations with BATMAN

For the purposes of this paper, the goal was to show that there may have been observational bias on our representative fits since it's harder to detect lower radii exoplanets. As discussed above, these lower radii planets ($<8\ R_{\oplus}$) stray away from the fit, meaning we may not be seeing the whole picture. With BATMAN, we only varied the planetary radius to show how noise could prevent the detection of lower radii planets, however, there are other parameters that were unused which required more information on stellar properties such as limb-darkening. By studying a host star's stellar properties and with the use of these other parameters, we would be able to come to further conclusions on how noise may affect the ability to detect exoplanets.

5 Conclusions

In conclusion, one of the most common methods for exoplanet discovery, the study of transiting light curves, is heavily biased towards planets with larger radii due to the larger deviations in host star flux that they cause during transit. After quantifying the M-R relationship of transit-discovered exoplanets, the logarithm of the relationship was found to obey the linear relation $y = 1.63x + 0.72$ most consistently for planets of radii $>1.6M_{\oplus}$. By taking into account the limiting magnitude of the telescope, observation bias was then examined through simulated transit light curves of planets of relatively small radii and was found to affect the M-R relation at lower-scale radii.

Acknowledgements

The authors of this paper would like to thank Professor Adrian Liu for his outstanding help in the execution of this project.

References

[1] Bashi, Dolev, Helled, Ravit, Zucker, Shay, and Mordasini, Christoph, "Two empirical regimes of the planetary mass-radius relation," *A&A*, vol. 604, p. A83, 2017. [Online]. Available: https://doi.org/10.1051/0004-6361/201629922 1, 2

[2] "Nasa exoplanet archive." [Online]. Available: https://exoplanetarchive.ipac.caltech.edu/index.html 2

[3] "Batman: Bad-ass transit model calculation." [Online]. Available: http://lkreidberg.github.io/batman/docs/html/index.html 3

[4] J. J. Lissauer, R. I. Dawson, and S. Tremaine, "Advances in exoplanet science from kepler," *Nature*, vol. 513, no. 7518, p. 336–344, 2014. 3

[5] G. R. Ricker, "Transiting exoplanet survey satellite (tess)," *SPIE Newsroom*, 2014. 3

A Appendix 1- Simulated Transit Curves

Figure 4 shows the transit curve simulations of several theoretical planetary radii of increasing size. This demonstrates how the relative stellar flux decreases significantly as the planet's size increases. The transit curves are also affected by several other parameters relating to the planet's orbit, such as its period, semi-major axis, inclination, eccentricity, and longitude of periastron. As these are theoretical systems, for the purposes of this simulation, as well as the rest of the simulations in this paper, all parameters are kept constant except for the planetary radius relative to its host star's radius.

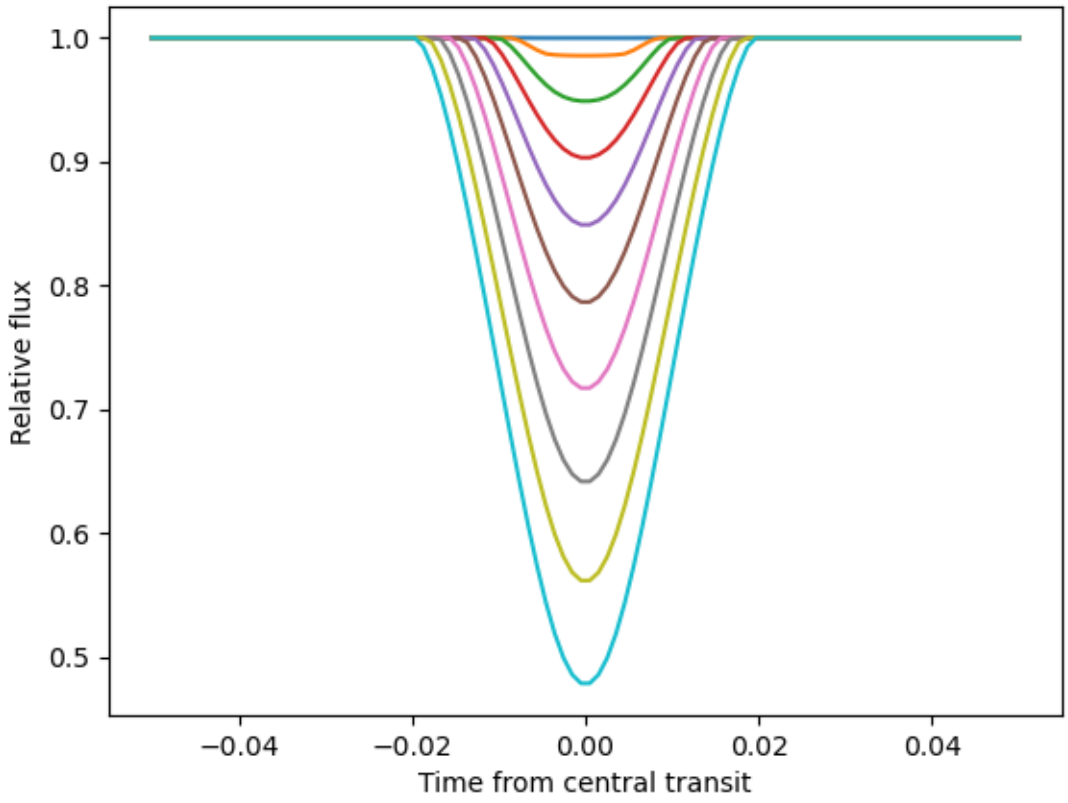


Figure 4: Transit light curves several theoretical planetary radii (varying from 0.01 to 1 stellar radii).