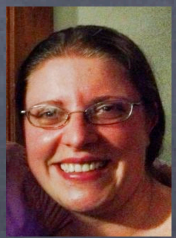


Just How Earth-like are Extrasolar Super-Earths?

Constraints on H+He Envelope Mass Fractions from *Kepler*'s Planet Candidates

Angie Wolfgang & Eric Lopez, UC Santa Cruz

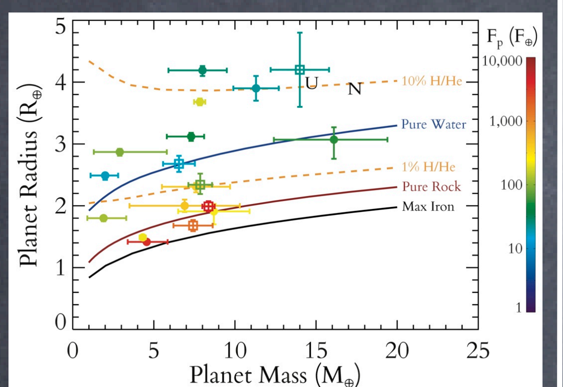
With 3500 planetary candidates discovered in its first 3 years of data, the Kepler Mission promises to answer one of the most fundamental questions posed in exoplanetary research: what kinds of planets occur most often in our Galaxy? As Kepler primarily yields planetary radii and orbital periods, it has enabled numerous studies of the occurrence rate of planets as a function of these variables. Unfortunately, the full mass distribution, and thus a direct measure of these planets' possible compositions, remains elusive due to the unsuitability of these faint targets for radial velocity follow-up and the relative rareness of transit timing variations. We show, however, that relatively straightforward models of planetary evolution in an irradiated environment can make some progress without this full mass distribution towards understanding bulk compositions of the abundant Super-Earth/Sub-Neptunes that Kepler has discovered. In particular, we constrain the distribution of envelope fractions, i.e. the fraction of a planet's mass that is in a gaseous hydrogen and helium envelope around its rocky core, for this exoplanet population that has no analogs in our Solar System.

This research builds on collaborations between astronomers and statisticians forged during a three week workshop on "Modern Statistical and Computational Methods for Analysis of Kepler Data" at SAMSI in June 2013.

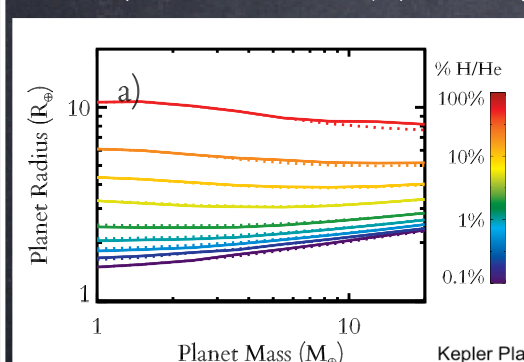
The Concept

For a given extrasolar planet, measurements of both mass and radius can begin to constrain its bulk composition. This is especially interesting for planets with masses between Earth and Neptune, as the transition from primarily rocky to gaseous compositions lies somewhere in this mass range, which has no Solar System analogs. For this work, we consider the fraction of a planet's mass that is in a gaseous hydrogen and helium envelope around a rocky core of Earth-like composition.

To date, less than 2% of all known planets with radii between Earth and Neptune, most of which are Kepler planetary candidates, have measured masses, as these measurements require a significant amount of telescope time and occur at the detection limit of the radial velocity technique. Transit timing variations can also yield masses, but require the planets to be in strong orbital resonances.

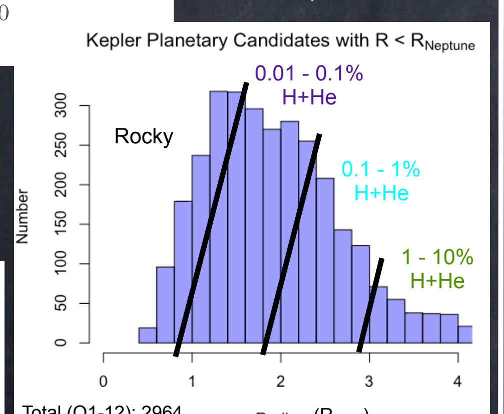


Fortunately, recent advances in modeling the thermal evolution of these super-Earth/sub-Neptunes' internal structures provides an opportunity to constrain these planets' compositions without masses (Lopez & Fortney, 2013; see Monday talk):



The weak dependence of these model radii on mass suggests that mass measurements are not necessarily required to get a sense for a planet's H+He envelope mass fraction. Therefore, when we apply these models to Kepler's planetary candidate radii, we can constrain the distribution of these super-Earths' bulk compositions.

For the first time we can predict the range of possible gas mass fractions for super-Earth/sub-Neptunes as a function of their size.



References:

Lopez, E.D., & Fortney, J.J. 2013, submitted.

Lopez, E.D., Fortney, J.J., & Miller, N. 2012, ApJ, 761, 59.

Wolfgang, A. & Laughlin, G. 2012, ApJ, 750, 148.

Methods

Because we aim to constrain the distribution of gaseous envelope mass fractions, our approach must easily enable calculation of uncertainties on this distribution. Bayesian statistical analyses naturally facilitate this. Furthermore, we need an analysis method which best accounts for the fact that this compositional distribution is not directly measured and incorporates the observational uncertainty on those quantities that are. Accordingly, we adopt the following approach:

Hierarchical Bayesian Modeling

If D = data and Θ = the parameters of a model to fit to data, Bayes' Theorem states that the posterior, i.e. the probability of the parameters given the data, is proportional to the likelihood of the data given certain values for the model parameters (assuming the model itself is correct) and the prior probability of those parameters:

$$P(\Theta|D) \propto P(D|\Theta)\, P(\Theta)$$

Hierarchical models extend Bayes' Theorem to multiple tiers of prior distributions. If the parameters we're trying to fit to the data have a range of possible values allowed by nature, then the parameters themselves have a distribution that can be fit with a model that contains its own parameters, called hyperparameters (α):

$$P(\Theta|D) \propto P(D|\Theta,\alpha)\, P(\Theta|\alpha)\, P(\alpha)$$

This hierarchical Bayesian model is represented by:

$$D|\Theta,\alpha \sim P(D|\Theta,\alpha) = \text{likelihood}$$

$$\Theta|\alpha \sim P(\Theta|\alpha) = \text{prior on the parameters}$$

$$\alpha \sim P(\alpha) = \text{prior on the hyperparameters}$$

where ~ means "drawn from the distribution". Our hierarchical model is as follows:

$$R_{obs,i} \,|\, \sigma_{obs}, M_{true,i}, f_{env,i}, p_i, \alpha_M, \alpha_f, k \sim \text{Normal}(\, R_{true,i} = g(M_{true,i}, f_{env,i}, p_i), \sigma_{obs} \,|\, \alpha_M, \alpha_f, k\,)$$

$$f_{env,i} \,|\, \alpha_f, k, p_i, M_{true,i} \sim \text{Beta}(\, \alpha_f, \beta_f = h(\alpha_f, k, p_i, M_{true,i})\,)$$

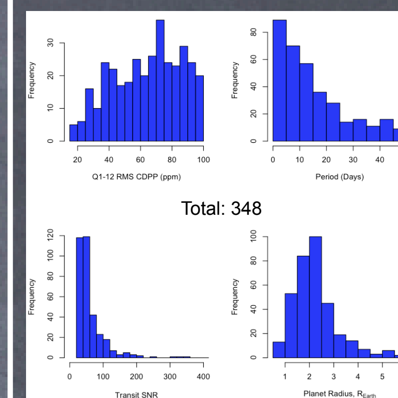
$$M_{true,i} \,|\, \alpha_M \sim \text{Pareto}(\, \alpha_M, 1\,)$$

$$\alpha_M \sim \text{Uniform}(0, 1) \qquad \alpha_f \sim \text{Uniform}(0.1, 10) \qquad k \sim \text{Uniform}(1, 50)$$

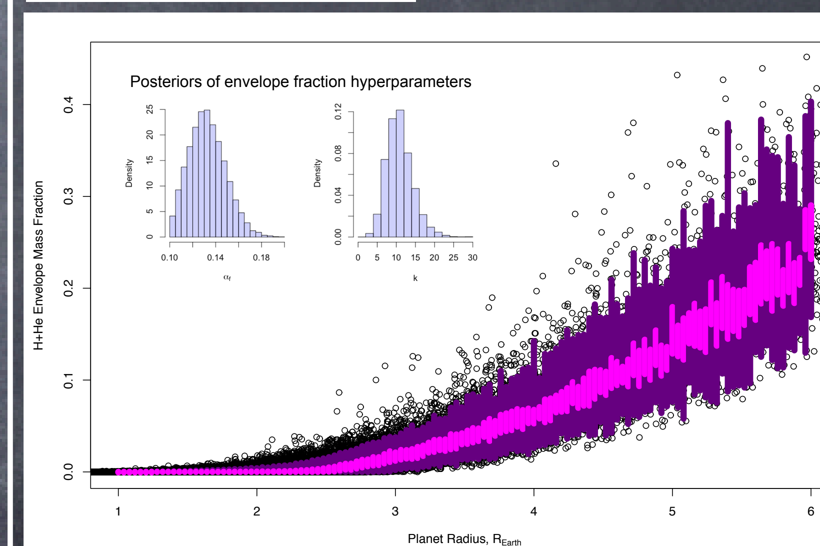
where $\mathbf{R_{obs,i}}$ is the observed radius of each planet and $\mathbf{p_i}$ is the observed period of each planet, which together compose our data. The measurement error on period is very low (< 0.01%), so we don't incorporate any observational uncertainty for it; however, this is not the case for planet radius, so we construct a likelihood for the radius data based on the planet's true radius given by the structural models of Lopez 2013 ($\mathbf{g(M_{true,i}, f_{env,i}, p_i)}$) and Gaussian noise with typical measurement uncertainty $\mathbf{\sigma_{obs}}$. $\mathbf{f_{env,i}}$ is the envelope fraction of the planet, and $\mathbf{\alpha_f}$ and $\mathbf{\beta_f}$ are the hyperparameters on this distribution; $\mathbf{\alpha_f}$, which controls the variance in this parameterization, is allowed to vary freely to be constrained by the data, while $\mathbf{\beta_f}$ is a function of $\mathbf{\alpha_f}$ and another parameter **k** which controls the mean of the distribution and is informed by the observational density threshold unearthed in Lopez, Fortney, & Miller, 2012 as a function of mass and period. Finally, the planetary mass, $\mathbf{M_{true,i}}$, is modeled as a power law distribution, with corresponding hyperparameter $\mathbf{\alpha_M}$.

We use the R package JAGS (Just Another Gibbs Sampler) to perform Markov Chain Monte Carlo sampling and arrive at the posteriors numerically.

Preliminary Results

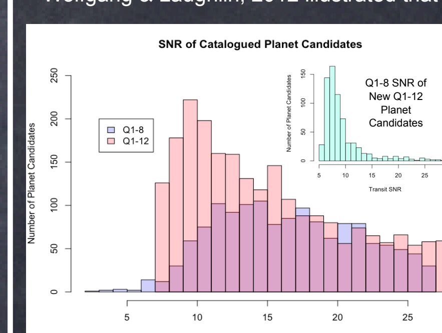


A subsample of the ~ 3000 sub-Neptune-sized Q1-12 planet candidates (left) was chosen for preliminary tests of our hierarchical model. This sample consists of the highest fidelity small planet candidates, with transit SNR > 30, P < 50 days, and the lowest existing stellar noise estimates (CDPP < 100 ppm). The completeness of the detected planet candidates in this subsample is also expected to be high. Below is the preliminary results of our MCMC, including samples from the hyperparameter posteriors (inset) to calculate the 1-σ (pink) and 2-σ (purple) spread in the population's envelope mass fraction as a function of radius.



Future Work: Accounting for Detection Biases

Wolfgang & Laughlin, 2012 illustrated that carefully accounting for Kepler's selection and detection biases is absolutely necessary for any accurate statistical analyses, so we must incorporate non-detections probabilistically as functions of period, radius, and host star into our hierarchical model.



However, these completeness corrections are not always straightforward to model, as the distribution of observed planetary properties change in complicated ways with improvements to Kepler's planet detection algorithms.