How Rocky Are They?

The Composition Distribution of *Kepler*'s Sub-Neptune Planet Candidates (within 0.15 AU)

Angie Wolfgang, NSF Graduate Research Fellow
SAMSI Bayesian Exoplanet Populations Group, Eric Lopez, Gregory Laughlin
Planets with no Solar System analogs . . . what possible compositions?
From just Mass and Radius:

Have a wide range of possible compositions... what does the distribution look like?
Interpreting \((M,R)\)

Inferring a composition requires modeling these planets’ internal structures:

Continuity: \[ \frac{dm(r)}{dr} = 4\pi r^2 \rho(r) \]

Hydrostatic Equilibrium: \[ \frac{dP(r)}{dr} = \frac{-Gm(r)\rho(r)}{r^2} \]

Equation of State: \[ P(r) = f(\rho(r), T(r)) \]

Intrinsic luminosity is set to a constant value on a grid … but how to determine an astrophysically appropriate value?

Rogers et al. 2011
But a planet cools . . .

\[ \int_{M_{\text{core}}}^{M_p} \frac{dm}{dt} \frac{T \, dS}{dS} = -L_{\text{int}} + L_{\text{radio}} - c_v M_{\text{core}} \frac{dT_{\text{core}}}{dt} \]

. . . at different rates depending on its mass!

Low-mass planets, with low surface gravity, also cool fastest . . .
Flat Mass-Radius Relations!

Earth-composition rocky core, H+He envelope

Radius determined by composition, not mass!

Lopez & Fortney, 2014
Radius Proxy for Composition

Suggests a single mass-radius relationship may be misleading (little dependence on mass)

Fressin et al. 2013
Why is this exciting?

Can sidestep the need for expensive mass measurements!
(but need to assume rocky core with H+He envelope)

Radial Velocities:

Planets must be in resonances, need high data cadence & long time baselines

Requires years of observations for good phase coverage; majority of Kepler targets too faint

Marcy et al. 2014

Nesvorny et al. 2012
Errors dominated by uncertainty in stellar radius:

\[ \text{transit depth} \propto \left( \frac{R_{pl}}{R_*} \right)^2 \]
Well Suited to HBM!

“Regular” Bayes:

\[ p(\theta | y) = \frac{p(y | \theta) p(\theta)}{\int_{\Theta} p(y | \tilde{\theta}) p(\tilde{\theta}) \, d\tilde{\theta}} \]

= p(y) = constant

y = data

\( \theta \) = the parameters of a model that can produce the data

\( p() \) = probability density [distribution] of; \( | = “conditional on”, or “given” \)

\( p(\theta) \) = prior probability

(how probable are the possible values that you think \( \theta \) could take?)

\( p(y|\theta) \) = likelihood, or sampling distribution

(ties your model to the data probabilistically:

how likely is the data you observed given specific model parameters?)

\( p(\theta|y) \) = posterior probability

(a “new prior” distribution, updated with information contained by the data:

what is the distribution of \( \theta \) values given the data and your beliefs?)
Well Suited to HBM!

“Regular” Bayes:

\[ p(\theta|y) \propto p(y|\theta) \ p(\theta) \]

\( y \) = data
\( \theta \) = the parameters of a model that can produce the data
\( p() \) = probability density [distribution] of; \( | \) = “conditional on”, or “given”

\( p(\theta) \) = prior probability
(how probable are the possible values that you think \( \theta \) could take?)

\( p(y|\theta) \) = likelihood, or sampling distribution
(ties your model to the data probabilistically: how likely is the data you observed given specific model parameters?)

\( p(\theta|y) \) = posterior probability
(a “new prior” distribution, updated with information contained by the data: what is the distribution of \( \theta \) values given the data and your beliefs?)
HBM: What is it?

Hierarchical Bayes:

What if there isn’t just one “true” value of \( \theta \) for all the data? i.e. \( \theta \) has its own intrinsic distribution?

\[ p(\theta, \alpha | y) \propto p(y | \theta, \alpha) \ p(\theta | \alpha) \ p(\alpha) \]

can still use MCMC!

\( \alpha = \) hyperparameters
(the parameters that describe the distribution of \( \theta \) values)

\[ p(\theta | \alpha) = \text{this “intrinsic distribution” for the parameters} \]

\[ p(\alpha) = \text{prior probability of the hyperparameters} \]

Happens often for population studies: data tends to be “grouped” (hierarchical), and one value of \( \theta \) may not be appropriate for all groups

Adding another layer of probabilistic structure
Our Hierarchical Model:

- **Power Law**
  - \( \alpha \)
  - \( M_{\text{core},i} \)

- **Log-Normal**
  - \( \mu \)
  - \( \sigma \)
  - \( f_{\text{env},i} \)
  - \( F_i \)

- **Normal (i.e. \( \chi^2 \))**
  - \( R_{\star,i} \)
  - \( R_{\text{pl},i} \)
  - \( \sigma_{\delta,i} \)

- **Internal Structure Models**
  - \( \gamma \)

- **Parameters of the individuals**
  - \( \delta_i \)

- **Data**
  - \( N \)

We are interested in the parameters defining planet compositions.

Execute Metropolis-within-Gibbs MCMC (using JAGS in R)

But first, identify data!

\[
p(\theta, \alpha | X) = p\{R_{\text{pl},i}, M_{\text{core},i}, f_{\text{env},i}\}, \{\alpha, \mu, \sigma, \gamma\}|\delta_i, \sigma_{\delta,i}, F_i) \propto
\prod_{i=1}^{N} \left\{ p(\delta_i | \alpha, \sigma_{\delta,i}, R_{\text{pl},i}, R_{\star,i}, M_{\text{core},i}, f_{\text{env},i}, F_i, \alpha, \mu, \sigma, \gamma) \right\}
\times \prod_{i=1}^{N} \left\{ p(R_{\star,i})p(M_{\text{core},i}|\alpha)p(f_{\text{env},i}|\mu, \sigma) \right\}p(\alpha)p(\mu)p(\sigma)p(\gamma)
\]
Applying to *Kepler* planets

Extreme caution needed in interpreting the observed radius distribution!
All Q1–12 planet candidates
Q1–12 Sample selected to be complete
Fraction of all planet candidates in sample

Cuts:
P < 25 days
stellar noise < 100 ppm
Observed all 12 quarters
R\star < 1.2 R\small{Sun}

Now can run the MCMC to get some . . .
Results!!

First composition distribution:
~ 1% envelope mass fractions are the most likely
Wolfgang & Lopez, in prep.
How much to believe this?

Rhat = 1.08

Rhat = 1.00

Rhat = 1.36

Rhat = 1.03

Rhat ~ \frac{\text{Between-chain variance}}{\text{Within-chain variance}}

ACK!
How much to believe this?

This looks more promising

[Thinning factor set to 250 as $N_{\text{steps/chain}} = 500,000$]
How affects composition?

Marginalized over $\gamma$ (threshpowlaw)

$\gamma = 2$
One last sanity check

- $1.05 < \text{Gelman–Rubin Diagnostic} < 1.2$
- $\text{Gelman–Rubin Diagnostic} > 1.2$
Implications

“Radius as proxy for composition” works, with some spread (due to errors on radii).

Note: Kepler 10c not necessarily rocky: $f_{\text{env}} \sim 0.5\%$
Rocky-Gaseous Transition

% Mass in H+He

Incident Flux ($F_{\text{Earth}}$)

Radius ($R_{\text{Earth}}$)

A benefit of HBM: full posteriors on all the individual parameters, for free!!
An honest view of inference on individual planets from Kepler radii...
Mass-Radius Relationship

![Graph showing the relationship between planet radius and planet mass for various compositions.](image)
Mass-Radius PDF (probability density function)

Also some interesting trends with flux possible!
Summary of Implications

1) First composition distribution: ~1% envelope mass fractions are the most likely compositions for sub-Neptunes.

2) Radius uncertainties don’t destroy “radius is a proxy for composition”.

3) Need probabilistic treatment to “convert” radii into masses.

4) Rocky-gas transition around ~ 1.75 $R_{\text{Earth}}$ (agrees with Rogers et. al. 2014), and could be fuzzy.