Atmospheric Retrieval 101: Extracting Science from Spectra

Michael Line
Hubble Fellow Nasa Ames
Arizona State University (~1 month)

Outline

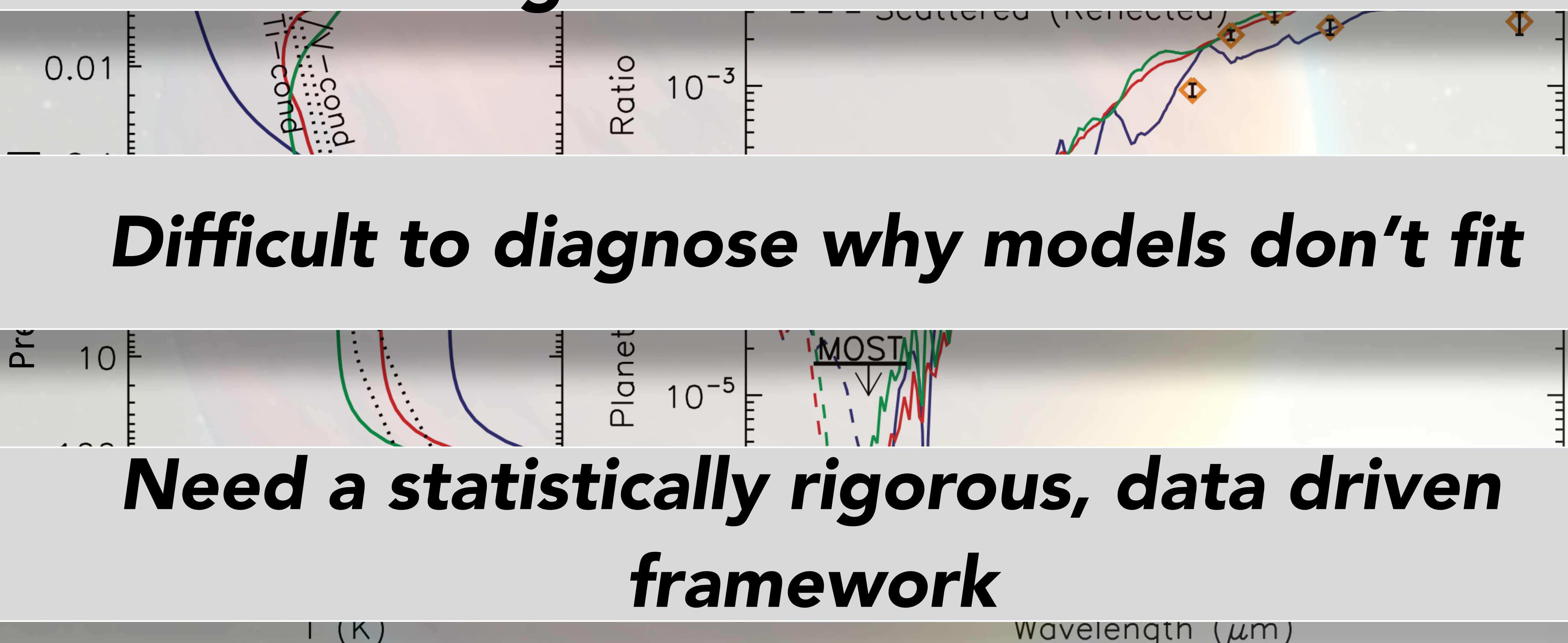
- **Retrieval Demystified**
- **Forward Models—the art of parameterizing complex atmospheres**
- **Bayesian Tools**
- **Some cool results**

Previous Method of Interpretation

No quantification of uncertainties-"eyeballed fits" resulting in model driven conclusions

Difficult to diagnose why models don't fit

Need a statistically rigorous, data driven framework



Fortney+2008

Retrieval in a Nutshell

The problem we are trying to solve is…

$$y = F(x) + \varepsilon$$

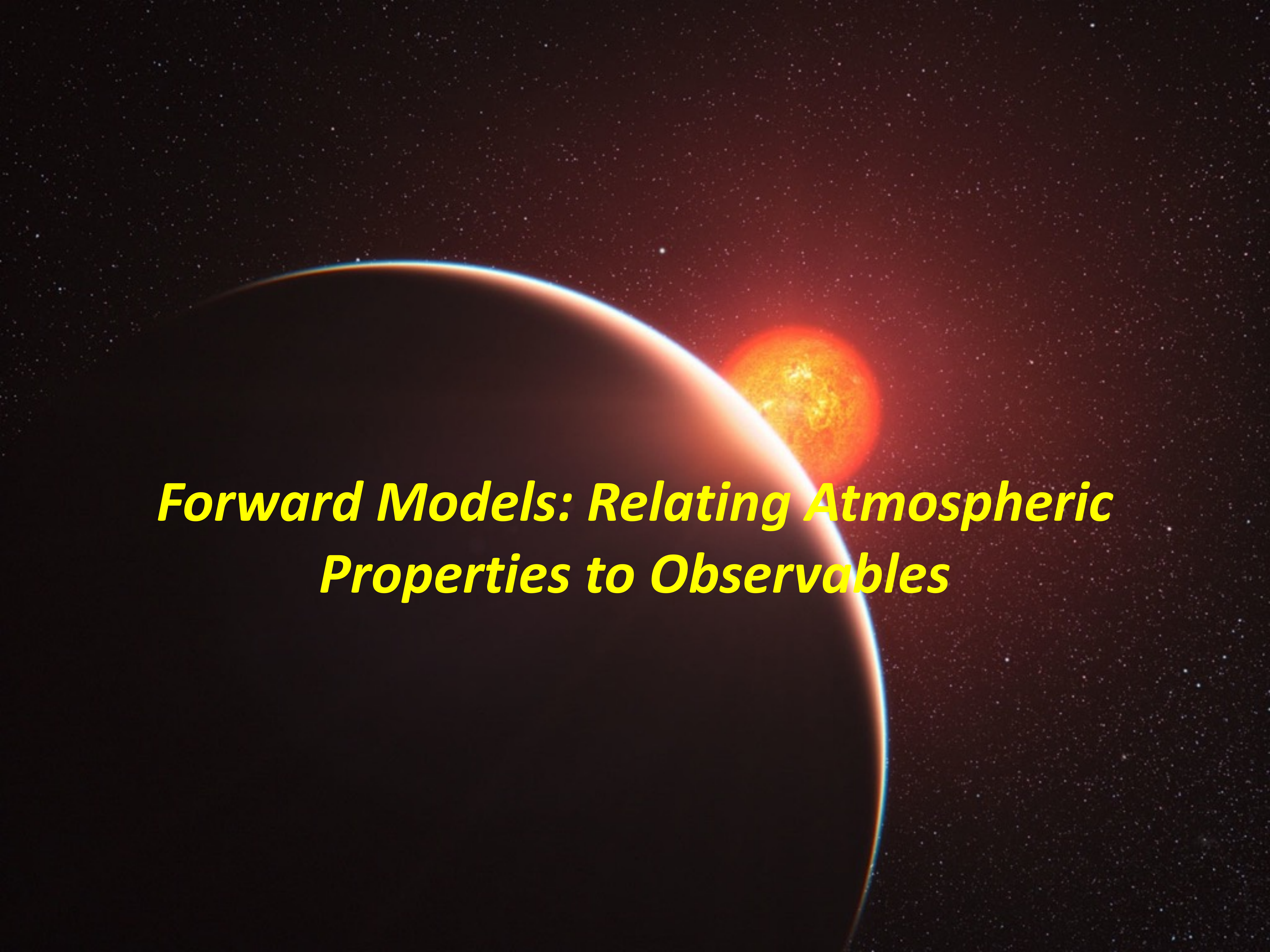
Measurements
(spectra)

Physics
(Radiative Transfer)

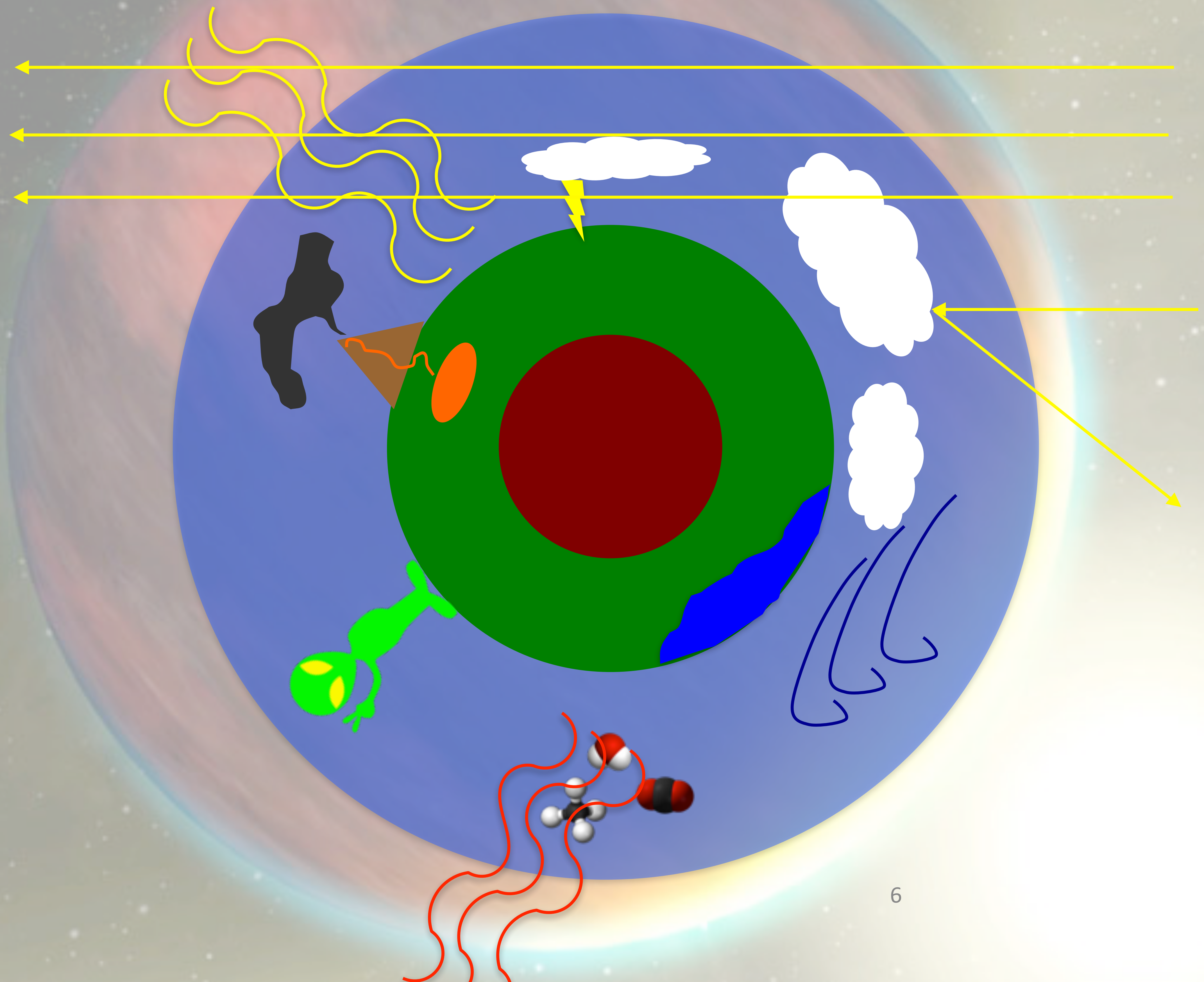
State (what we want)
(temperatures, abundances, clouds)

Noise ☹

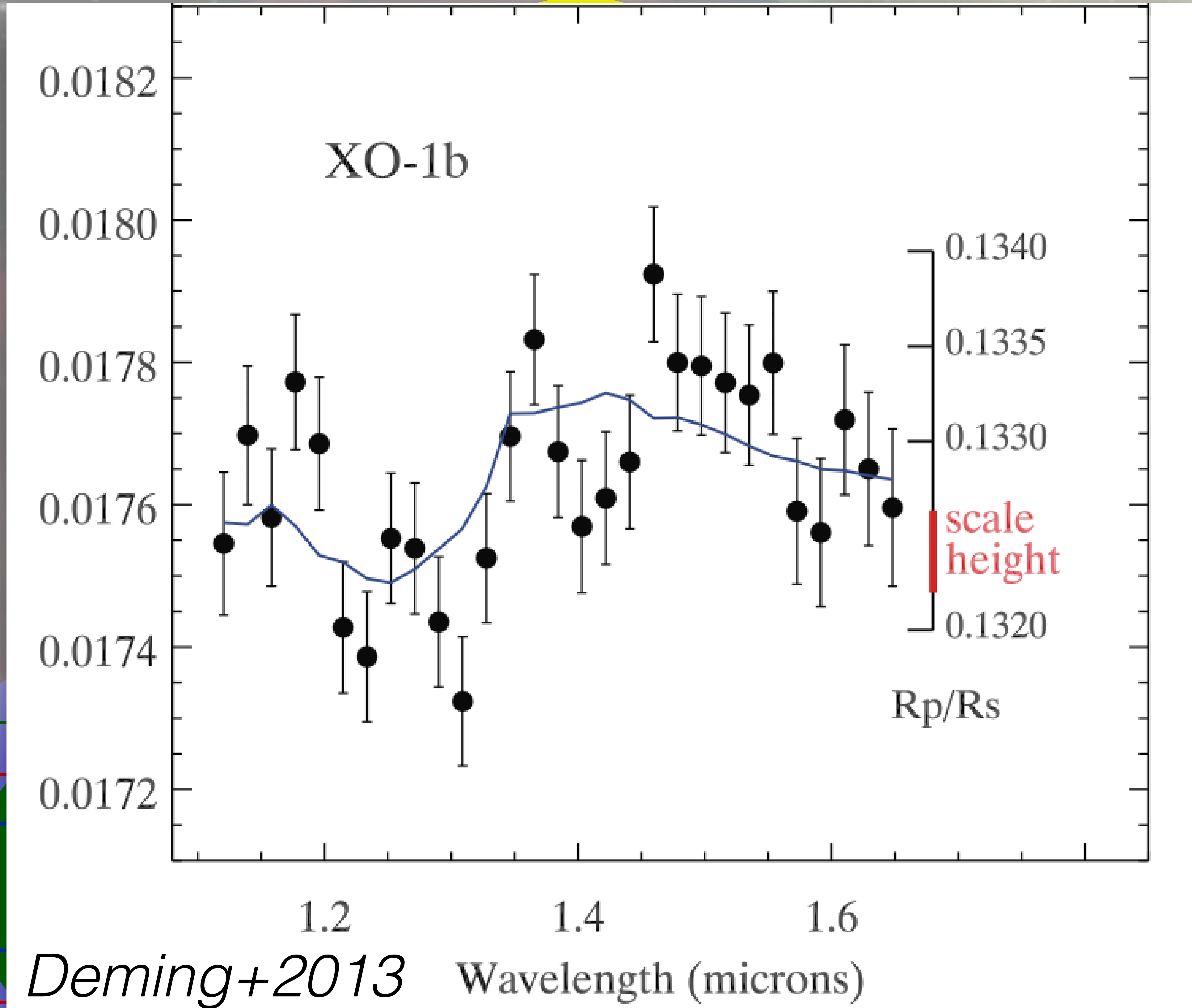
Forward Models: Relating Atmospheric Properties to Observables



What is **x**?

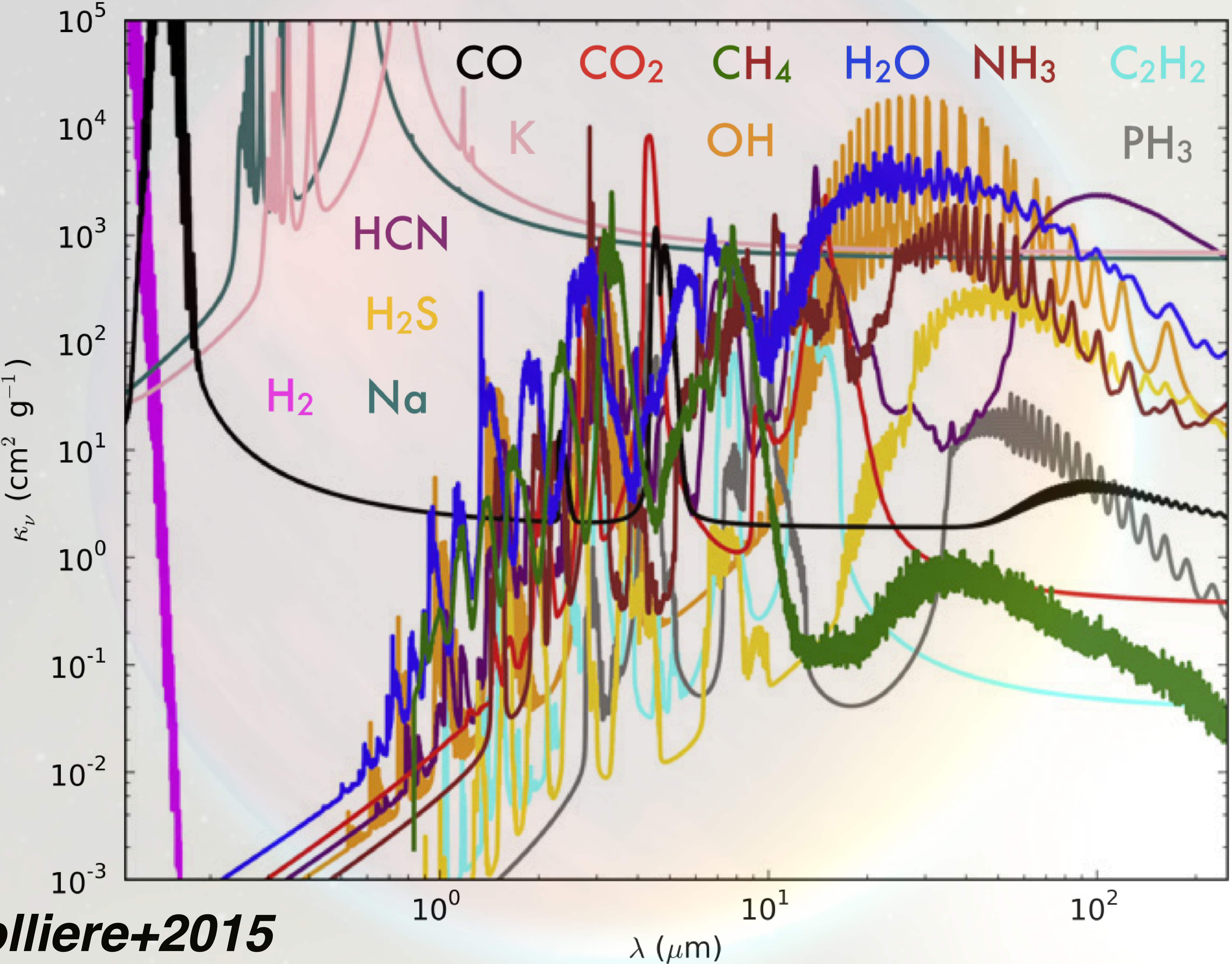


F(x): Transmission



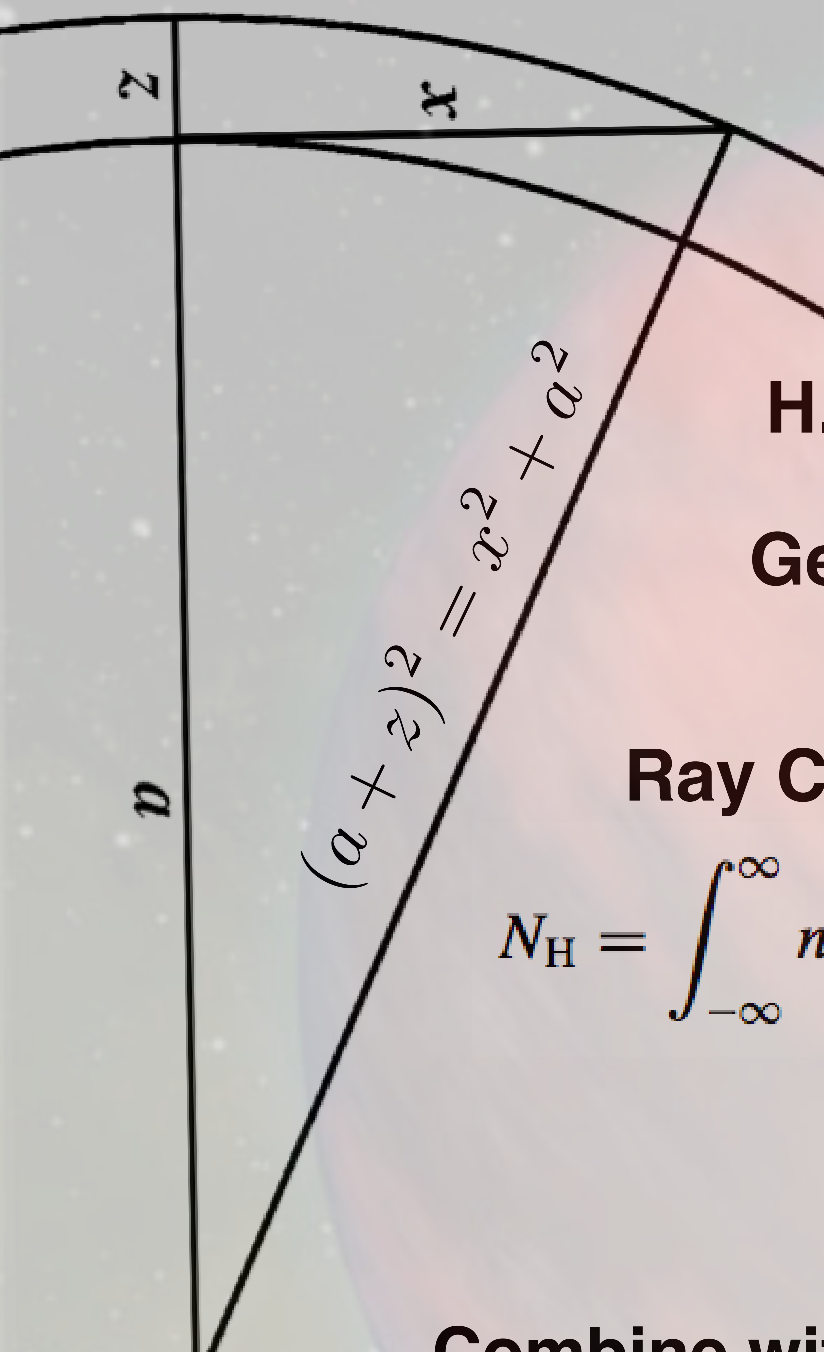
$\sim 10^{-4}$

Under the Rug: Opacities



Molliere+2015

F(x): Transmission



H.E: $n(z) = n_0 e^{-z/H}$

Geom, z<<a: $z \approx \frac{x^2}{2a}$

$$n(x) = n_0 \exp\left(\frac{-x^2}{2aH}\right)$$

$$H = \frac{k_b T}{\mu g}$$

Ray Column Abund.

$$N_{\rm H} = \int_{-\infty}^{\infty} n(x)\mathrm{d}x = n_0\sqrt{2\pi a H}$$

Optical Depth along tangent ray at some Z

$$\tau_\lambda(z) = \sigma_\lambda(T,P)\frac{P(z)}{k_B T}\sqrt{2\pi R_p H}$$

Combine with H.E. and re-arrange:

$$z(\lambda) = H \ln\left(\xi_{\rm abs} P_{z=0}\sigma_{\rm abs}(\lambda)/\tau_{\rm eq} \times \sqrt{2\pi R_{\rm p}/kT\mu g}\right)$$

Effective Eclipse Depth:

$$\alpha_\lambda \approx \left(\frac{R_p}{R_{\rm star}}\right)^2 + \frac{2R_p z_\lambda}{R_{\rm star}^2}$$

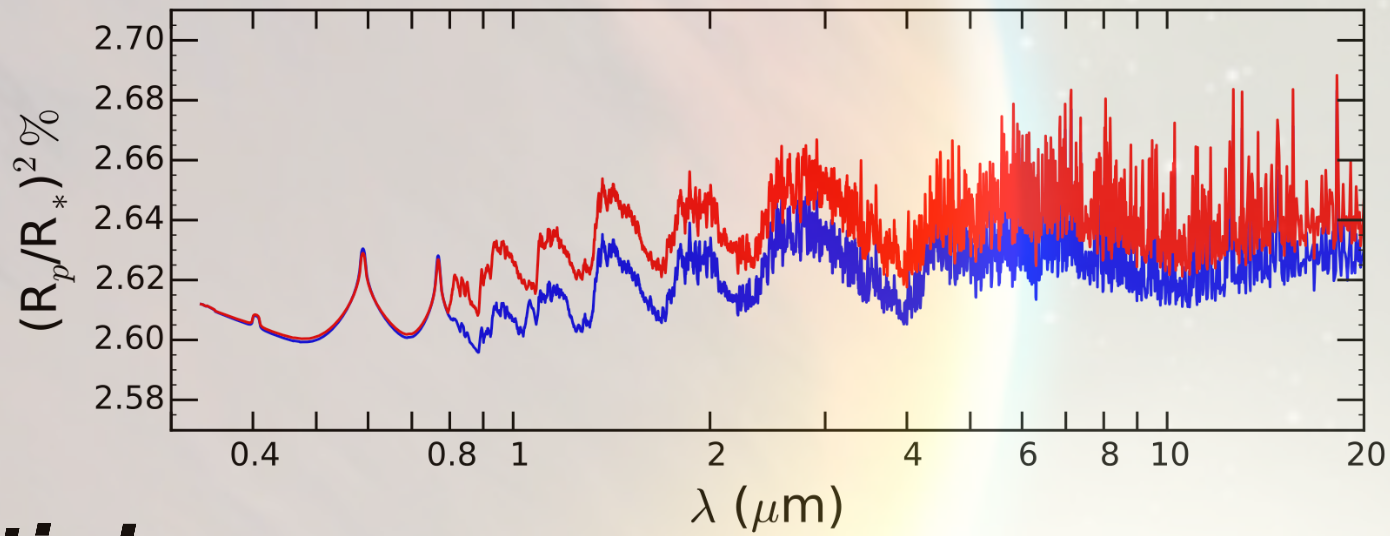
Fortney+2005 Lecavelier des Etangs+2008

F(x): Transmission

$$z(\lambda) = H \ln\left(\xi_{\rm abs} P_{z=0} \sigma_{\rm abs}(\lambda)/\tau_{\rm eq} \times \sqrt{2\pi R_{\rm p}/kT\mu g}\right) \qquad \alpha_\lambda \approx \left(\frac{R_p}{R_{\rm star}}\right)^2 + \frac{2R_p z_\lambda}{R_{\rm star}^2}$$

Spectral "shape" for a single gas

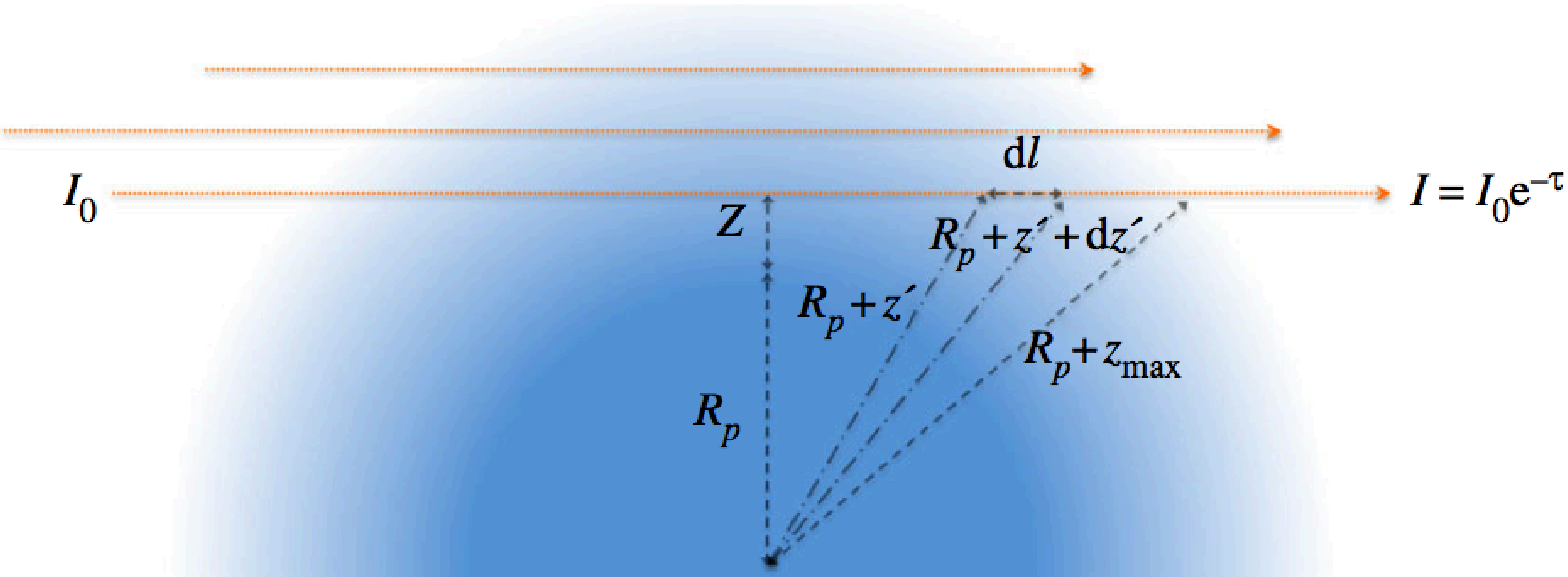
$$\frac{d\alpha_\lambda}{d\lambda} = \frac{2R_p}{R_{\rm star}^2} H \frac{d\ln(\sigma_\lambda)}{d\lambda}$$



Spectral "shape" for multiple gases

$$\frac{d\alpha_\lambda}{d\lambda} = \frac{2R_p}{R_{\rm star}^2} H \frac{1}{1 + \frac{\xi_2 \sigma_{\lambda,2}}{\xi_1 \sigma_{\lambda,1}}} \left(\frac{d\ln\sigma_{\lambda,1}}{d\lambda} + \frac{\xi_2 \sigma_{\lambda,2}}{\xi_1 \sigma_{\lambda,1}} \frac{d\ln\sigma_{\lambda,2}}{d\lambda} \right)$$

F(x): Transmission



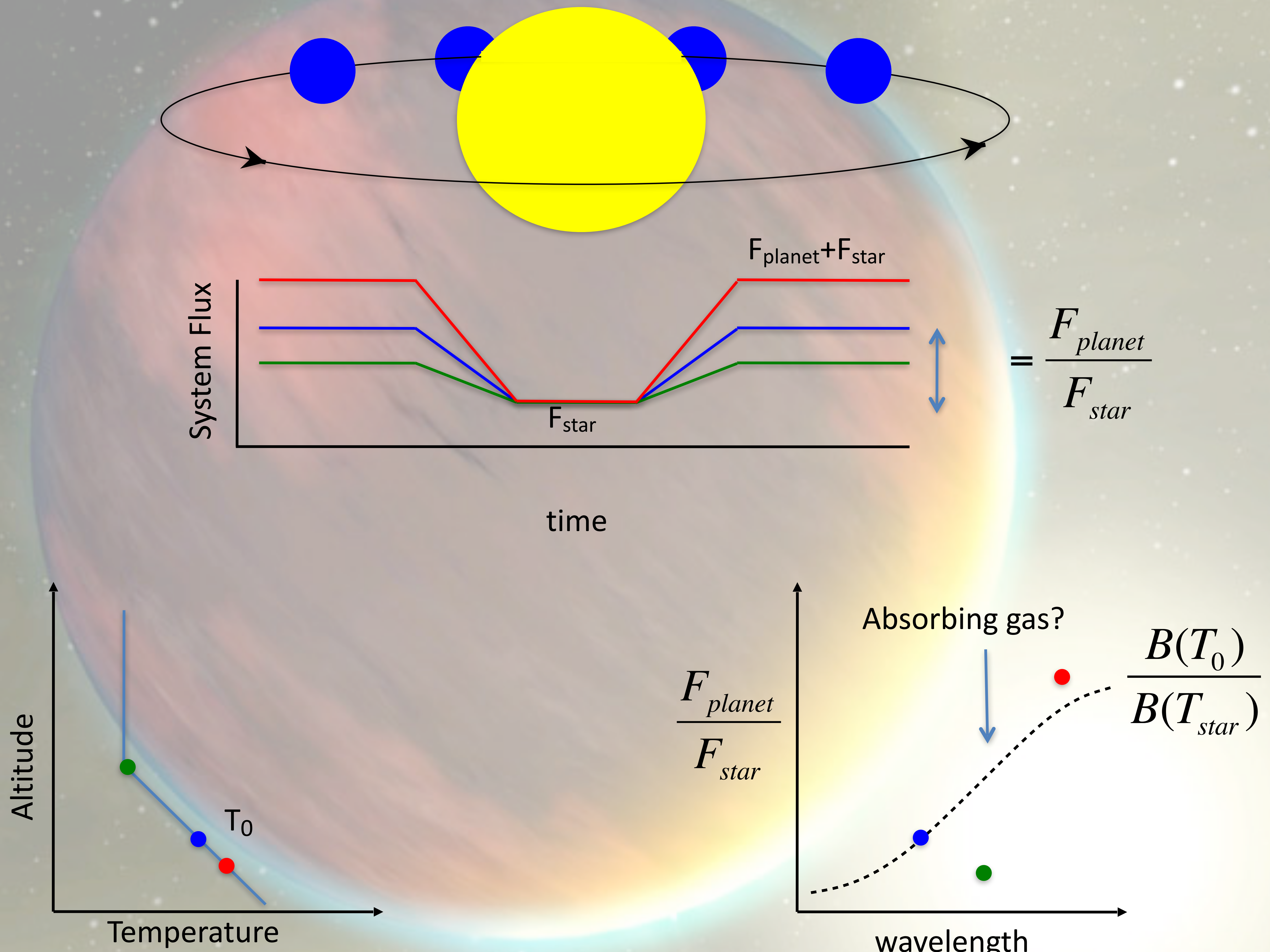
$$\tau_i(\lambda, z) = 2\int_0^{\ell(z)} \rho(z')\chi_i(z')\sigma_i(\lambda, T)\,\mathrm{d}\ell$$

$$\mathrm{d}\ell = \sqrt{(R_p + z' + \mathrm{d}z')^2 - (R_p + z)^2} - \sqrt{(R_p + z')^2 - (R_p + z)^2}$$

$$\kappa(\lambda) = \frac{R_p^2 + 2\int_0^{z_{\max}} (R_p + z)(1 - \mathrm{e}^{-\tau(z,\lambda)})\,\mathrm{d}z}{R_*^2}$$

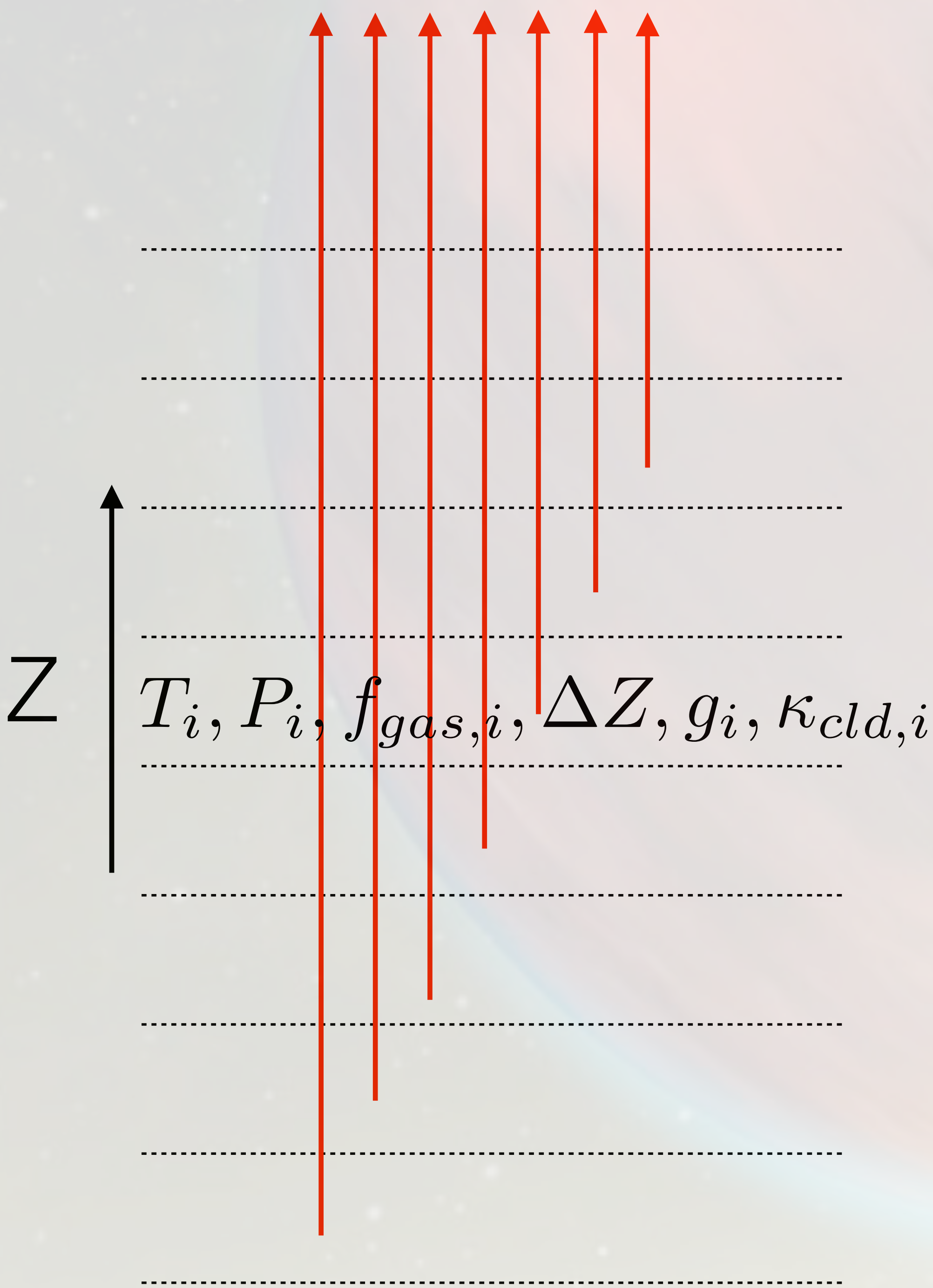
Brown 2001
Tinetti+2012

F(x): Emission



$F_{planet}+F_{star}$
System Flux
F_{star}
time
$= \frac{F_{planet}}{F_{star}}$
Altitude
T_0
Temperature
Absorbing gas?
$\frac{F_{planet}}{F_{star}}$
$\frac{B(T_0)}{B(T_{star})}$
wavelength

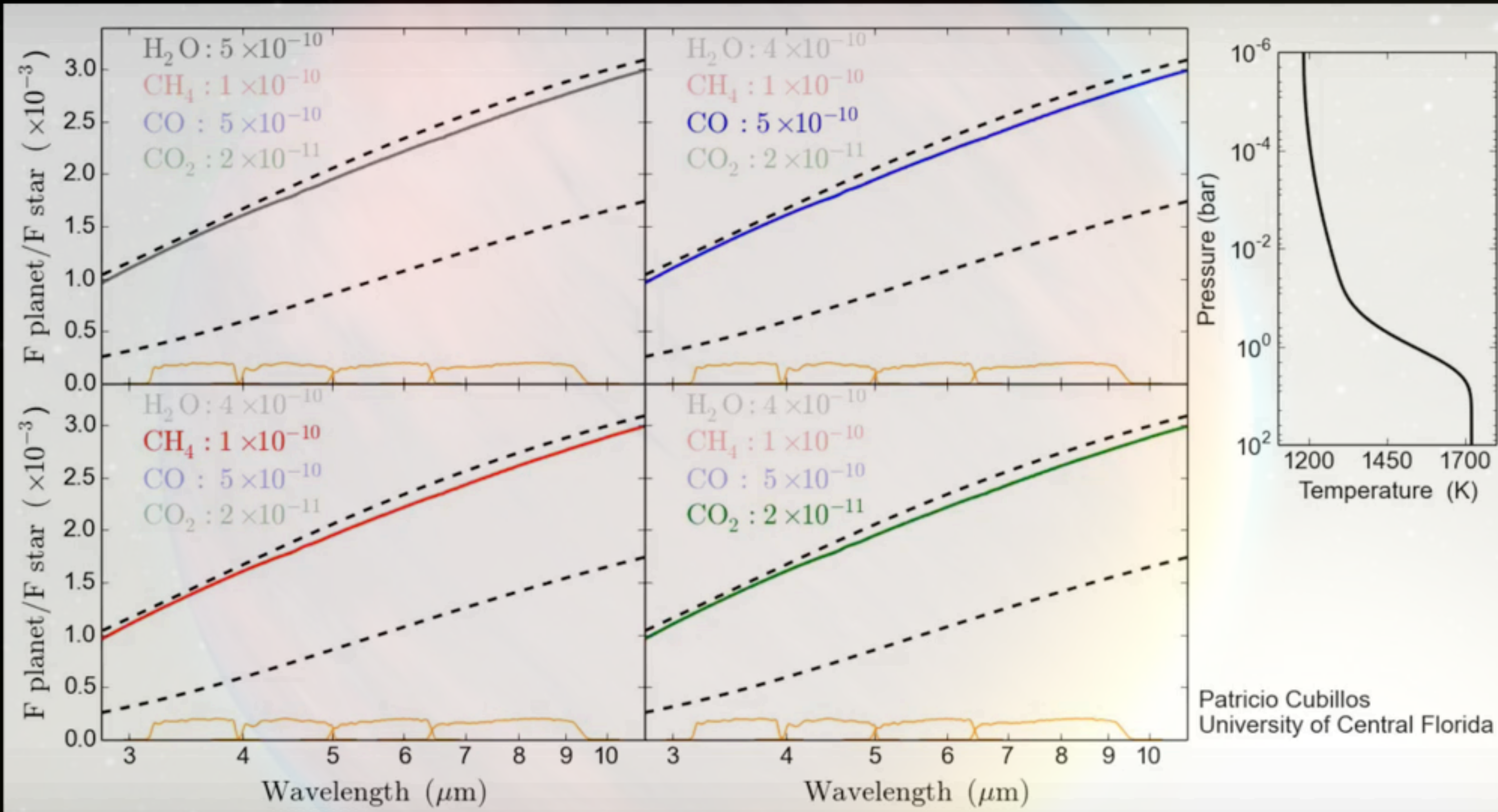
F(x): Emission



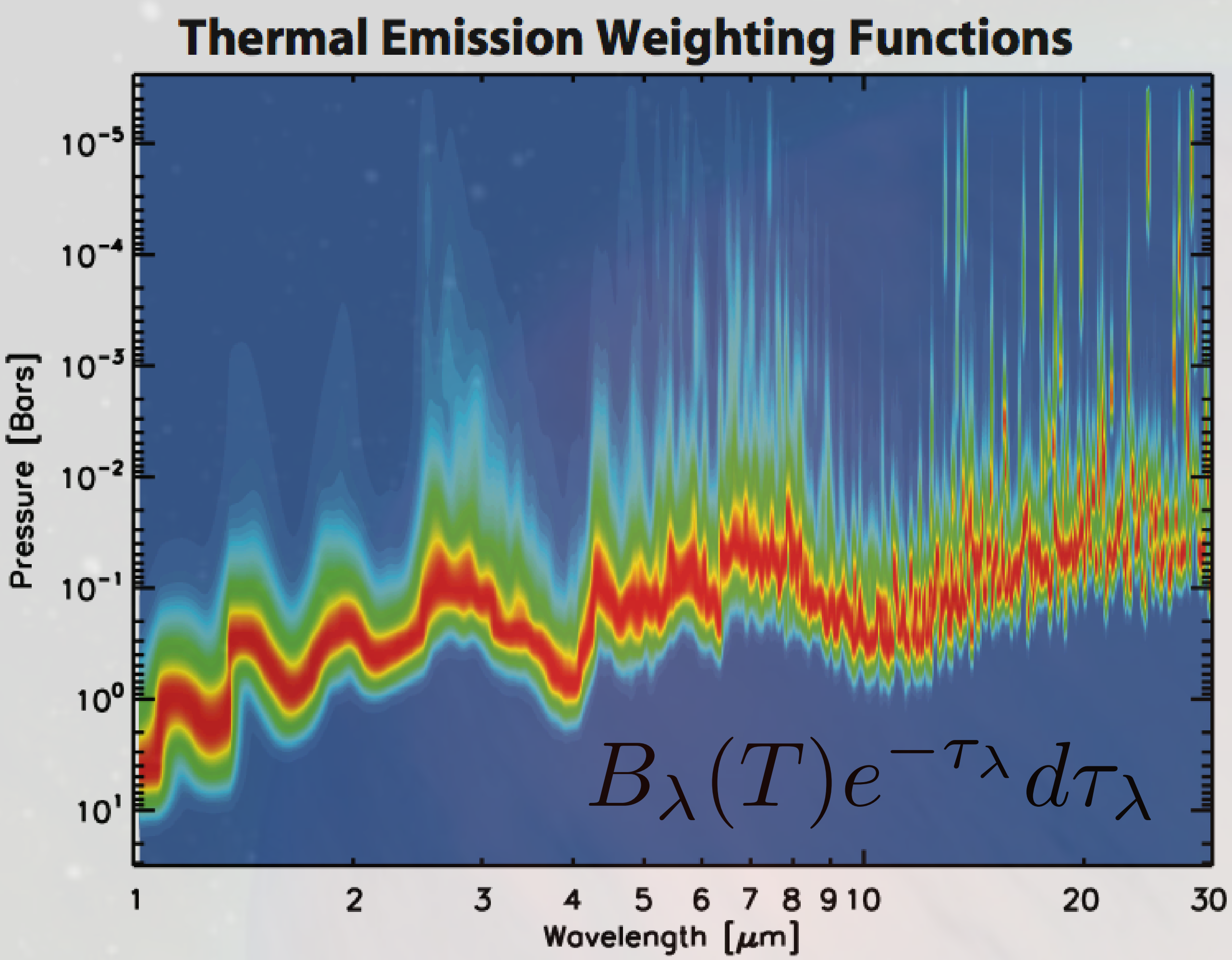
$$F_\lambda = \int_{d\Omega} \int_\tau B_\lambda(T) e^{-\tau_\lambda} d\tau_\lambda d\Omega$$

$$\Delta\tau_{k,z,\lambda} = f_{k,z}\sigma_{k,z,\lambda}\frac{\Delta P_z}{\mu_{\mathrm{atm}} g}$$

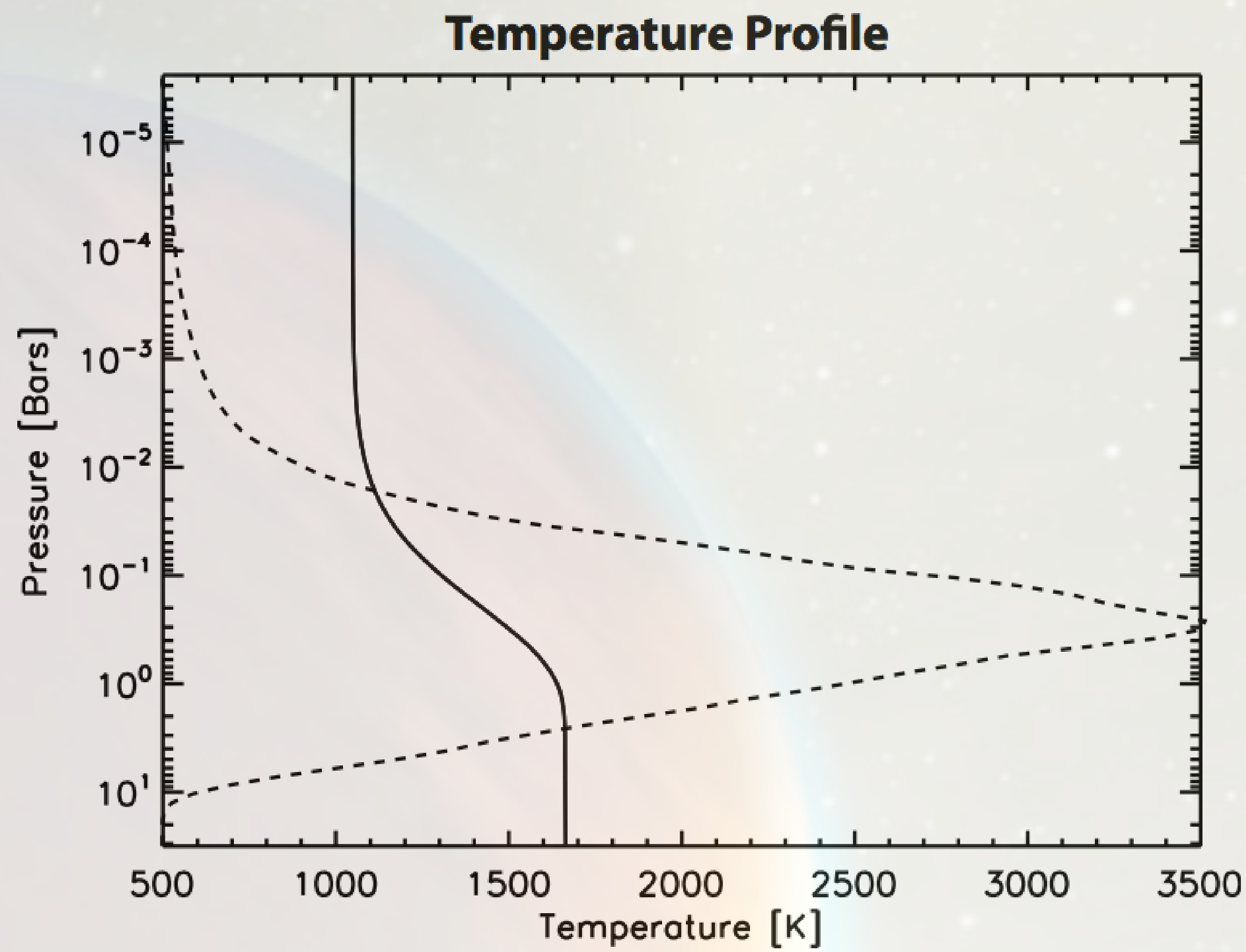
F(x): Emission



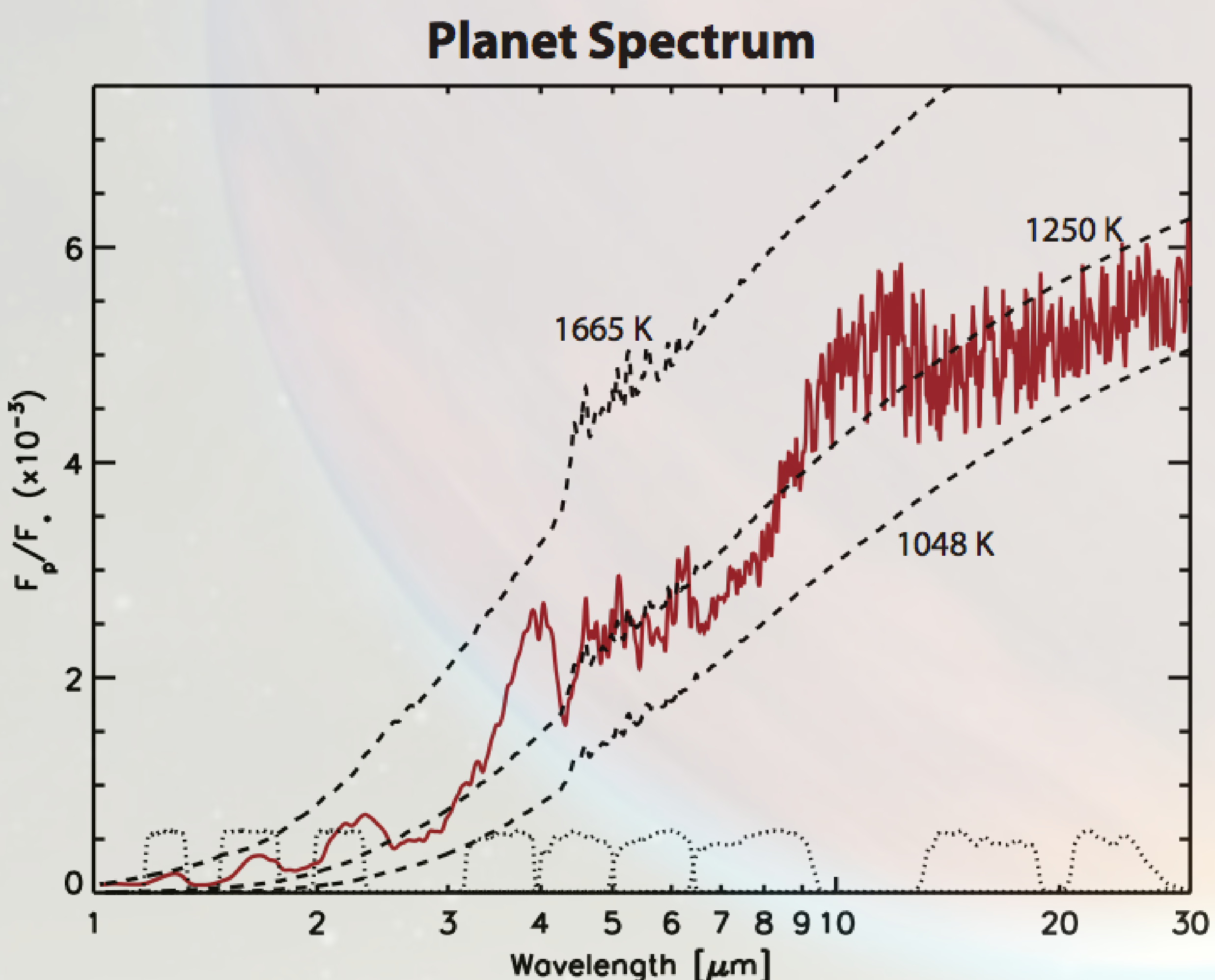
H$_2$O : 5 ×10^{-10}
CH$_4$: 1 ×10^{-10}
CO : 5 ×10^{-10}
CO$_2$: 2 ×10^{-11}
H$_2$O : 4 ×10^{-10}
CH$_4$: 1 ×10^{-10}
CO : 5 ×10^{-10}
CO$_2$: 2 ×10^{-11}
H$_2$O : 4 ×10^{-10}
CH$_4$: 1 ×10^{-10}
CO : 5 ×10^{-10}
CO$_2$: 2 ×10^{-11}
H$_2$O : 4 ×10^{-10}
CH$_4$: 1 ×10^{-10}
CO : 5 ×10^{-10}
CO$_2$: 2 ×10^{-11}
F planet/F star (×10^{-3})
Wavelength (μm)
Pressure (bar)
Temperature (K)
1200 1450 1700
Patricio Cubillos
University of Central Florida



Thermal Emission Weighting Functions
Pressure [Bars]
Wavelength [μm]
$B_\lambda(T)e^{-\tau_\lambda}d\tau_\lambda$



Temperature Profile
Pressure [Bars]
Temperature [K]



Planet Spectrum
F_p/F_* (x10^{-3})
1250 K
1665 K
1048 K
Wavelength [μm]

Basic Retrieval Setup

Typical Retrieved Parameters:

-Molecular Abundances: H_2O, CH_4, CO, CO_2 (sometimes NH_3, C_2H_2, HCN, H_2S, TiO, Na/K)

-Thermal structure in Emission (parameterized or L-b-L) (usually includes albedo/redistribution)

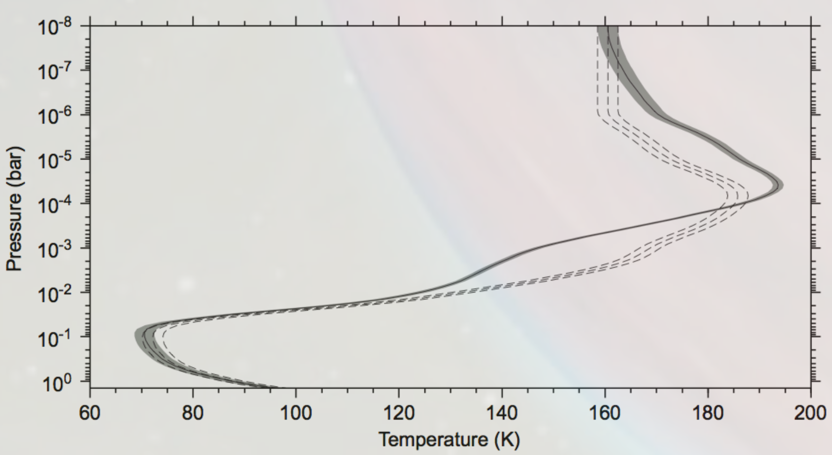
-Clouds—usually a "cloud-top-pressure" in transmission—generally ignored in emission

-C/O, [Fe/H] derived from Molecular abundances

$$\ln \mathcal{L}(\boldsymbol{y}|\boldsymbol{x}) = -\frac{1}{2}\sum_{i=1}^{n}\frac{\left(y_i - F_i(\boldsymbol{x})\right)^2}{s_i^2}$$

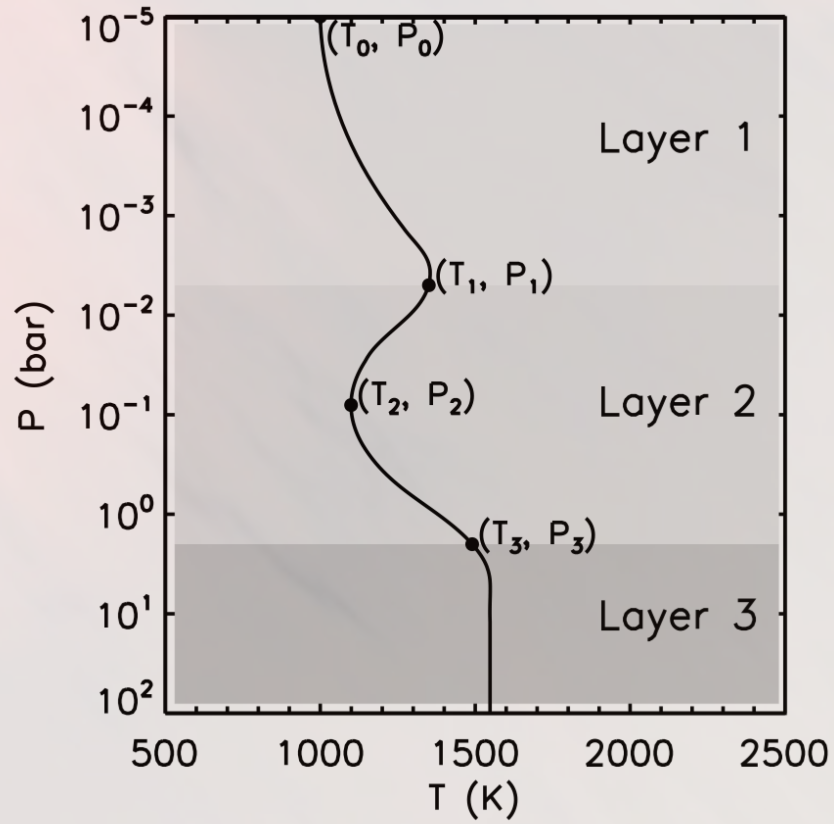
TP-Profile Parameterizations

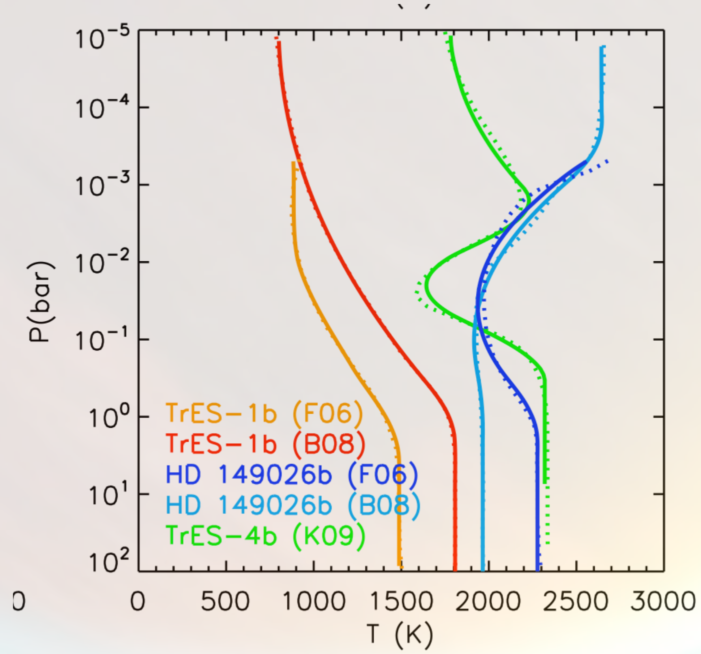
Classic-L-b-L



Irwin+2008

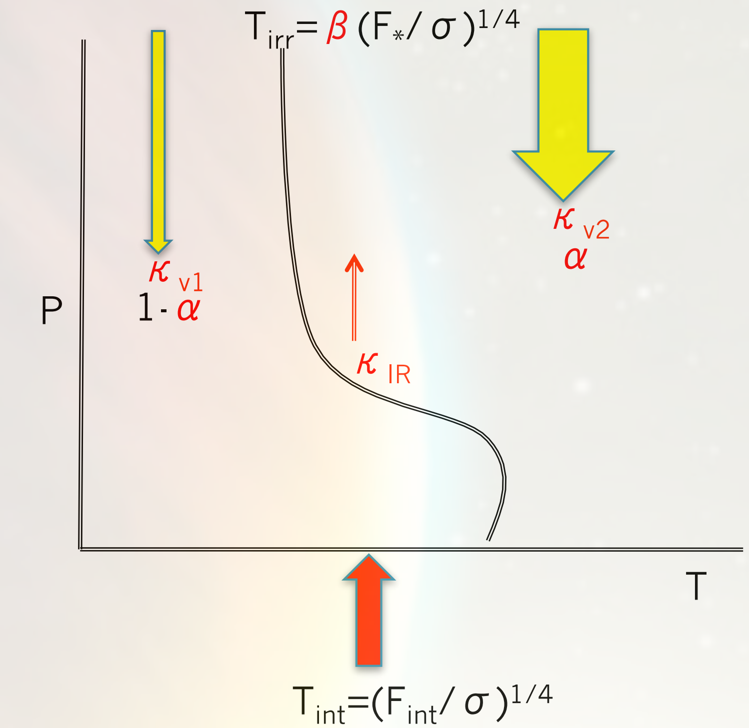
Parameterized

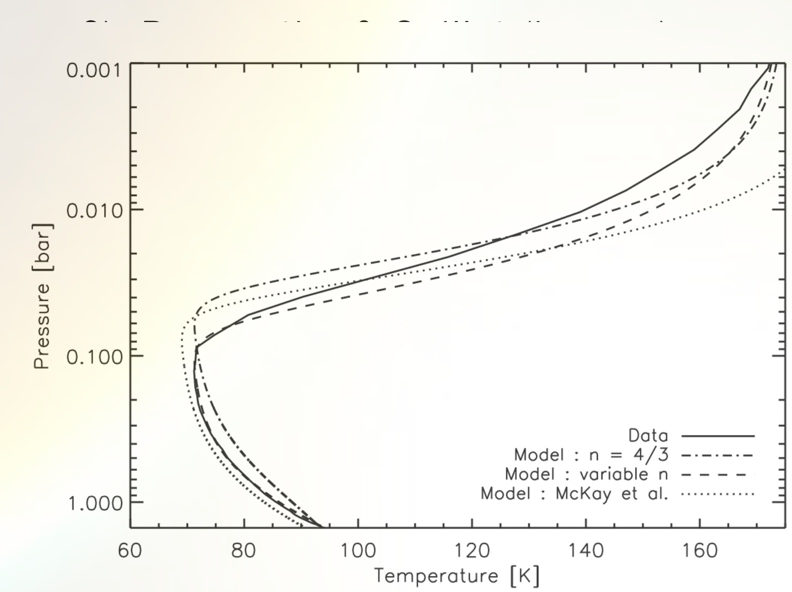




M&S 2009

Physically Motived Parameterization

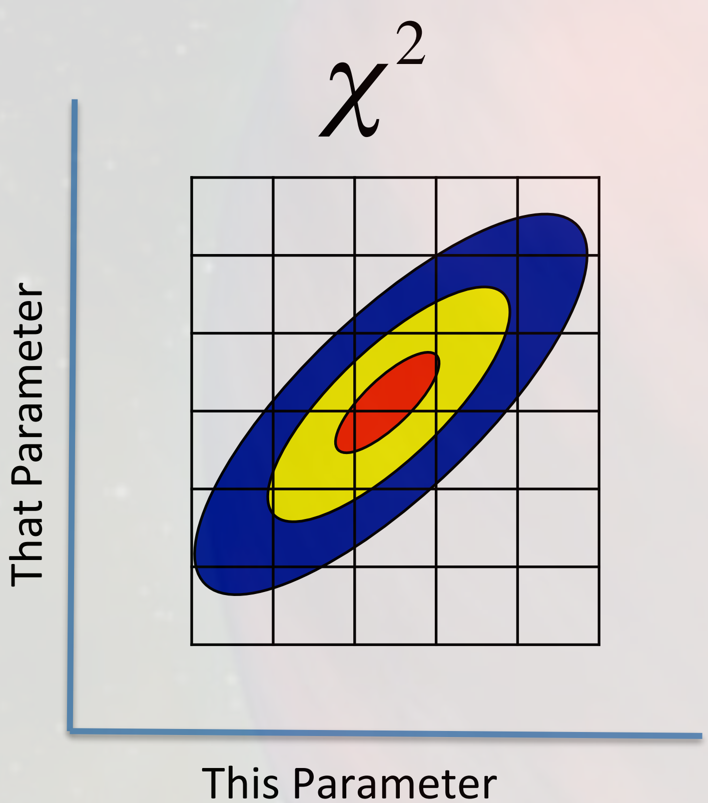




Guillot 2010; Robinson & Catling 2012

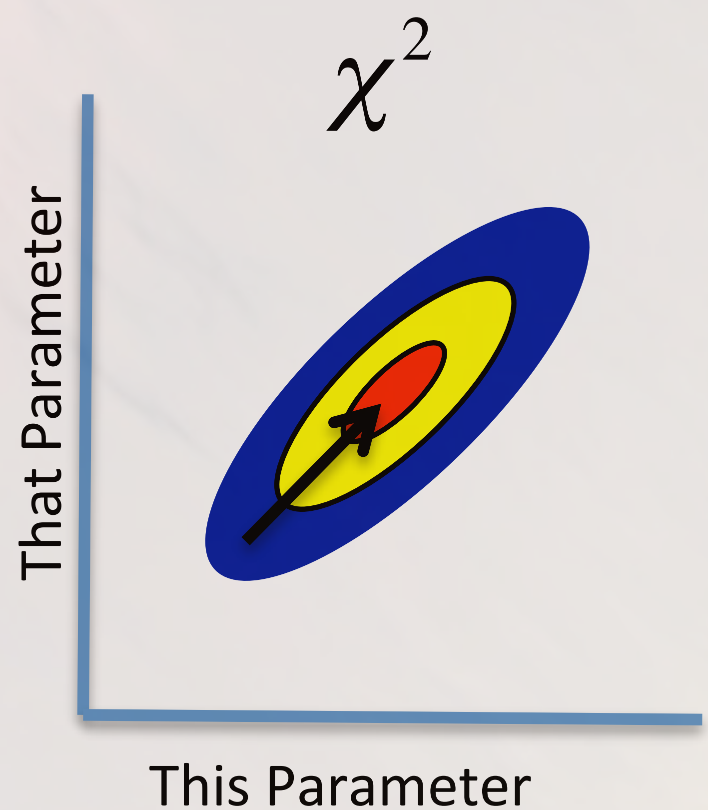
Parameter Estimation Tools

Grid Search

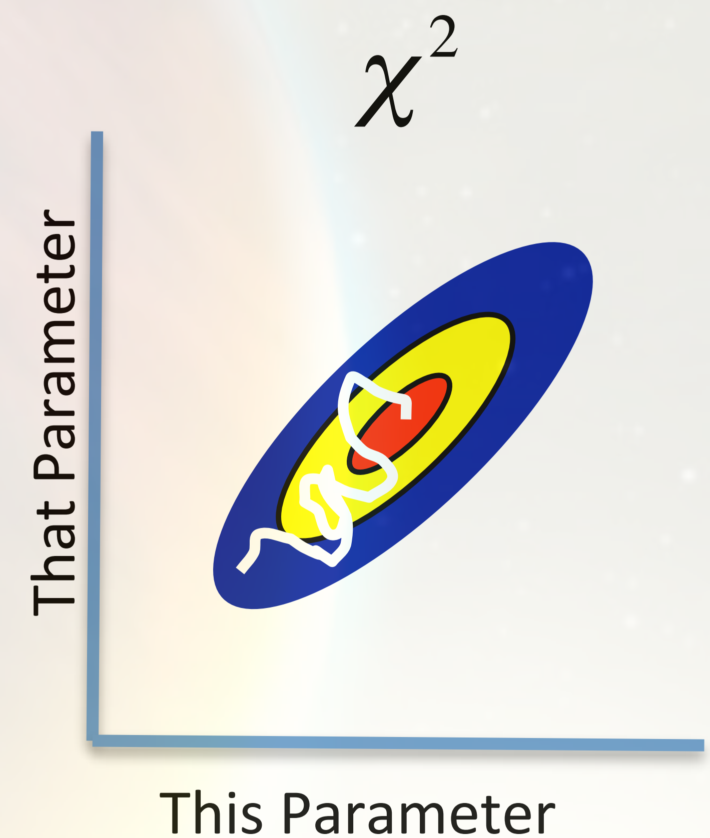


#models=#grids^#paramters

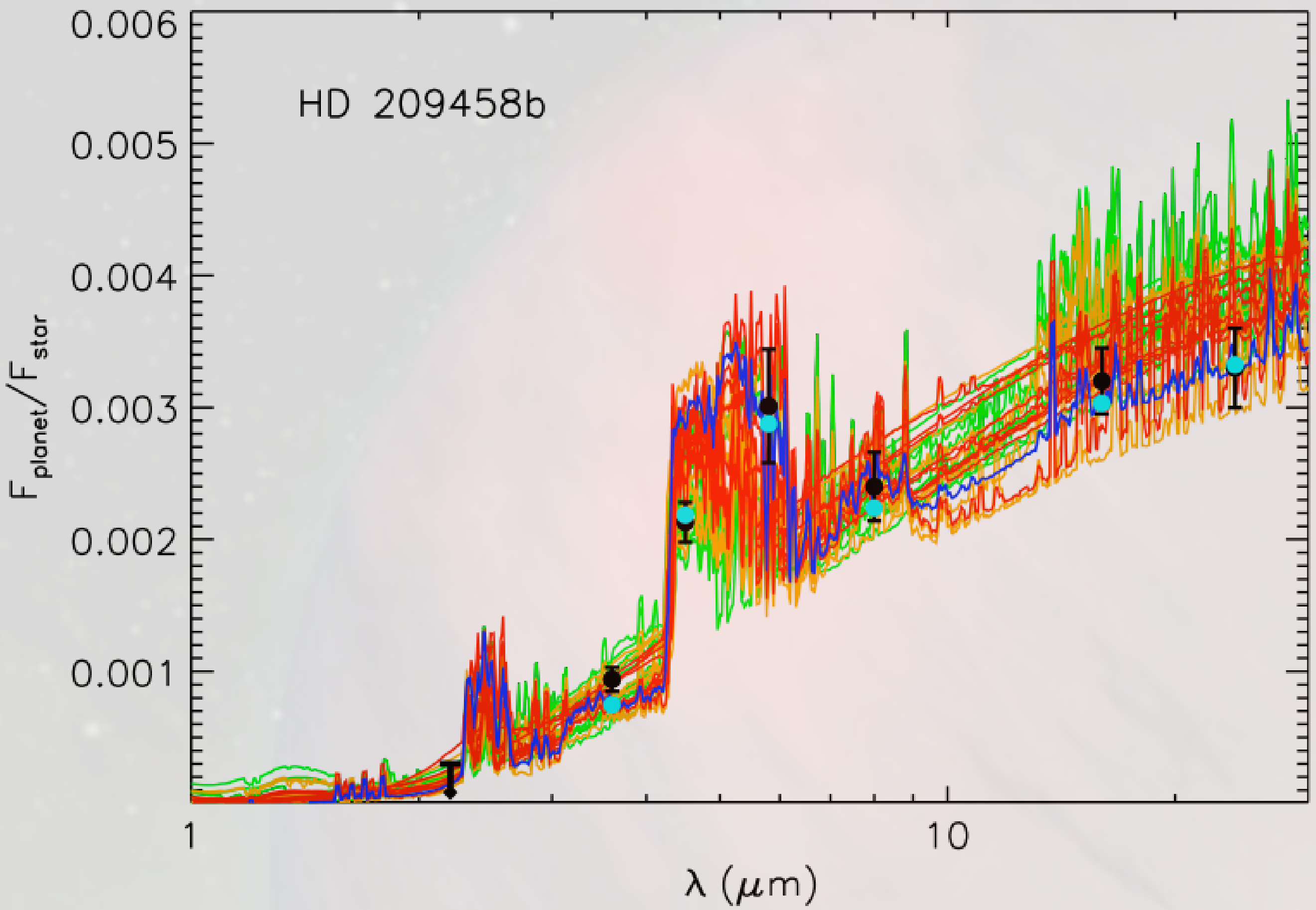
Optimal Estimation

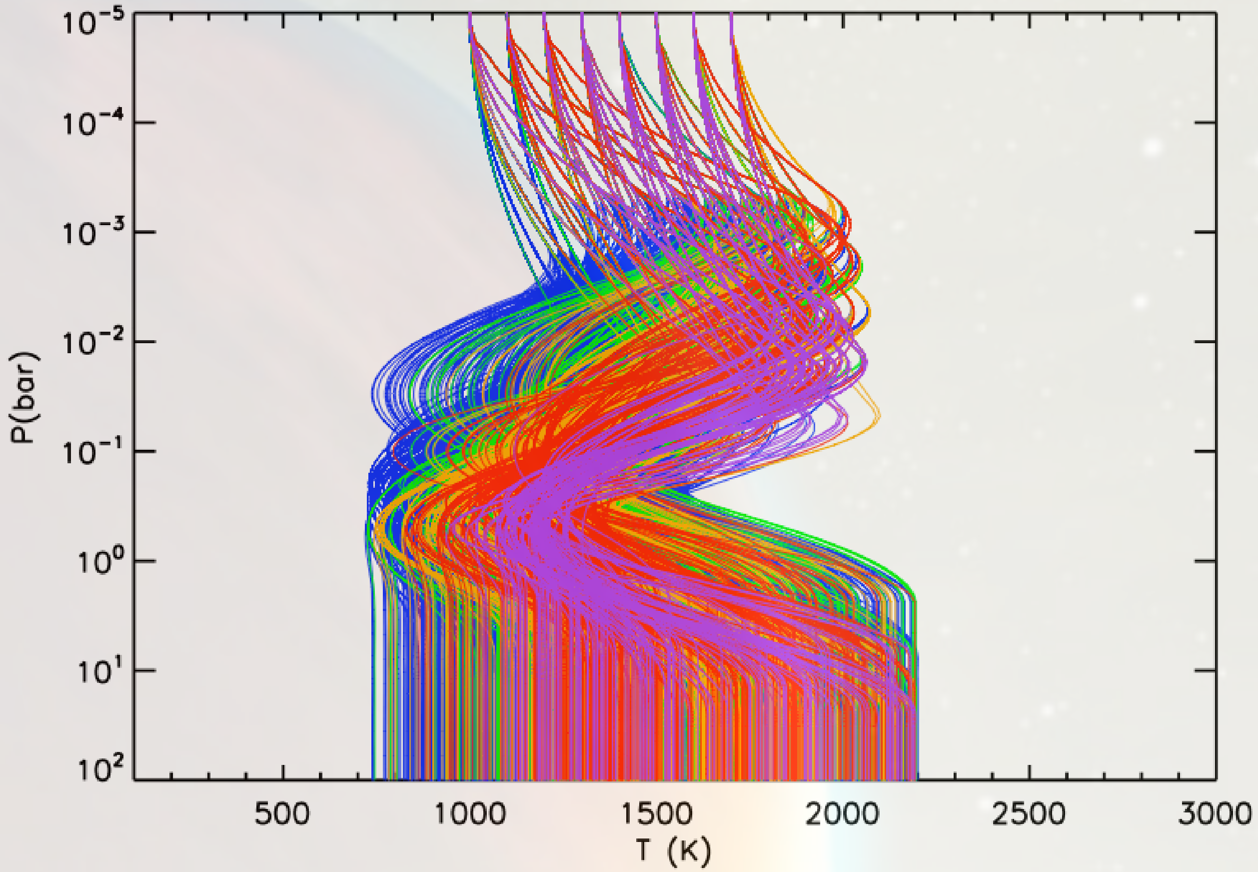


MCMC

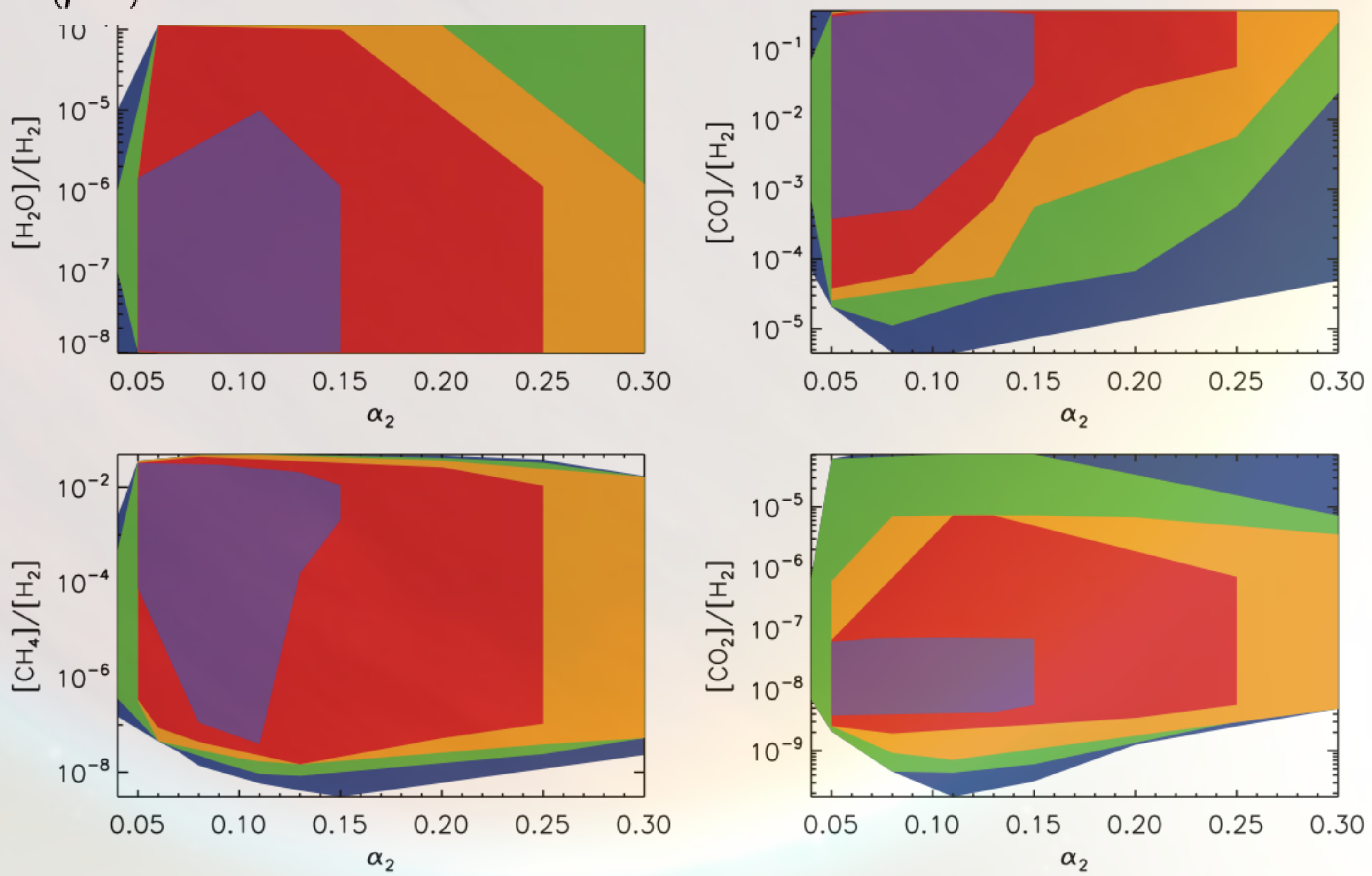


Grid Search: First Quantification of TP/Abundance uncertainties

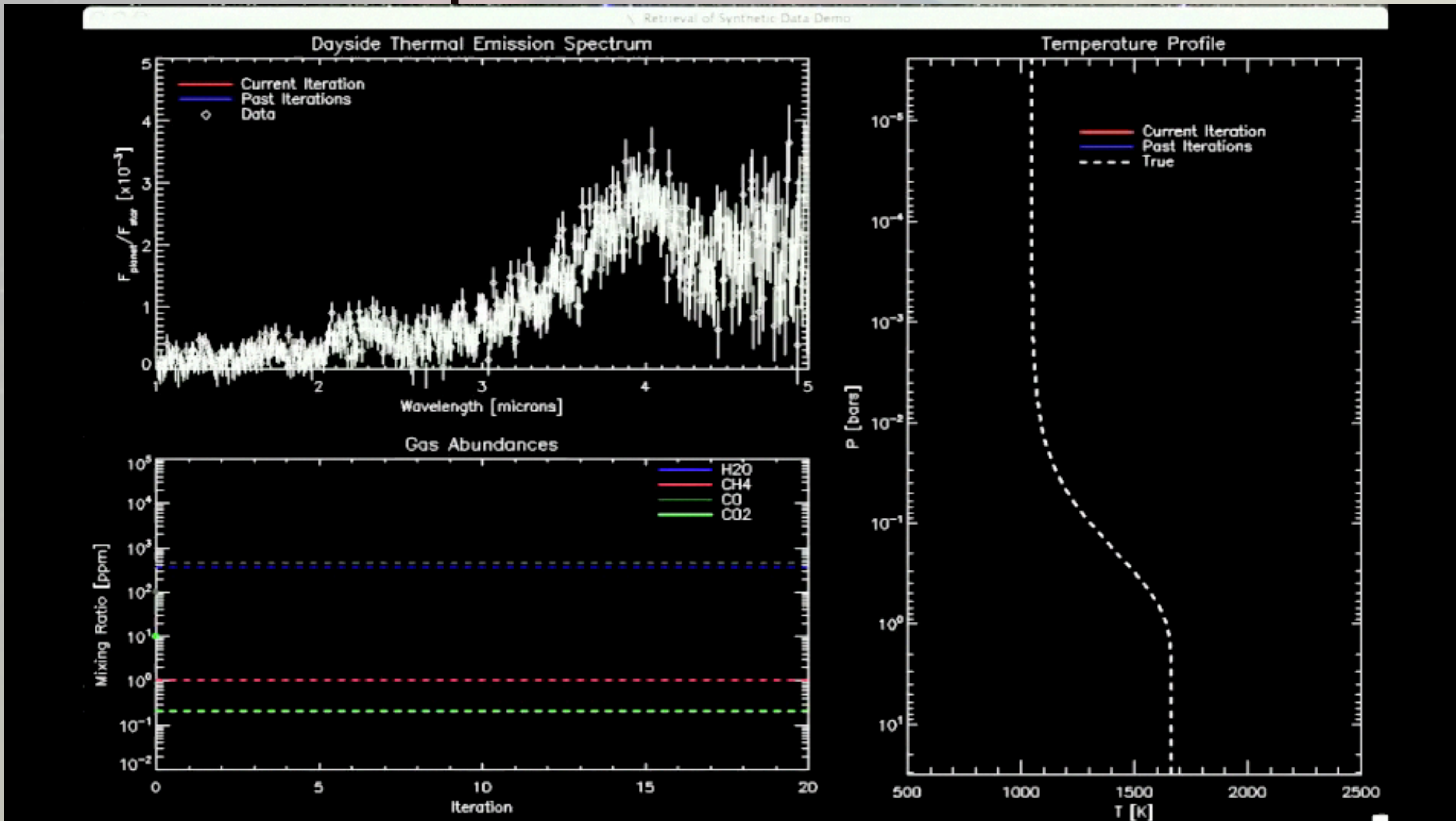




Madhusudhan
&
Seager 2009



Optimal Estimation



$$H = \frac{1}{2}\ln(|\hat{\mathbf{S}}^{-1}\mathbf{S}_a|)$$

Lee+2012;13 Line+2012;13;14 Barstow+2013;14;15

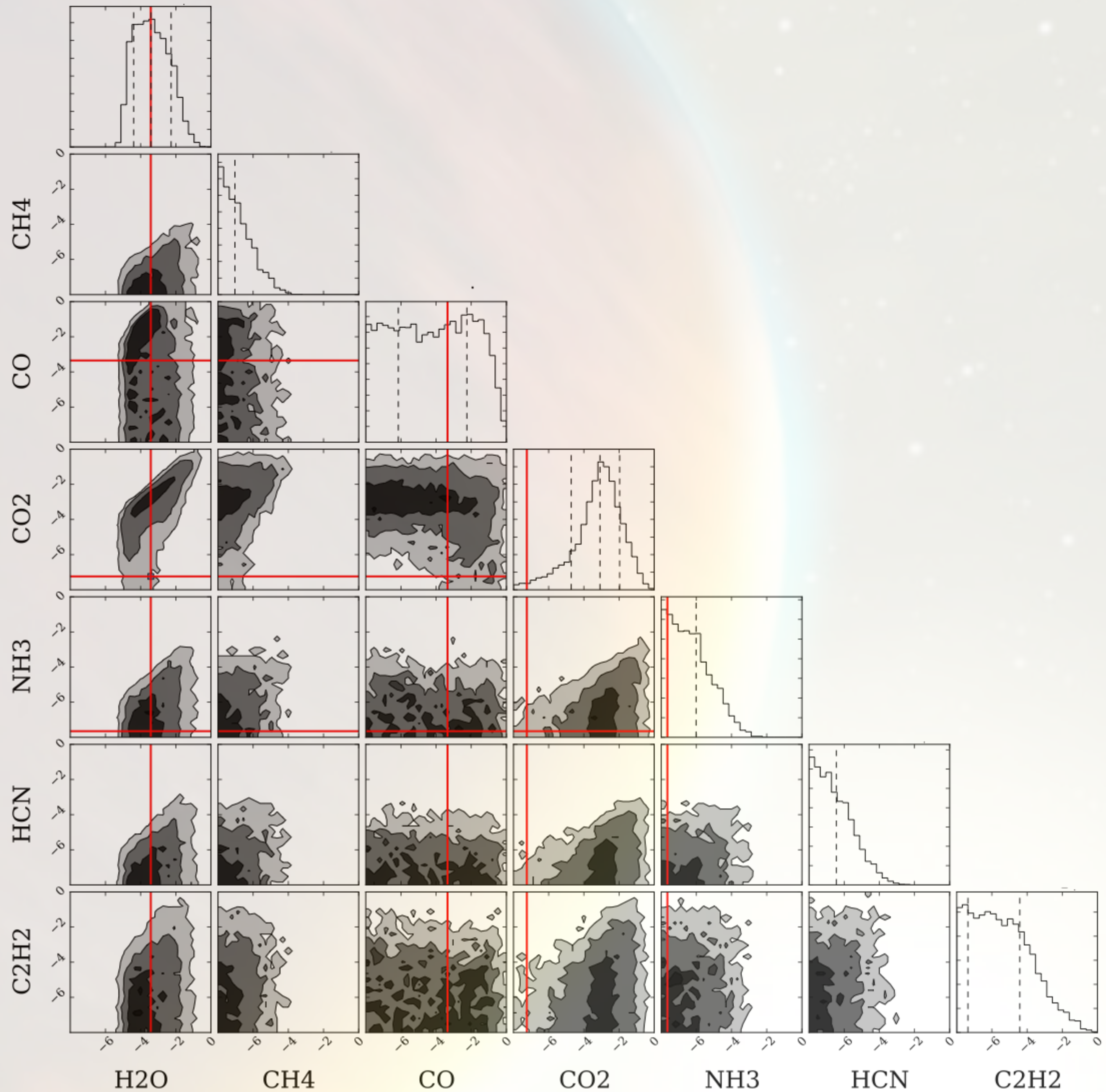
Markov Chain Monte Carlo

Formalism:

Bayes': $P(\mathbf{x}|\mathbf{y}) \propto \mathcal{L}(\mathbf{y}|\mathbf{x})P(\mathbf{x})$

$$\ln \mathcal{L}(y|x) = (\mathbf{y} - \mathbf{F}(\mathbf{x}))^T \mathbf{S}_\mathbf{e}^{-1}(\mathbf{y} - \mathbf{F}(\mathbf{x}))$$

*lnP(**x**)=whatever you want it to be*



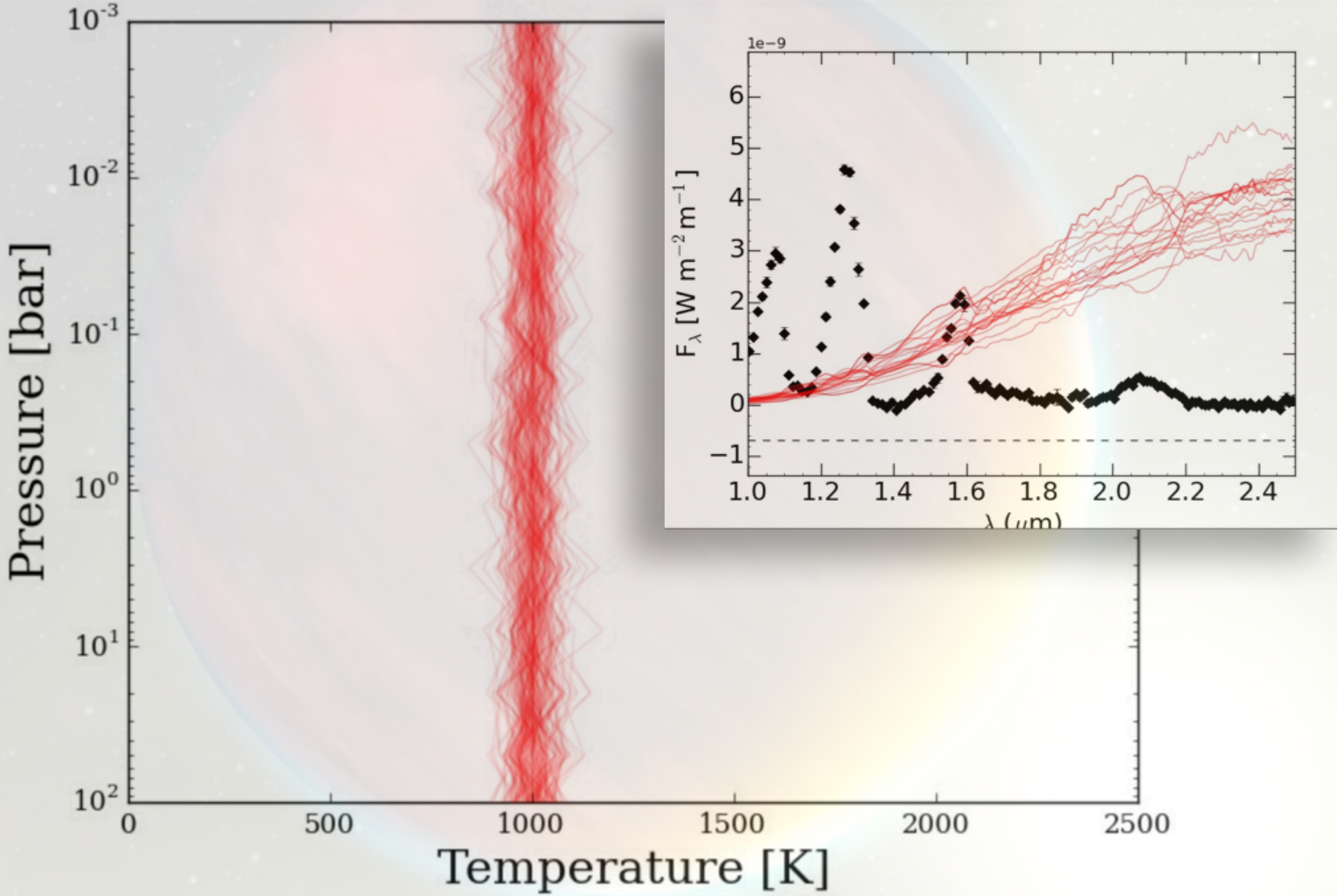
Pros:

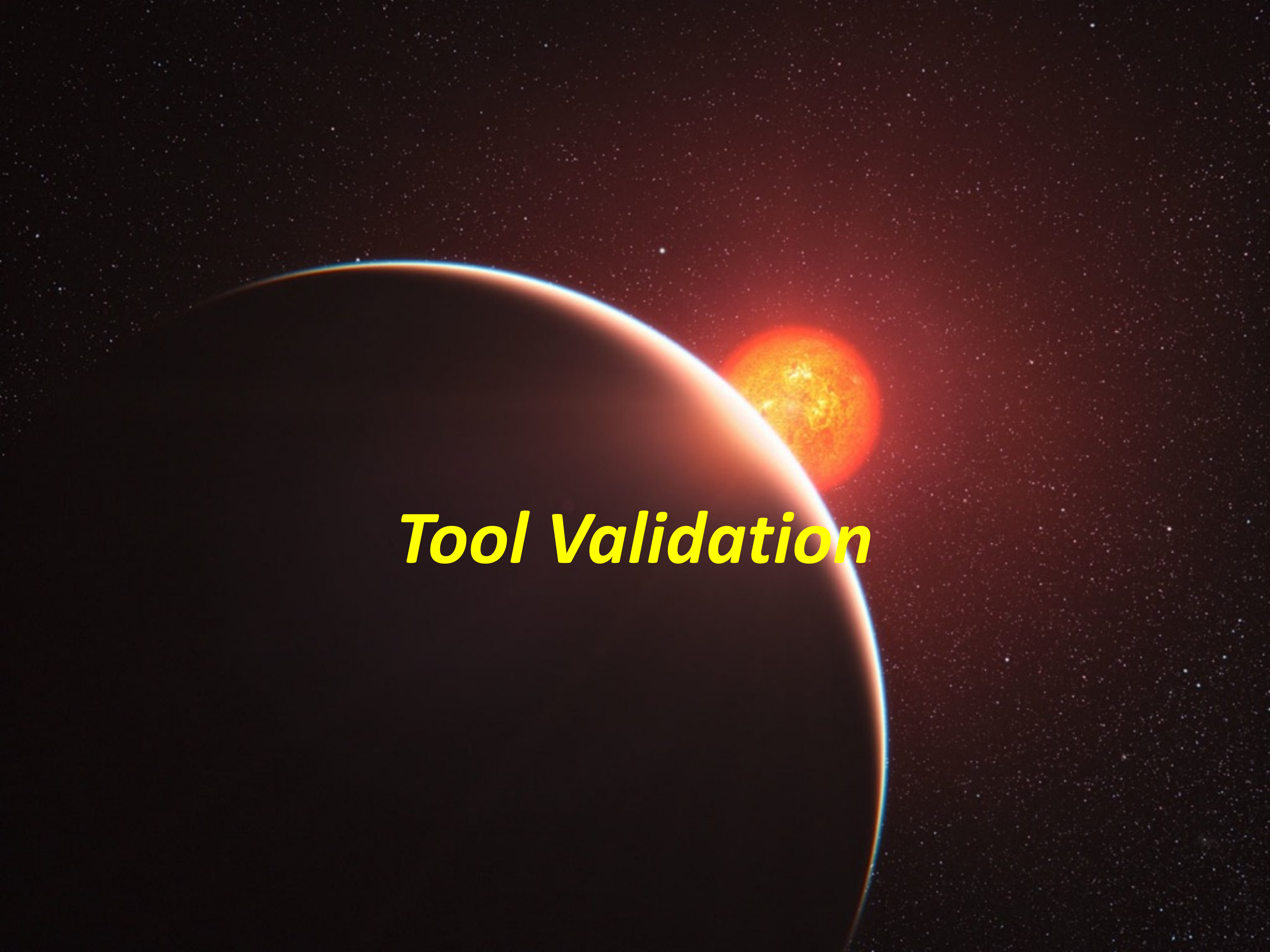
-Full exploration of parameter space
-no posterior shape assumptions
-arbitrary likelihood or prior functions
-can do "hierarchal" magic
-many off-the-shelf routines
(emcee, pymultinest etc)

Cons:

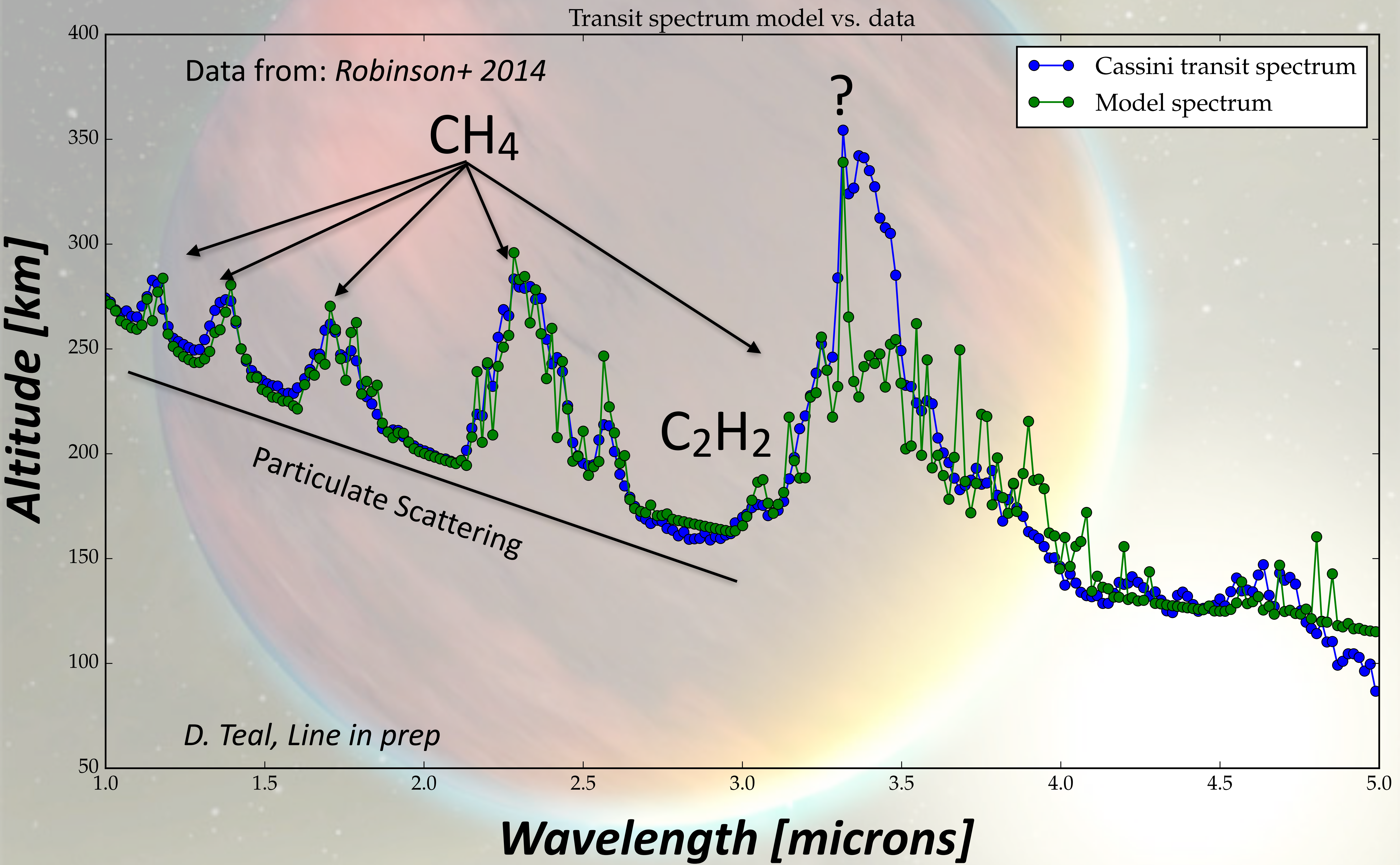
-oh so slow with complicated models

Madhusudhan+2011, Benneke & Seager 2012;13, Line et al. 2013;14;15;16, Waldmann et al. 2015

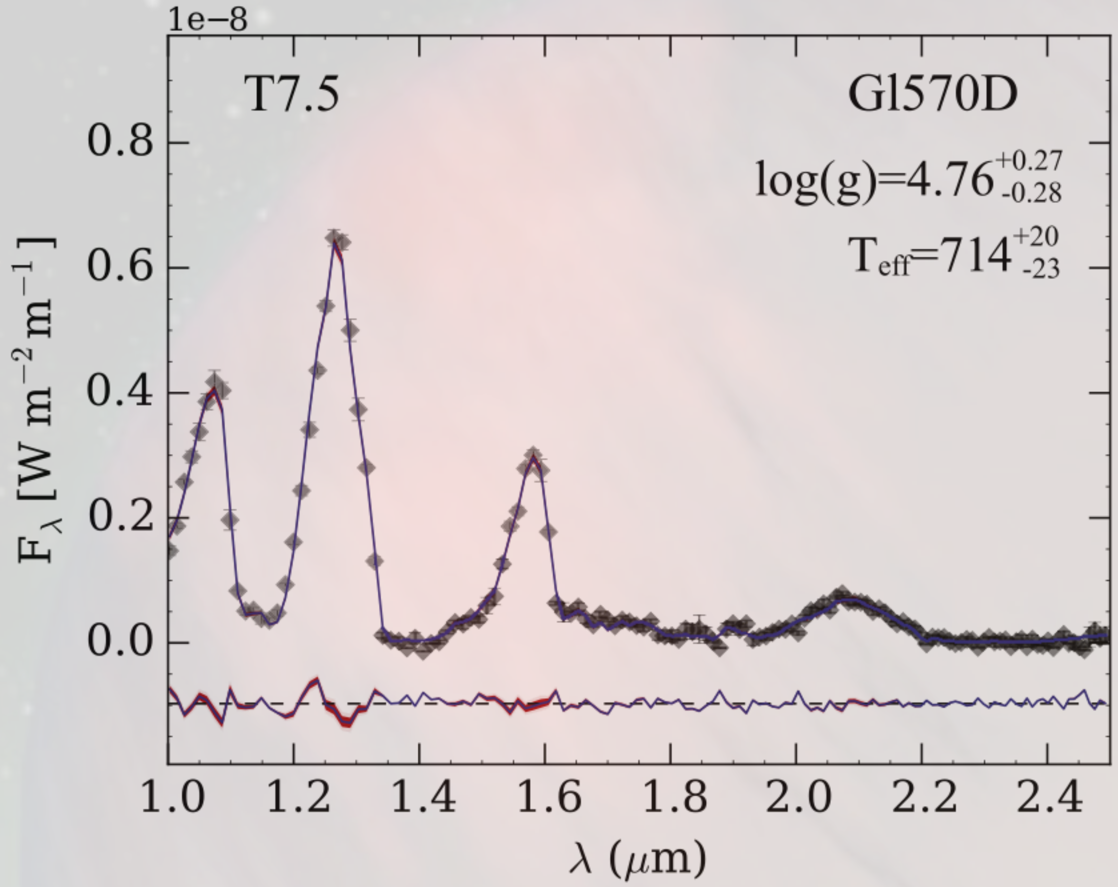

Pressure [bar]
Temperature [K]
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
0
500
1000
1500
2000
2500
1e−9
F_λ [W m^{-2} m^{-1}]
λ (μm)
6
5
4
3
2
1
0
−1
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4

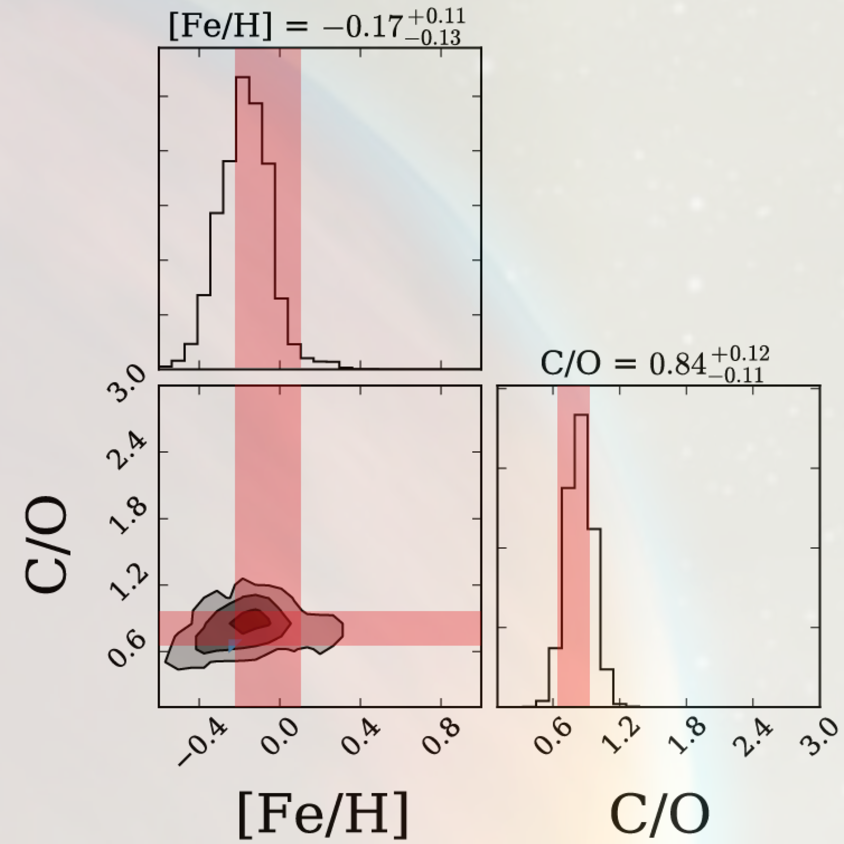

Tool Validation

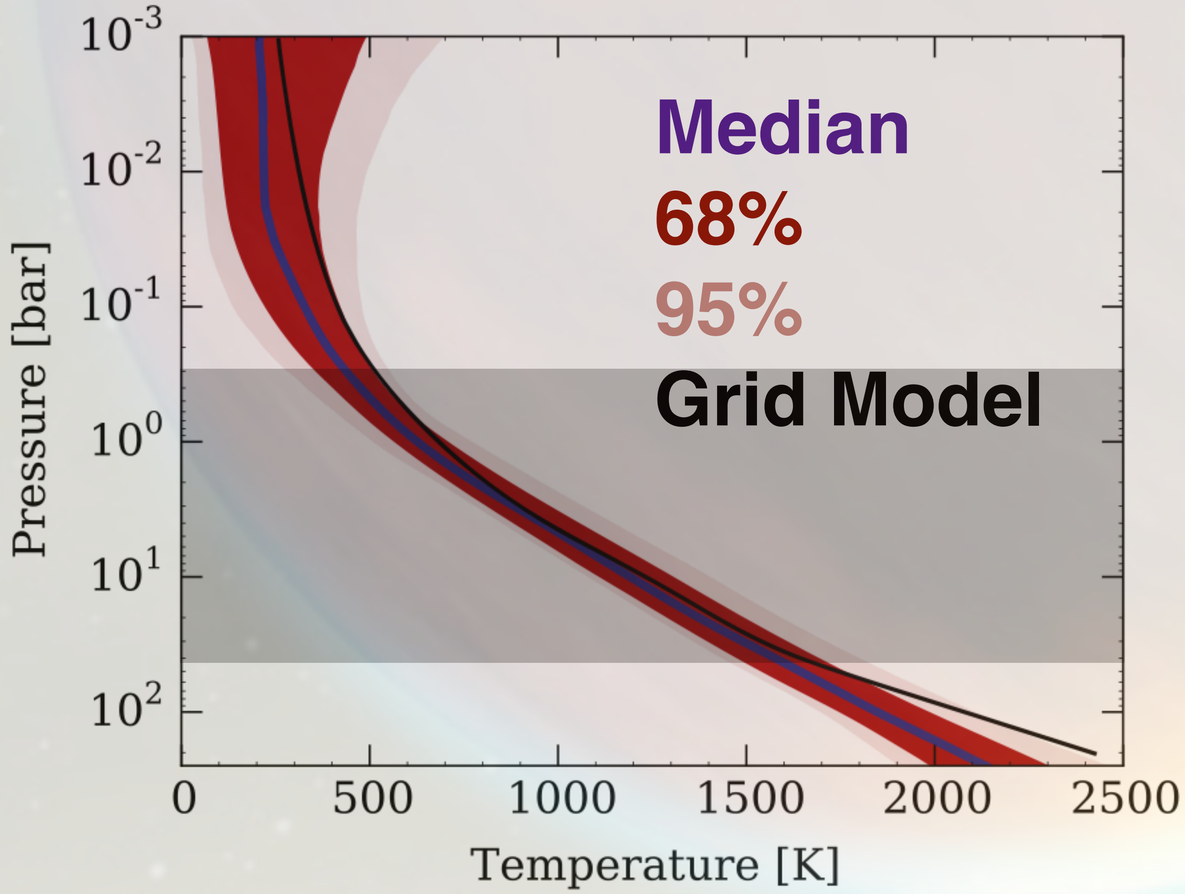
Tool Validation: Solar System

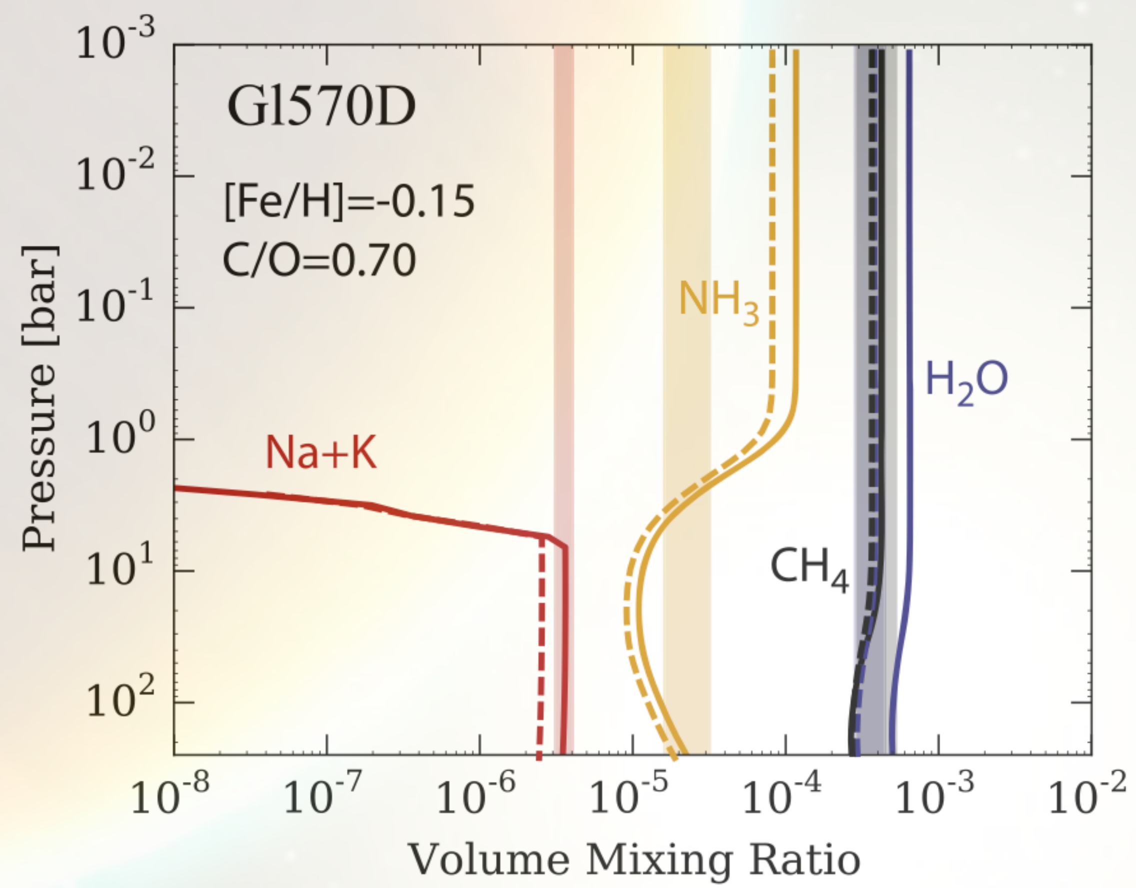


Tool Validation: Brown Dwarfs

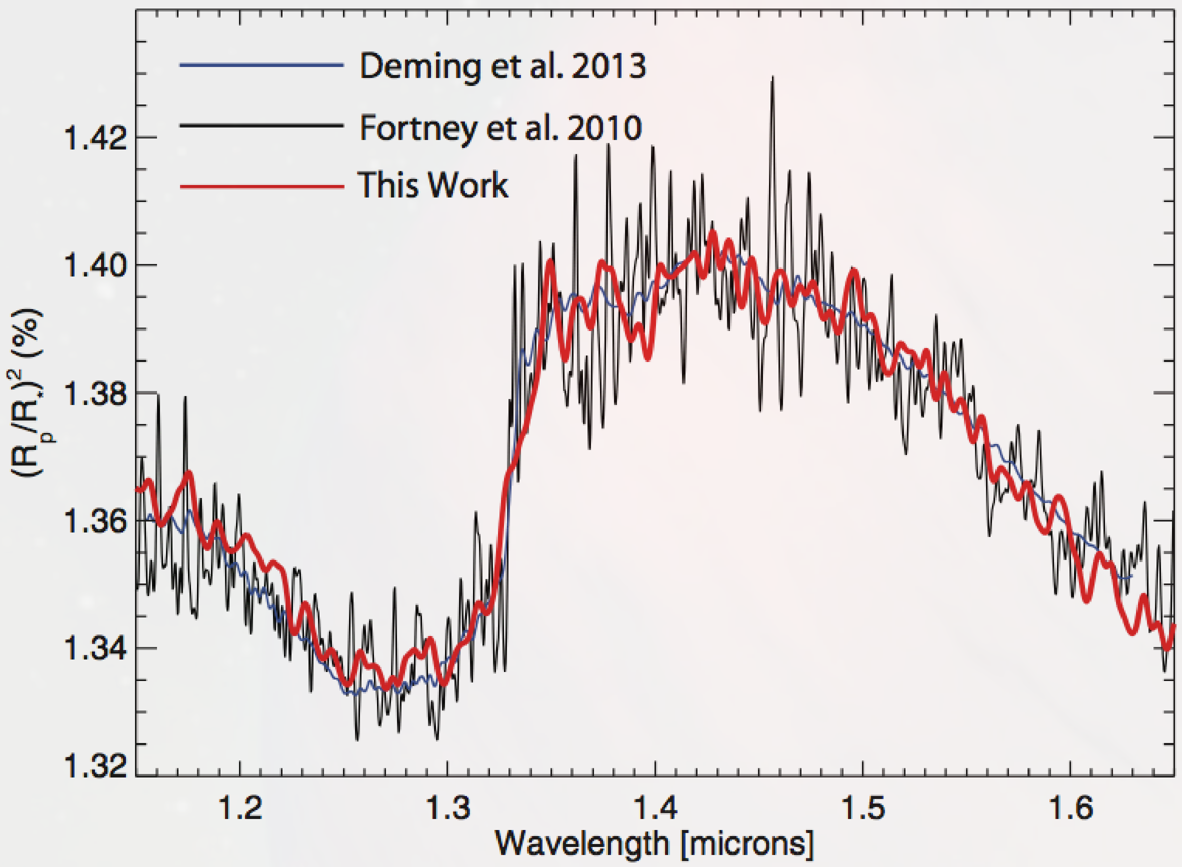


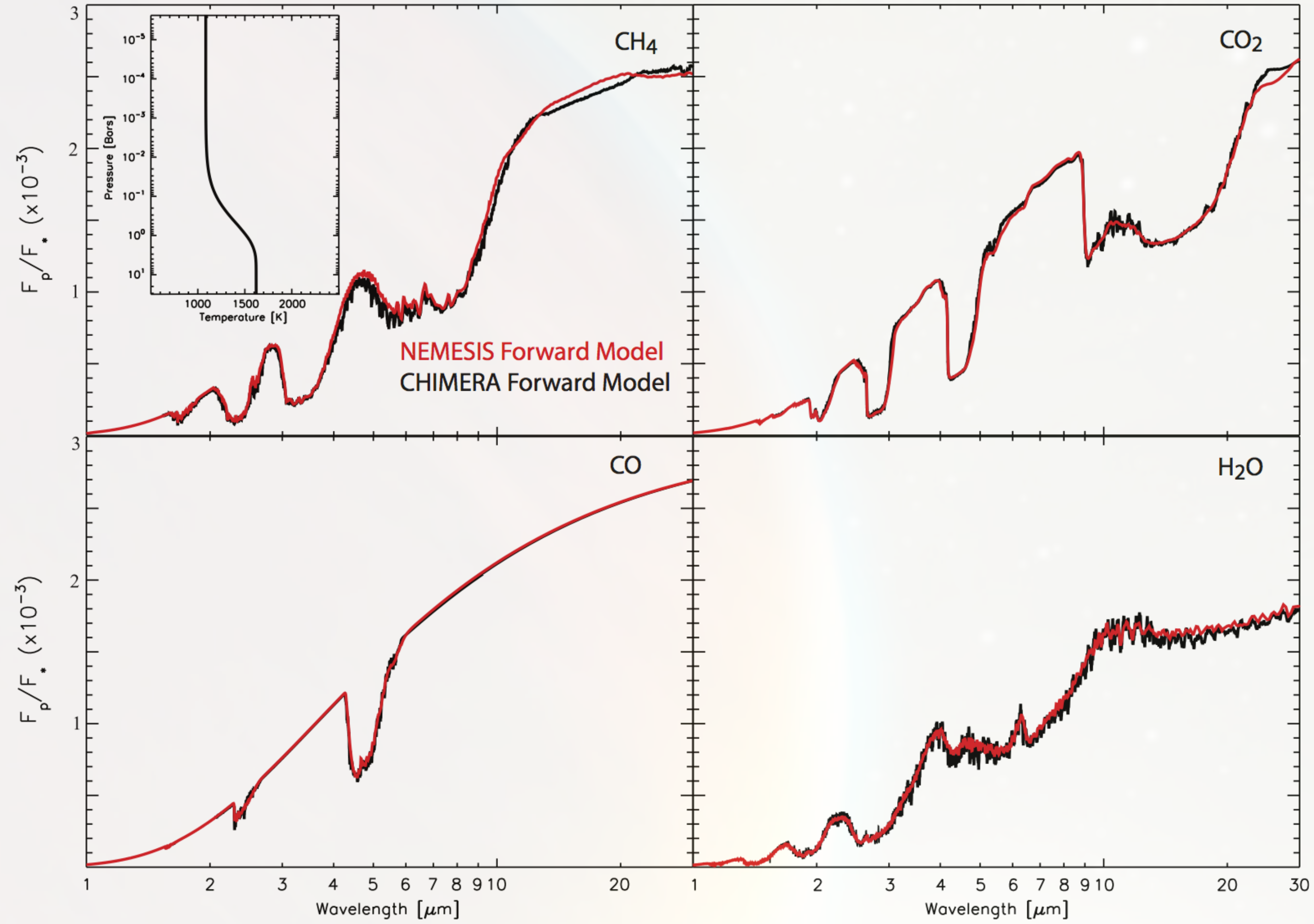


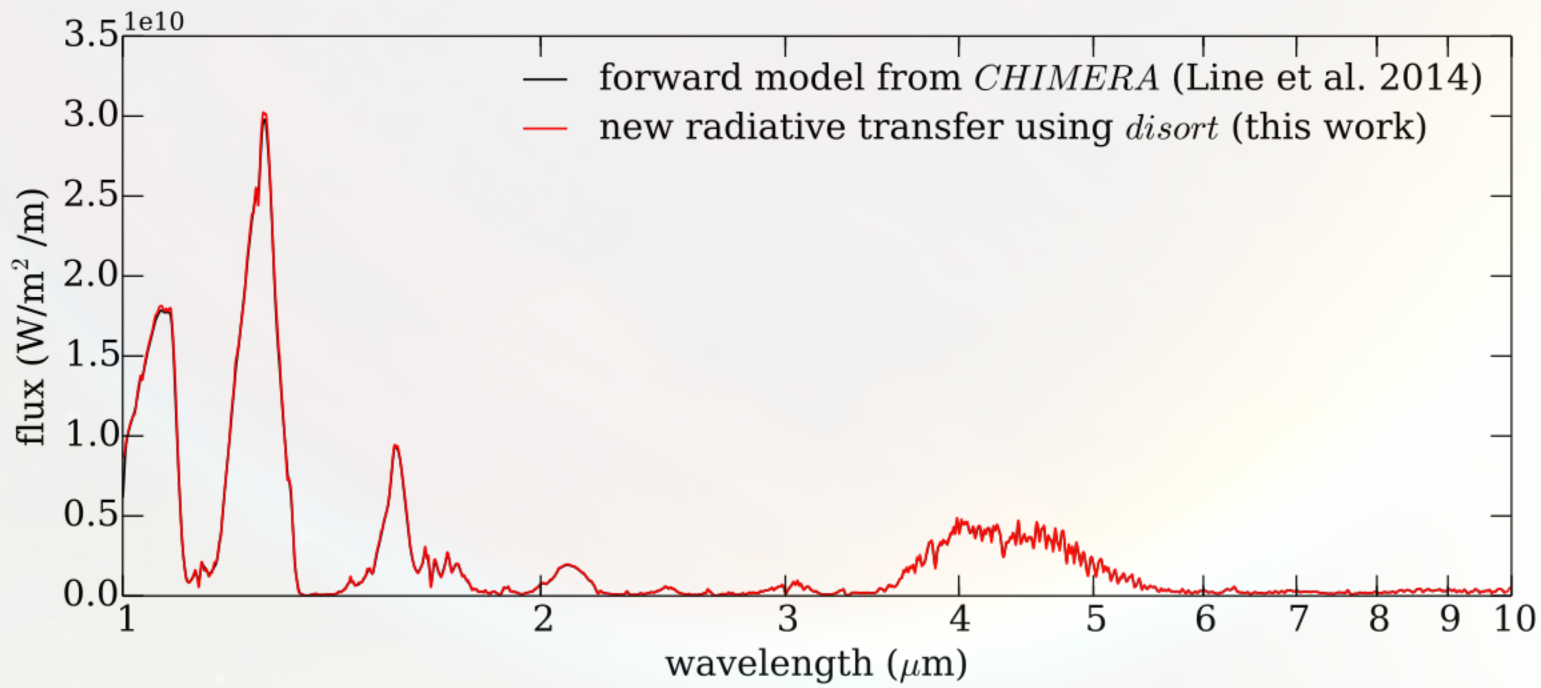




Tool Validation: Model Comparisons

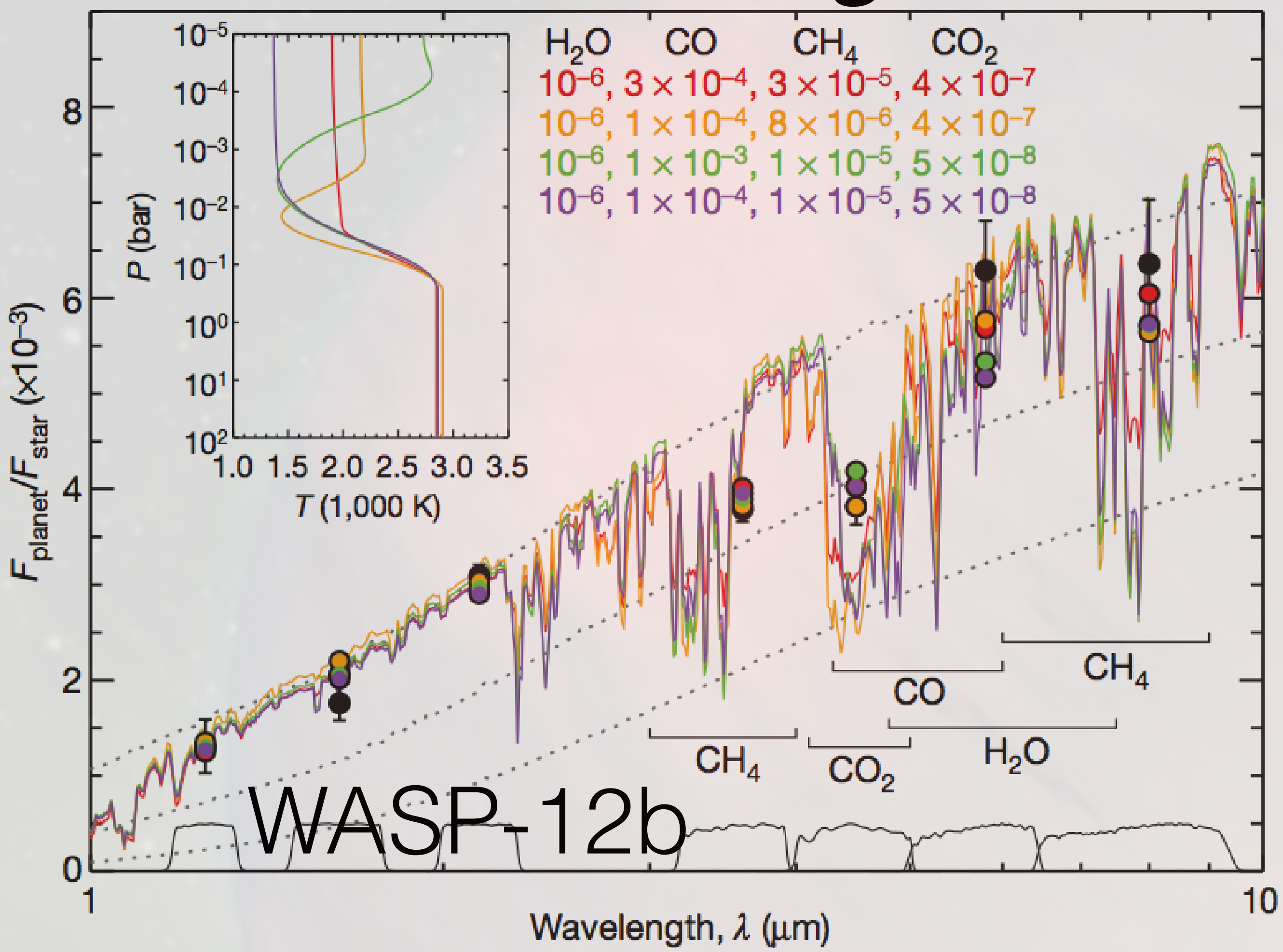


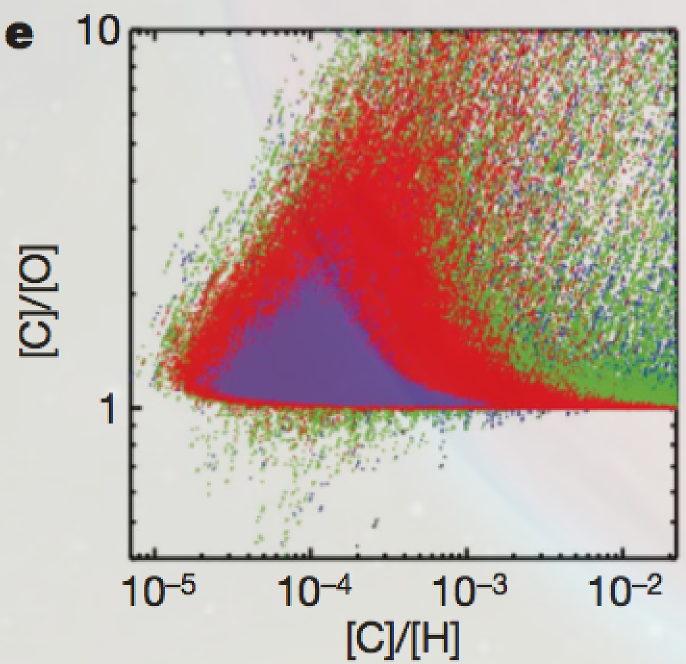





Cool Results

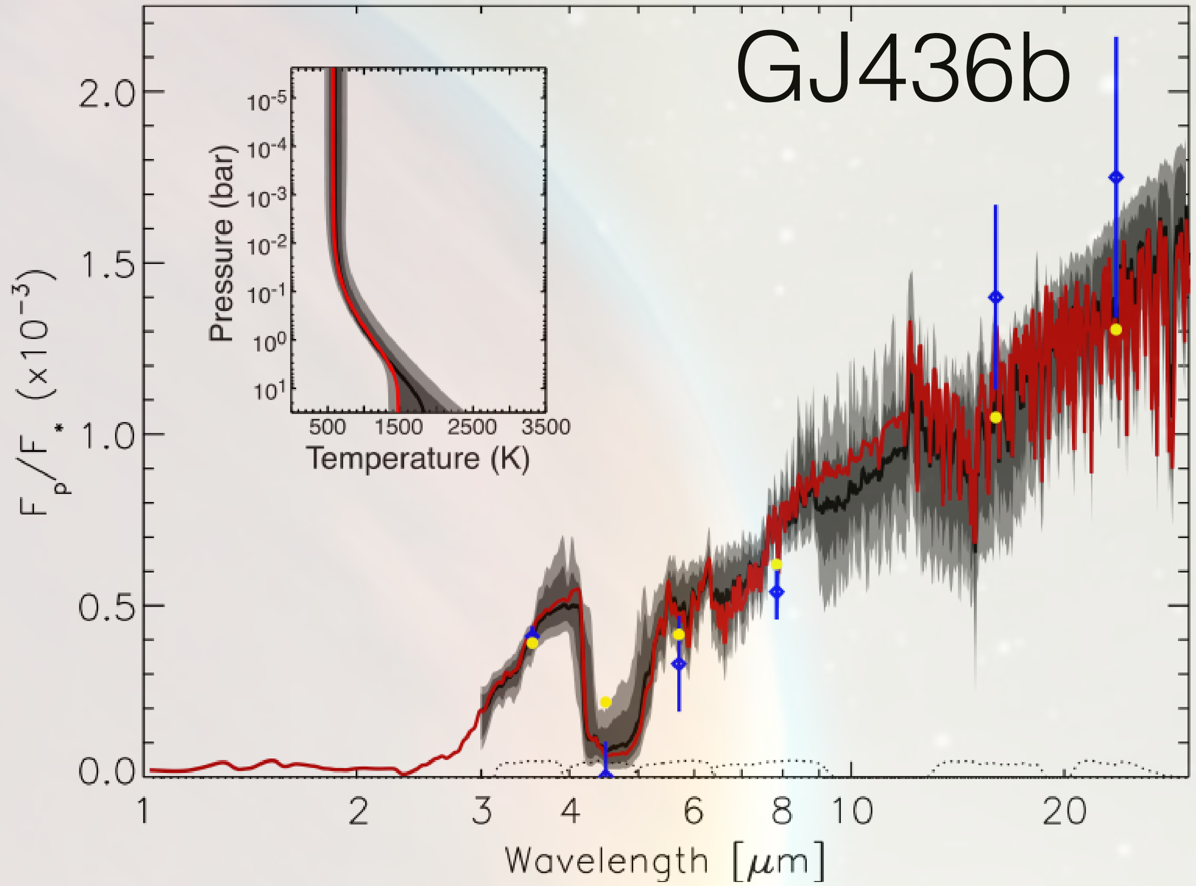
Evidence for High C/O

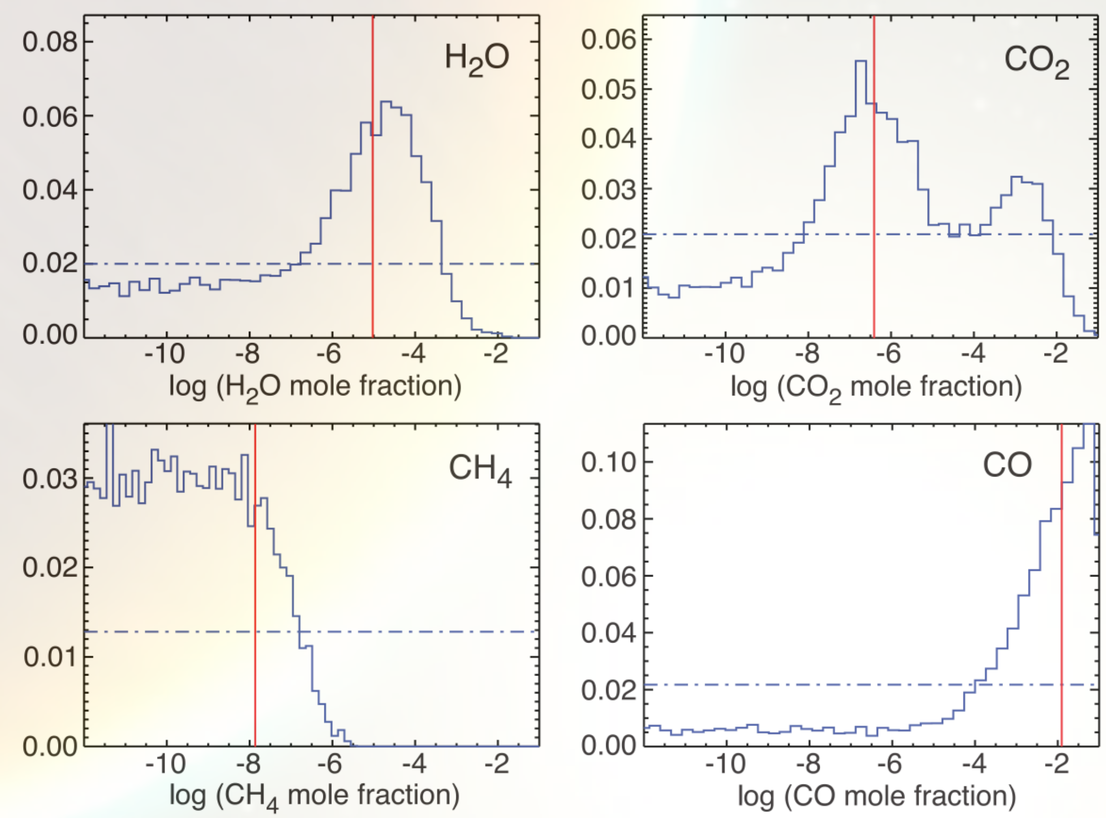




Madhusudhan+2011

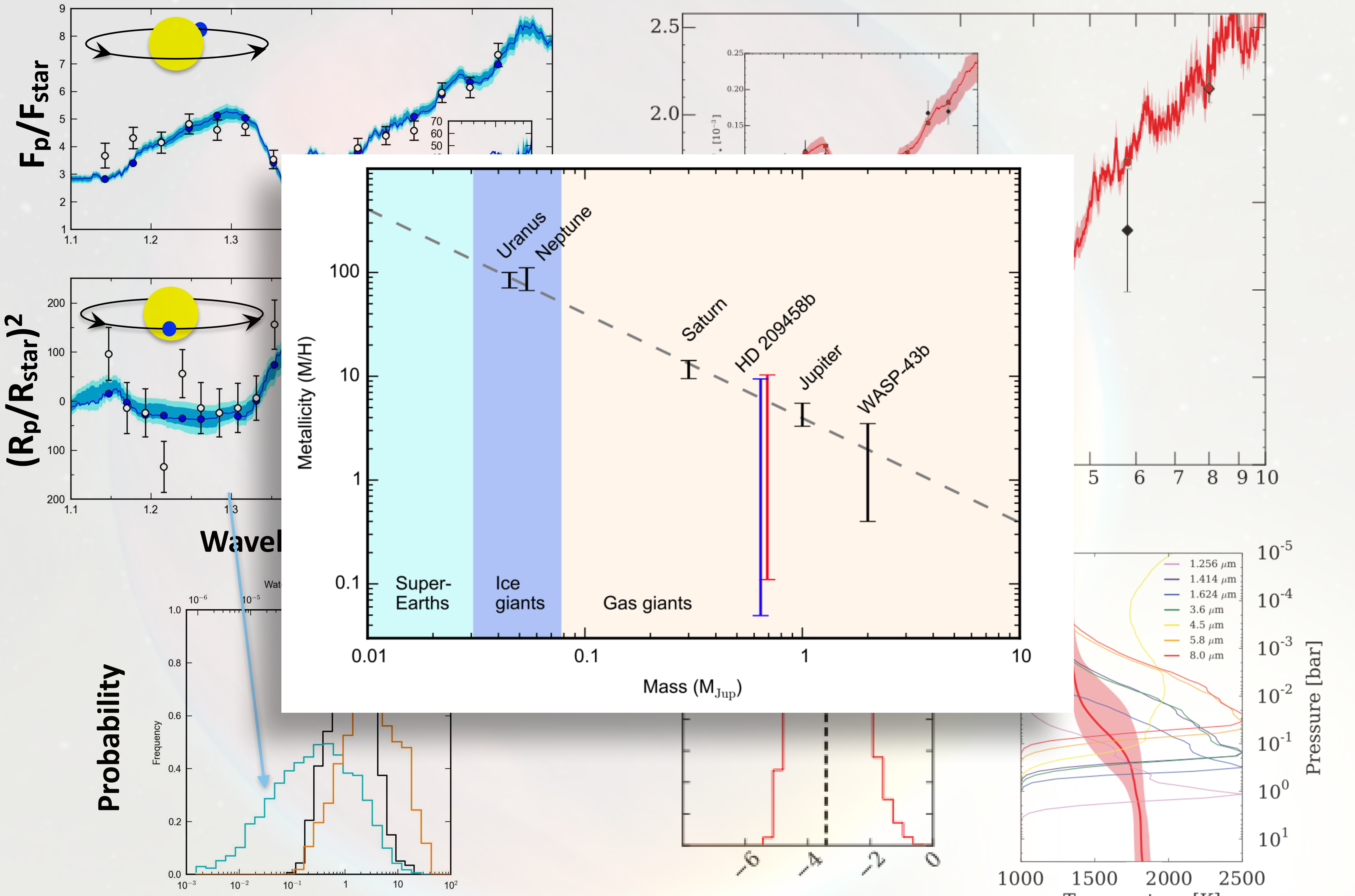
High Metallicity Atmospheres





Moses+2013

Towards "Precision" Abundance Determinations

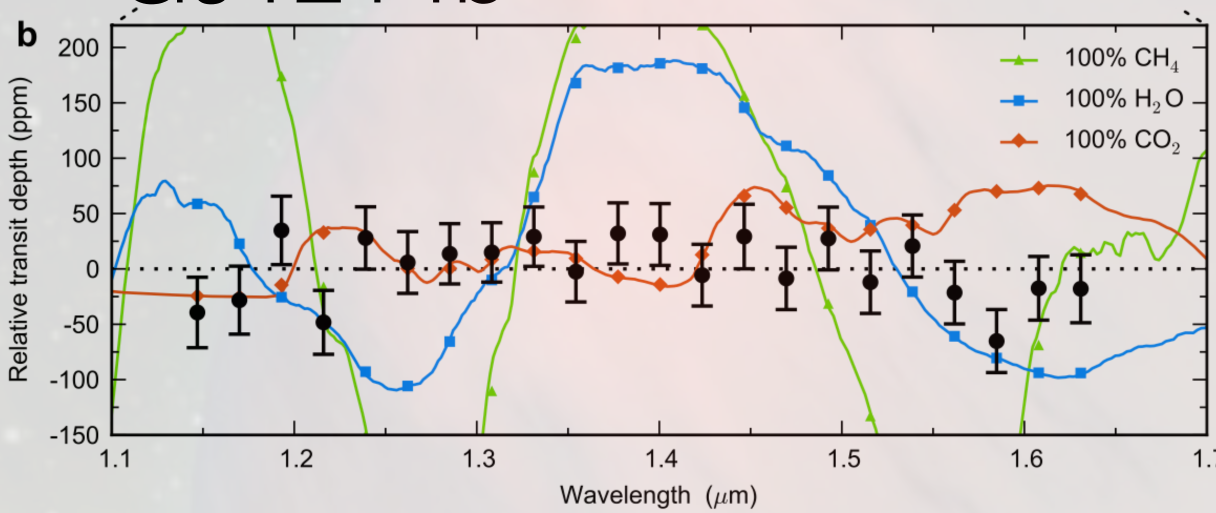


Kreidberg+2014

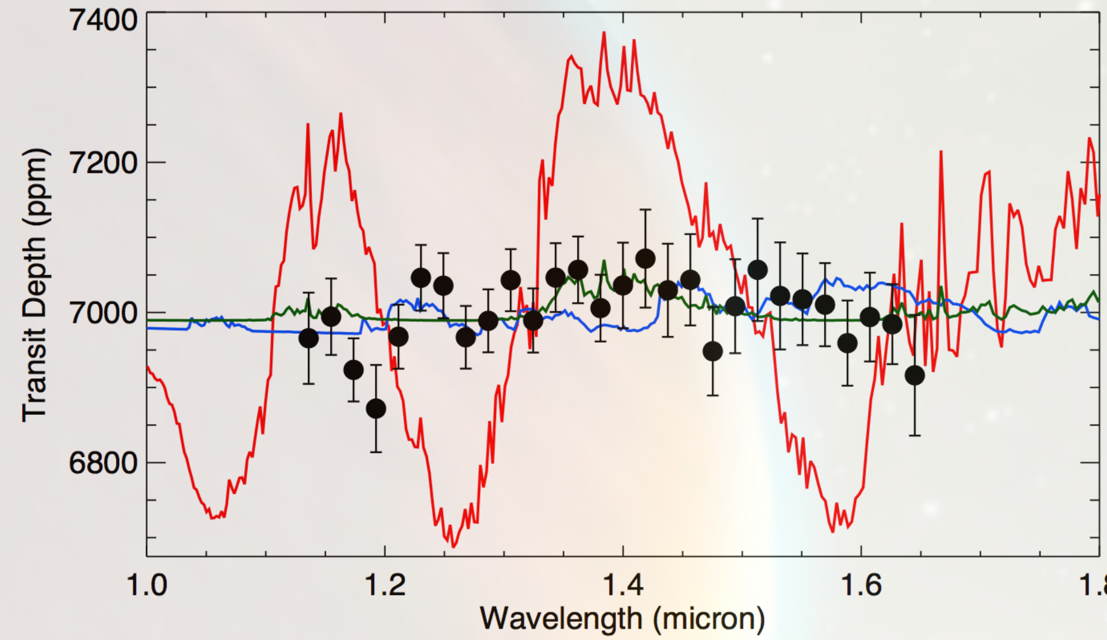
Line+2016

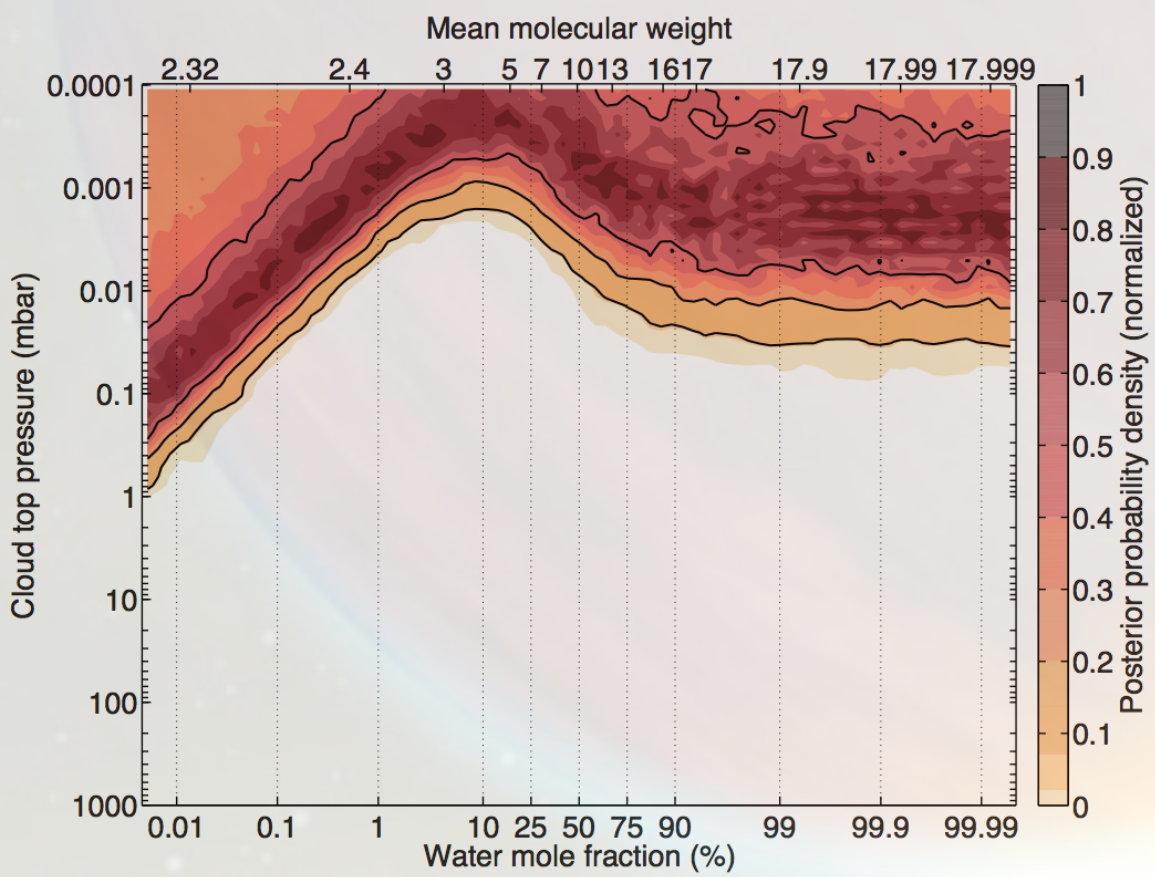
Clouds are just the Worst

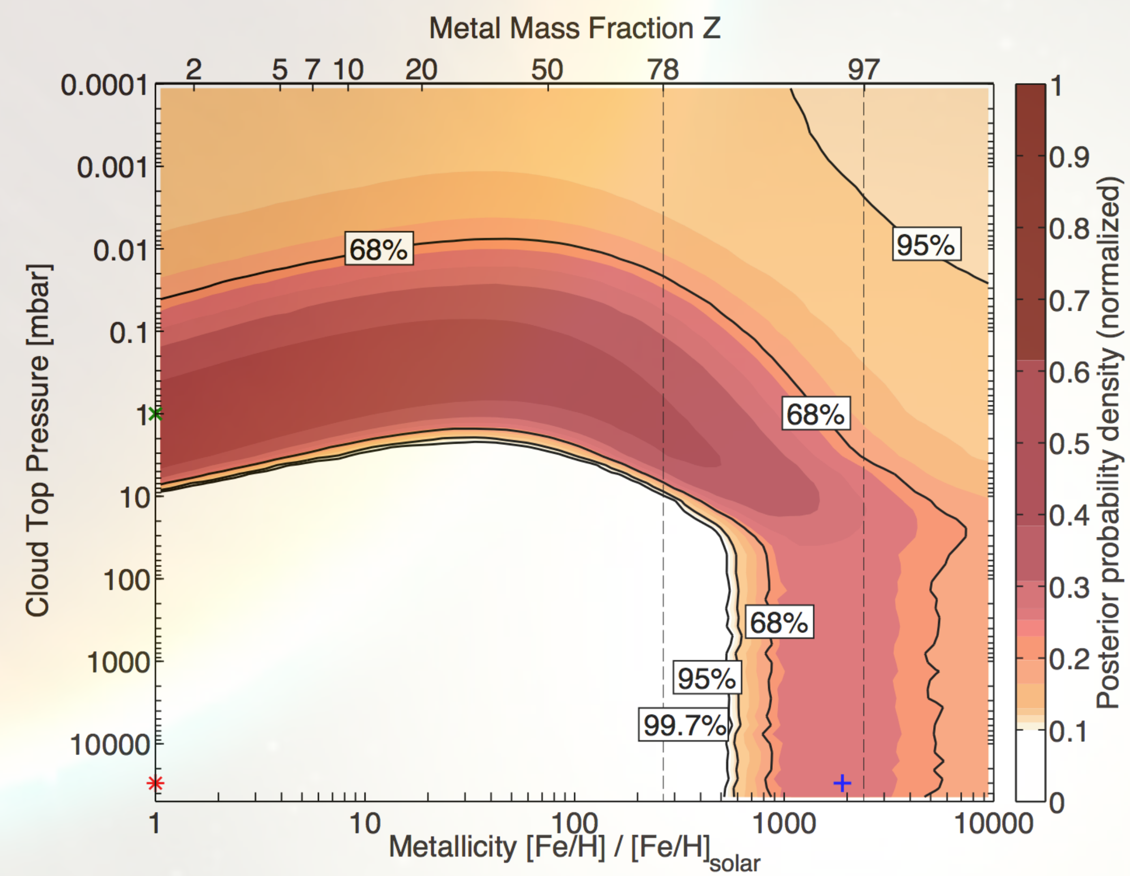
GJ1214b



GJ436b

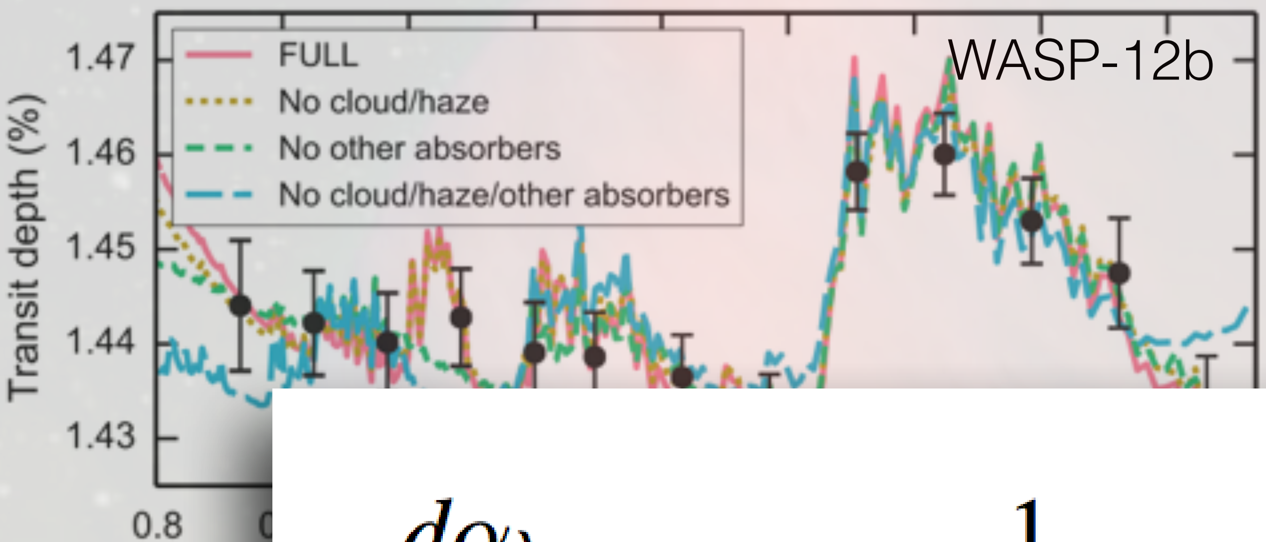


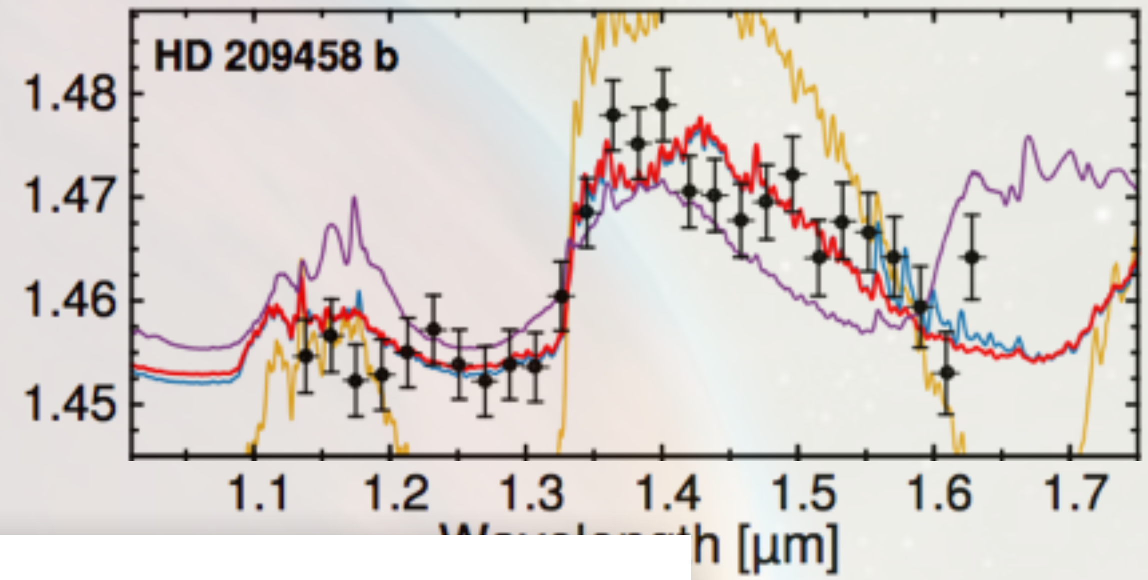




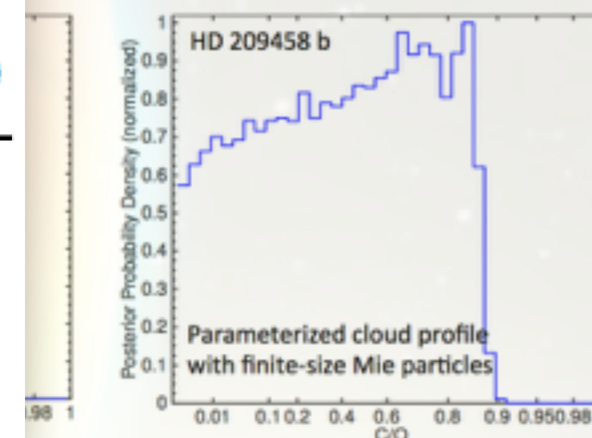
Kriedberg+2014; Knutson+2014

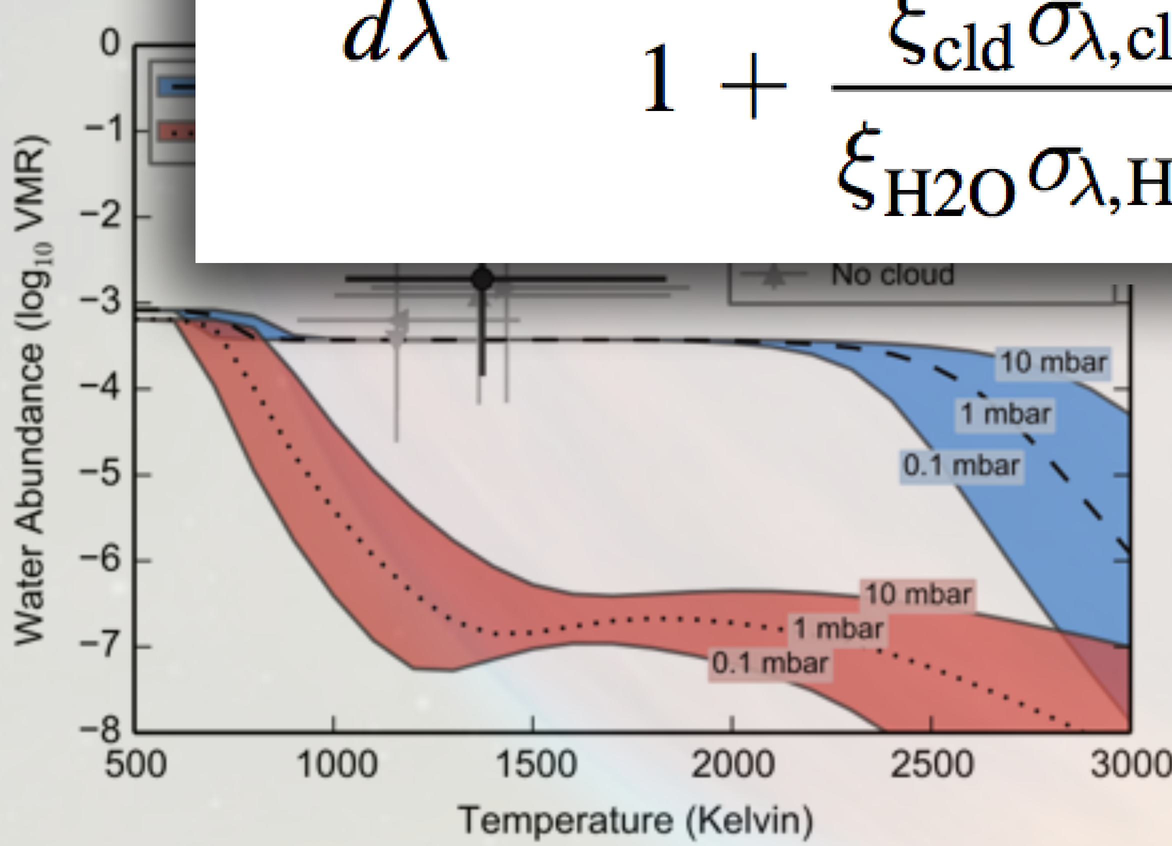
Quantifying Cloud Impacts

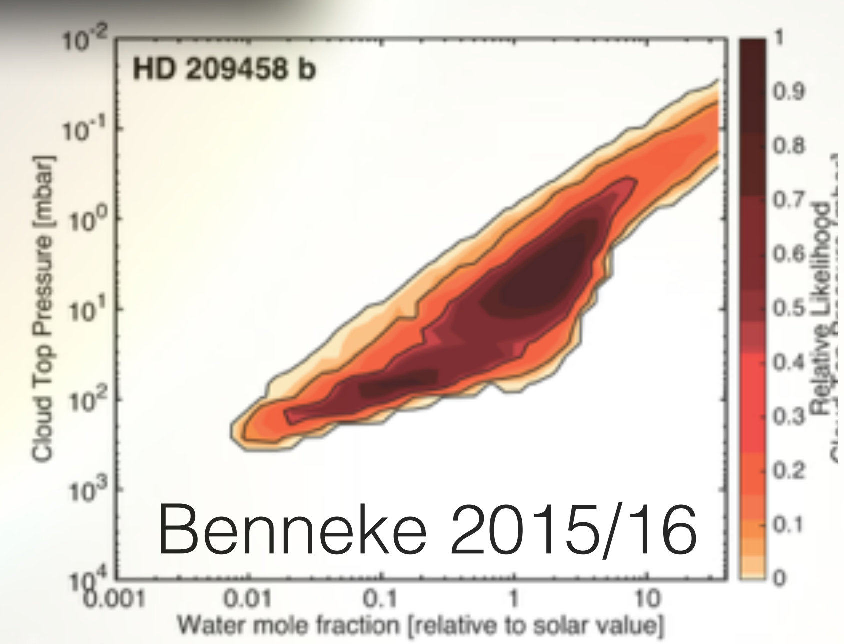




$$\frac{d\alpha_\lambda}{d\lambda} = \frac{1}{1+\frac{\xi_{\mathrm{cld}}\,\sigma_{\lambda,\mathrm{cld}}}{\xi_{\mathrm{H2O}}\,\sigma_{\lambda,\mathrm{H2O}}}}\frac{2R_p}{R_{\mathrm{star}}^2}H\frac{d\ln(\sigma_{\lambda,\mathrm{H2O}})}{d\lambda}$$

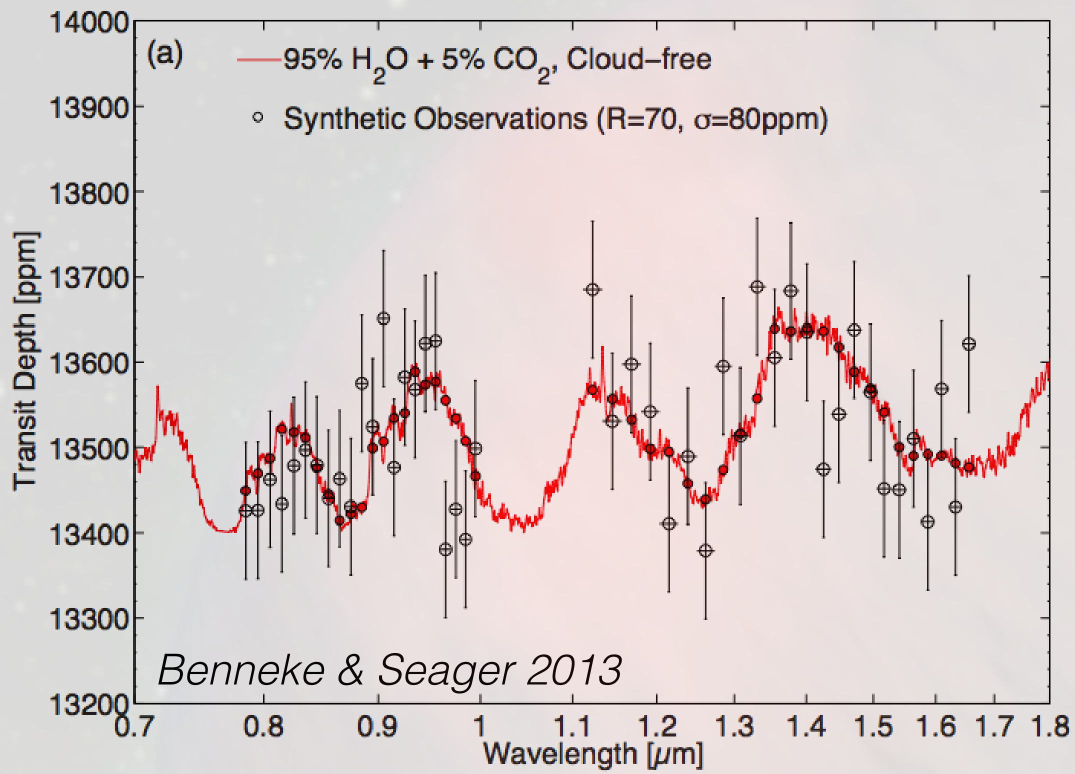


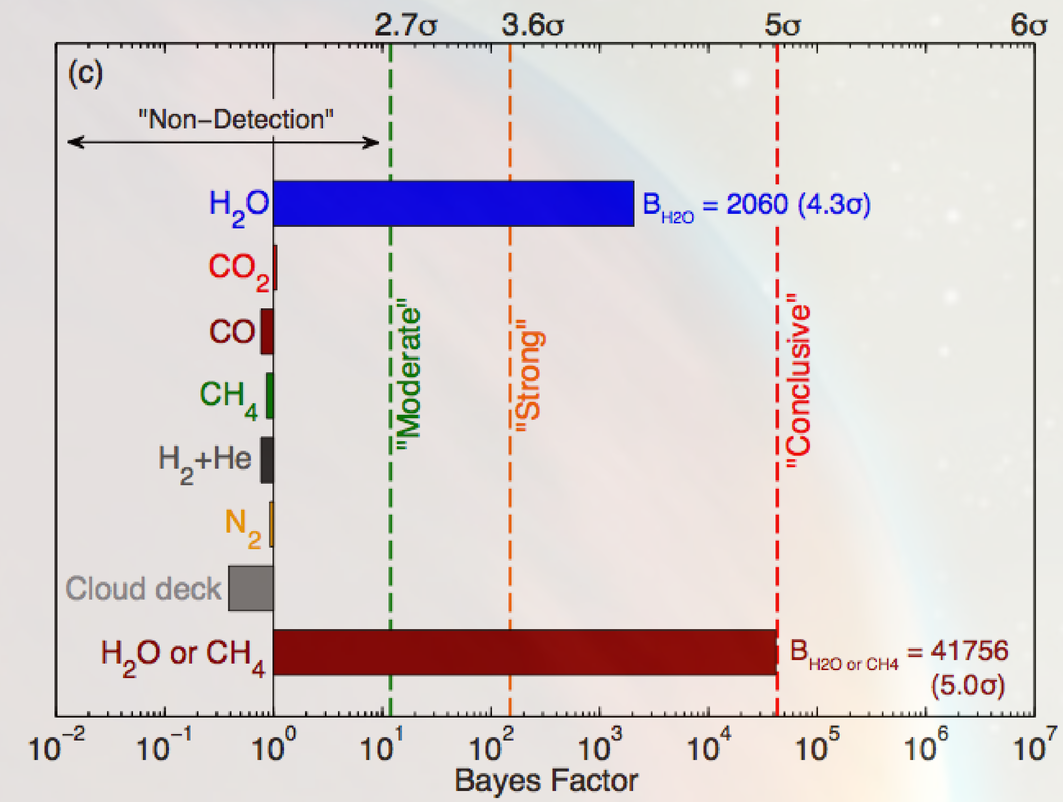


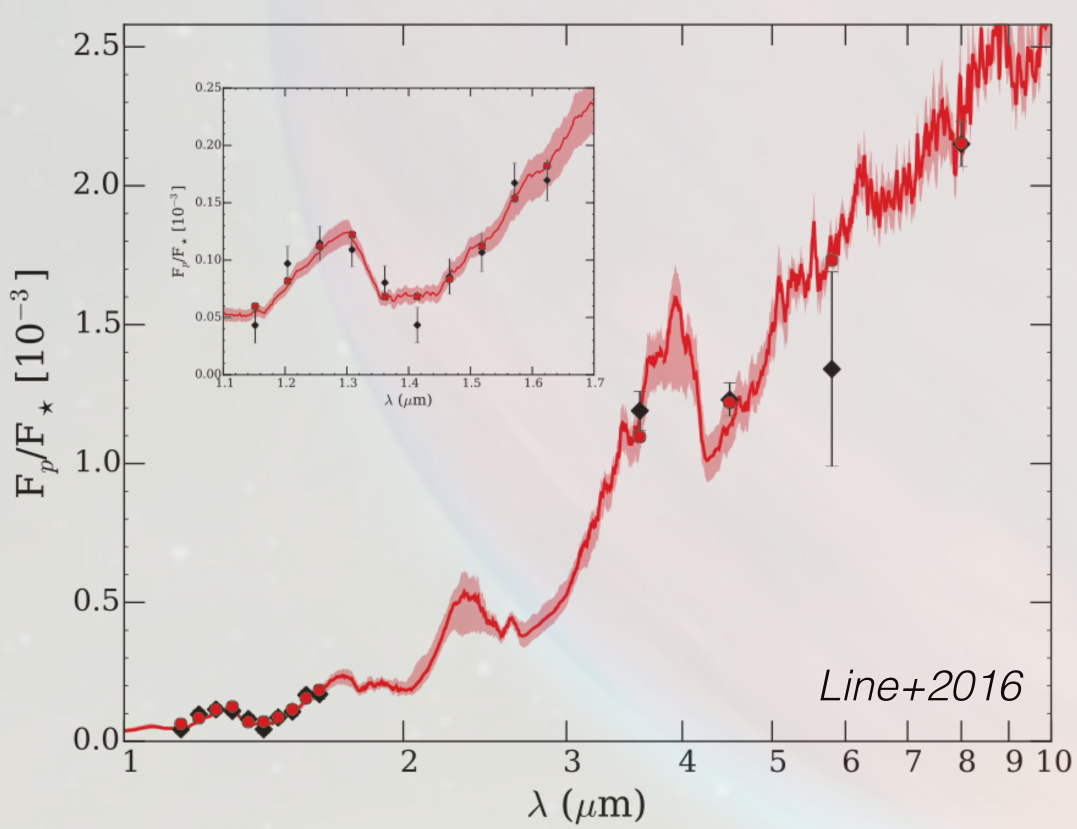


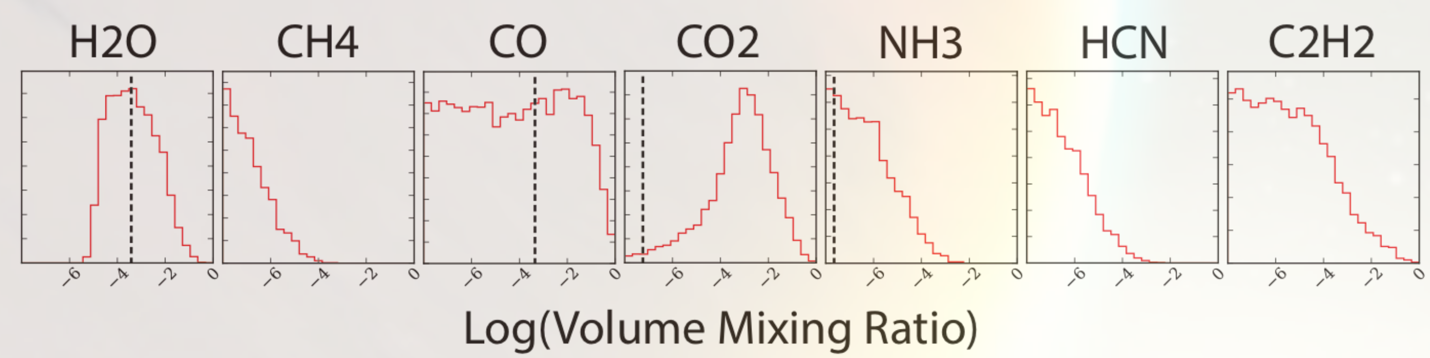
Kreidberg+2015

Is there a Molecule in my Spectrum?







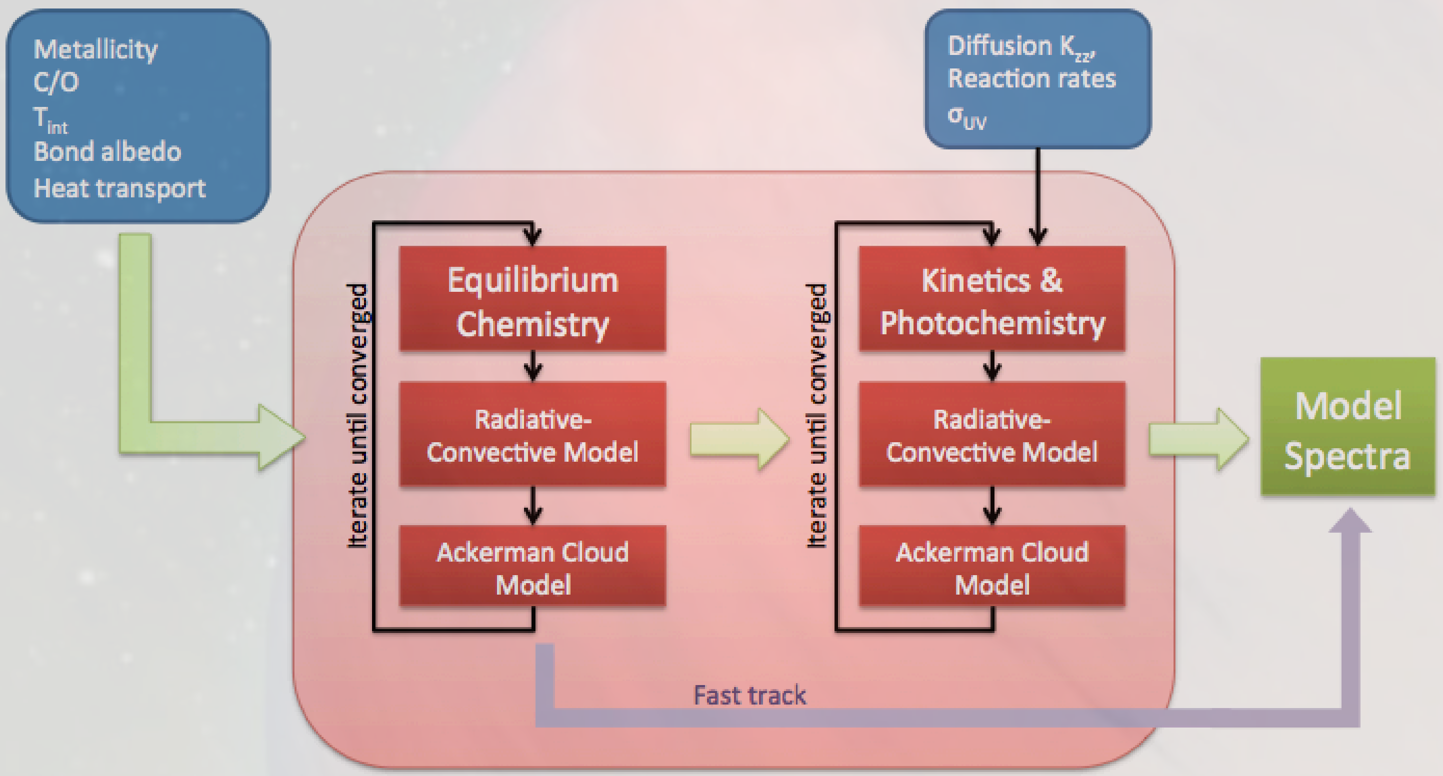


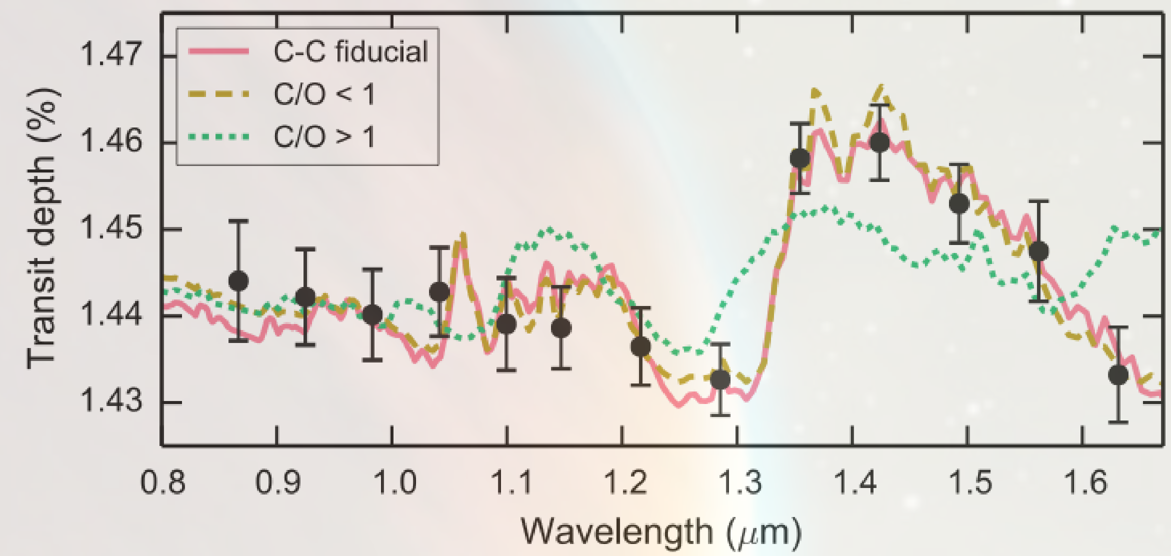
Scenario (included gases)	ln B	Det. Sig. (σ)
FULL (all gases)	-	-
Molecules other than Water (H_2O)	3.96	3.3
CO & CO_2 (H_2O, CH_4, NH_3, HCN, C_2H_2)	6.90	4.1
CH_4 & NH_3 (H_2O, CO, CO_2, HCN, C_2H_2)	-0.92	undefined
HCN & C_2H_2 (H_2O, CH_4, CO, CO_2, NH_3)	-0.71	undefined
H_2O (CH_4, CO, CO_2, NH_3, HCN, C_2H_2)	17.10	6.2
Monotonically Decreasing T (all gases)	27.10	7.7

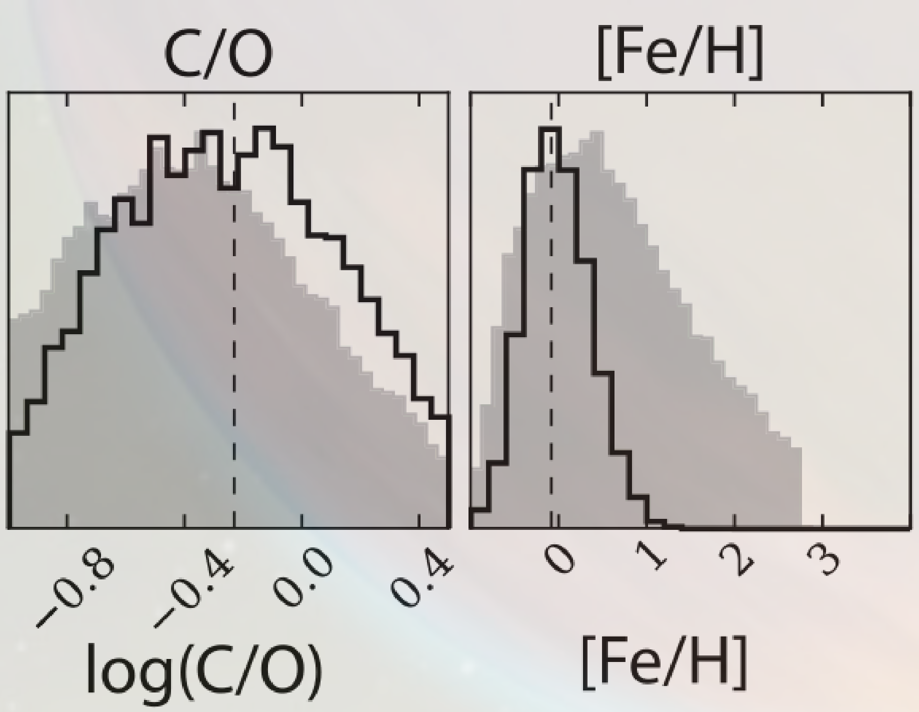
See also: Swain+2013;2014; Fraine+2014; Kreidberg+2014; 2015;Stevenson+2014; Waldmann+2015a,b; Line & Parmentier 2016

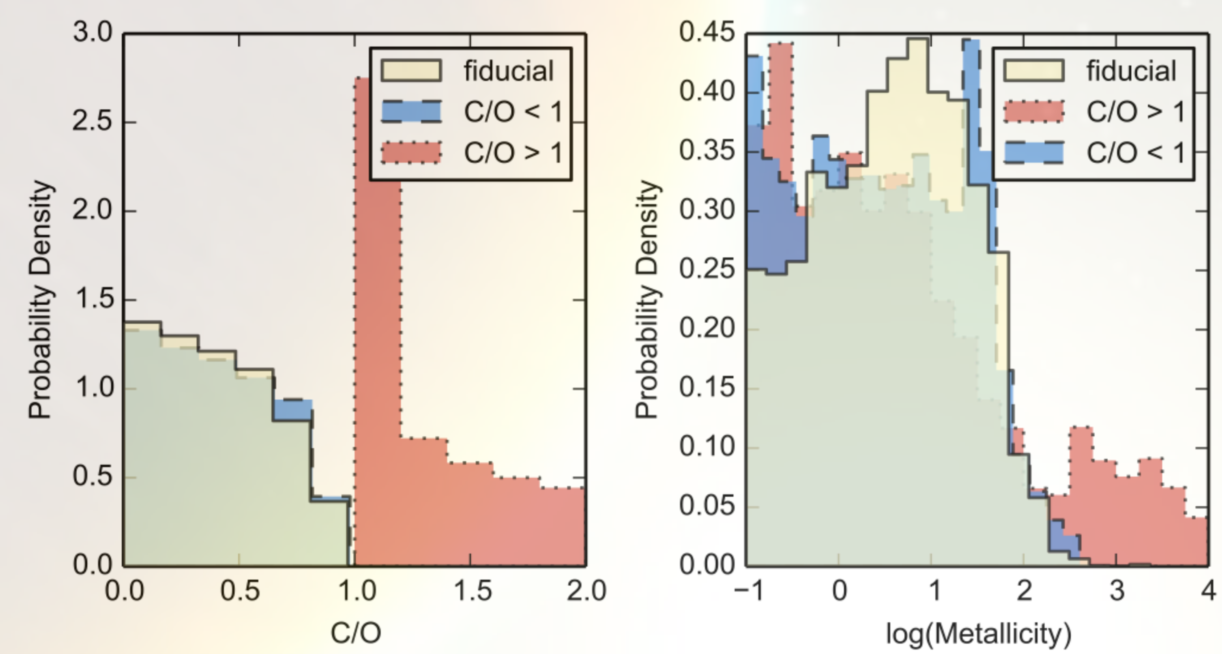
Semi-Self Consistent Methods

(a.k.a boosting "prior" information)



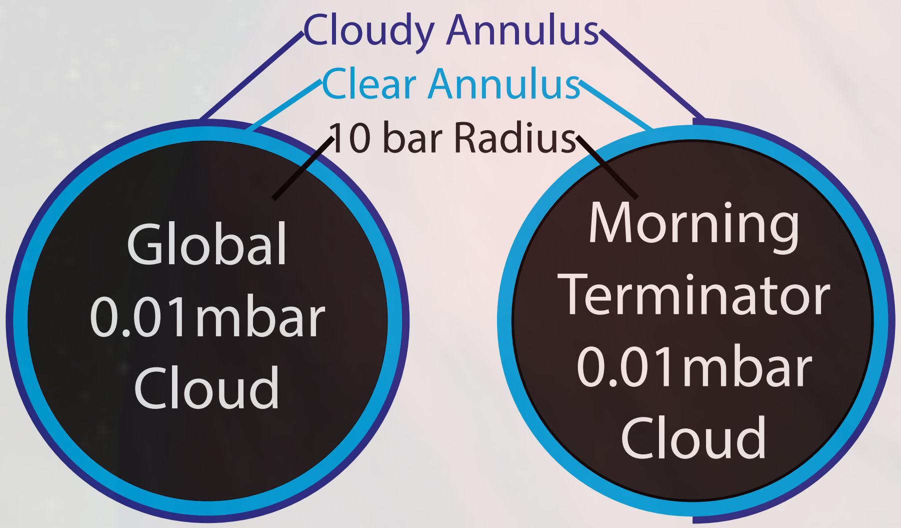


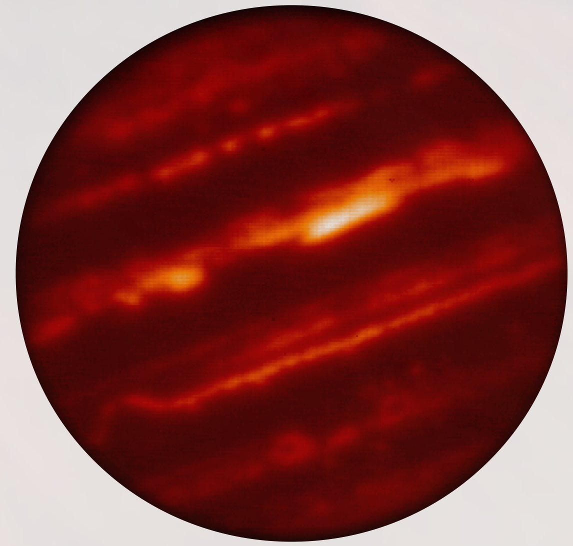


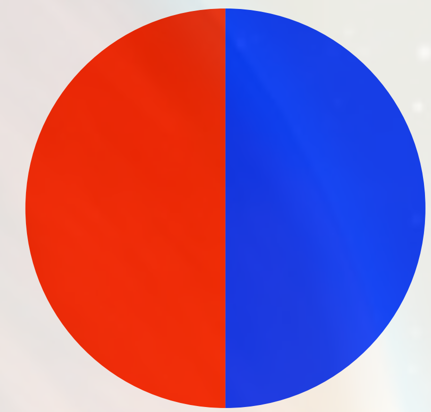


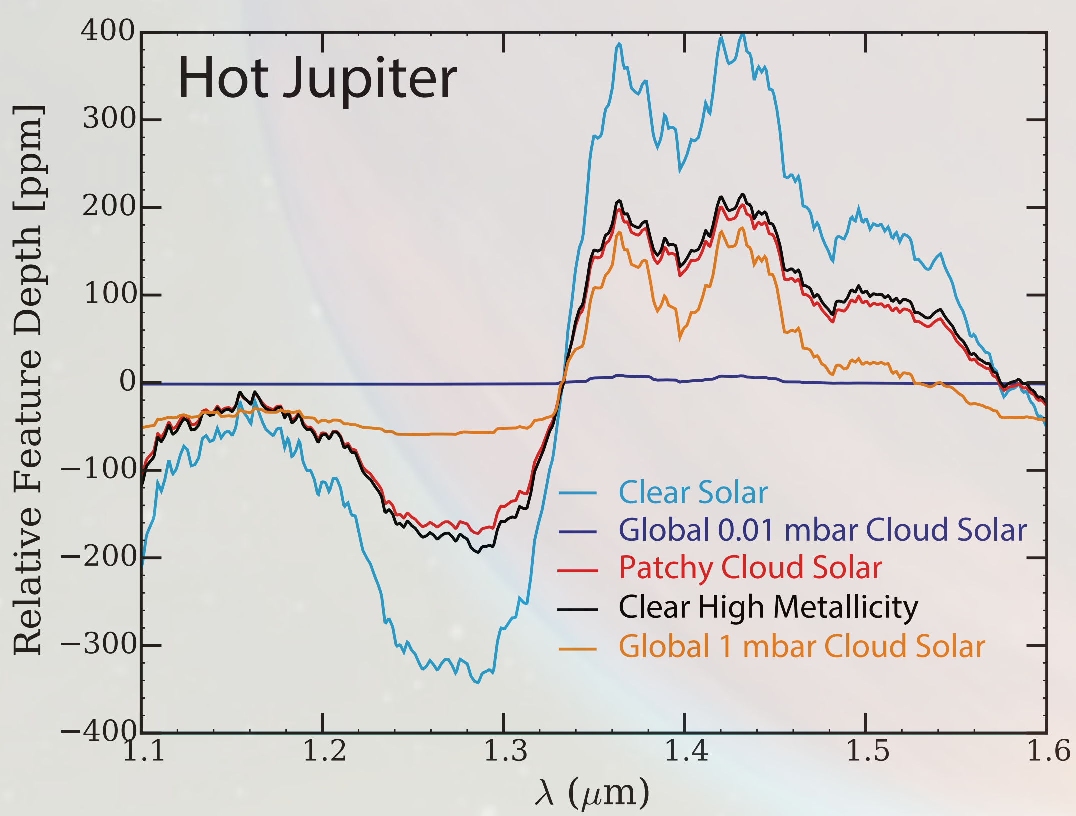
Fraine+2014; Benneke 2015; Kreidberg+2015; Greene+2016

Uh oh…Planets are 3D

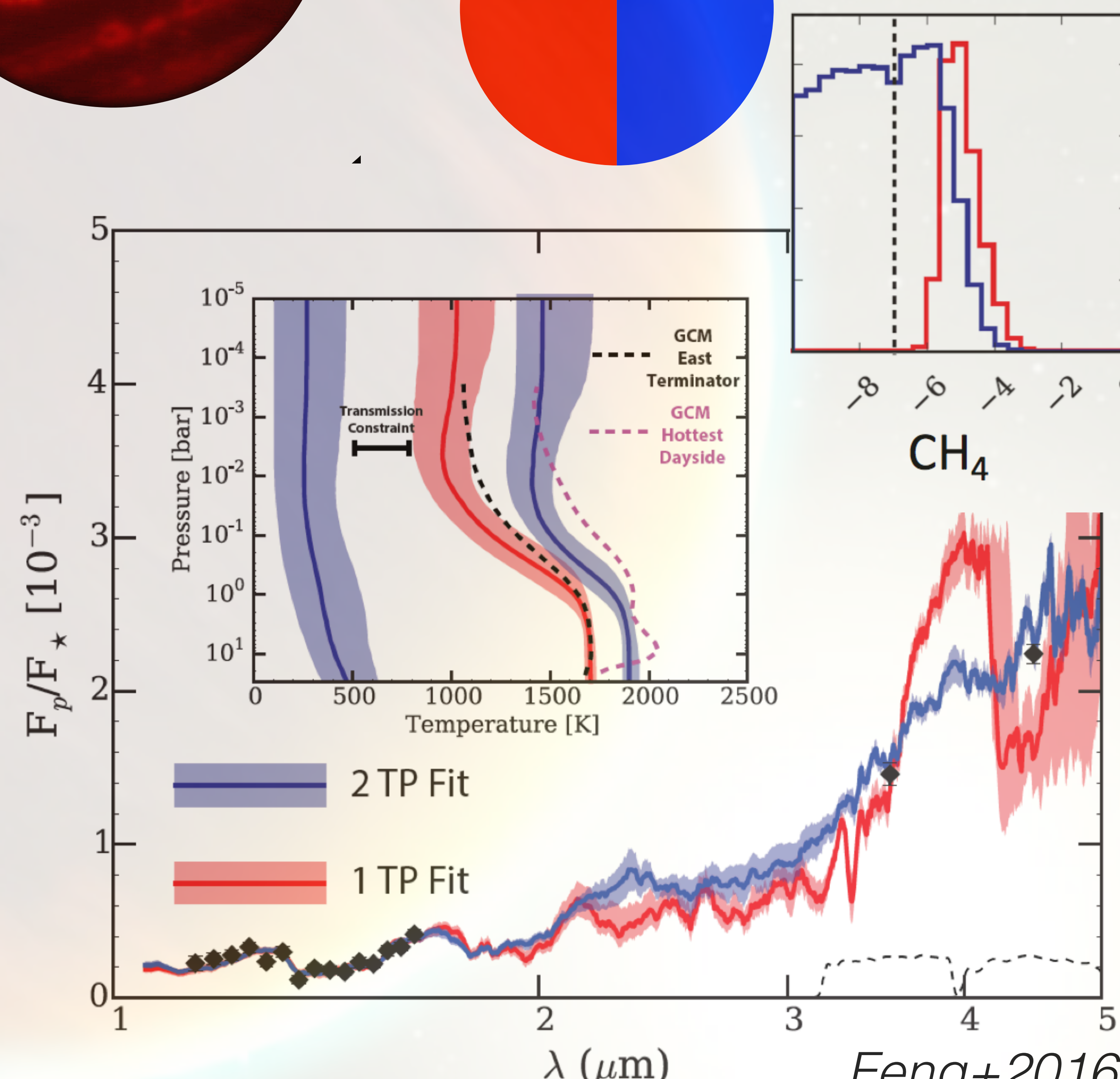








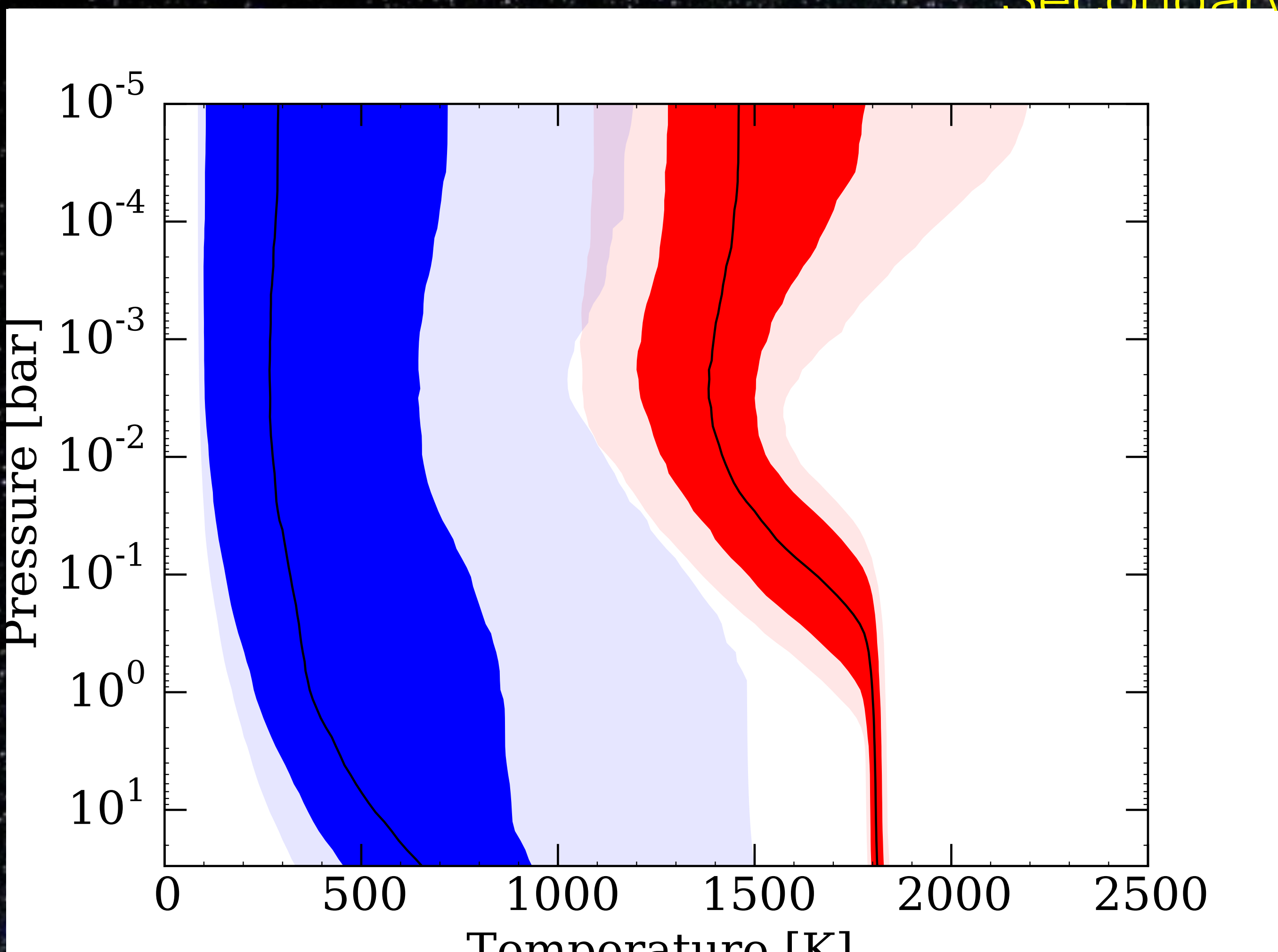
Line & Parmentier 2016

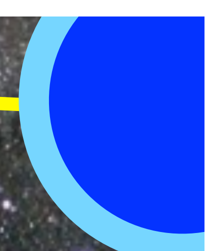


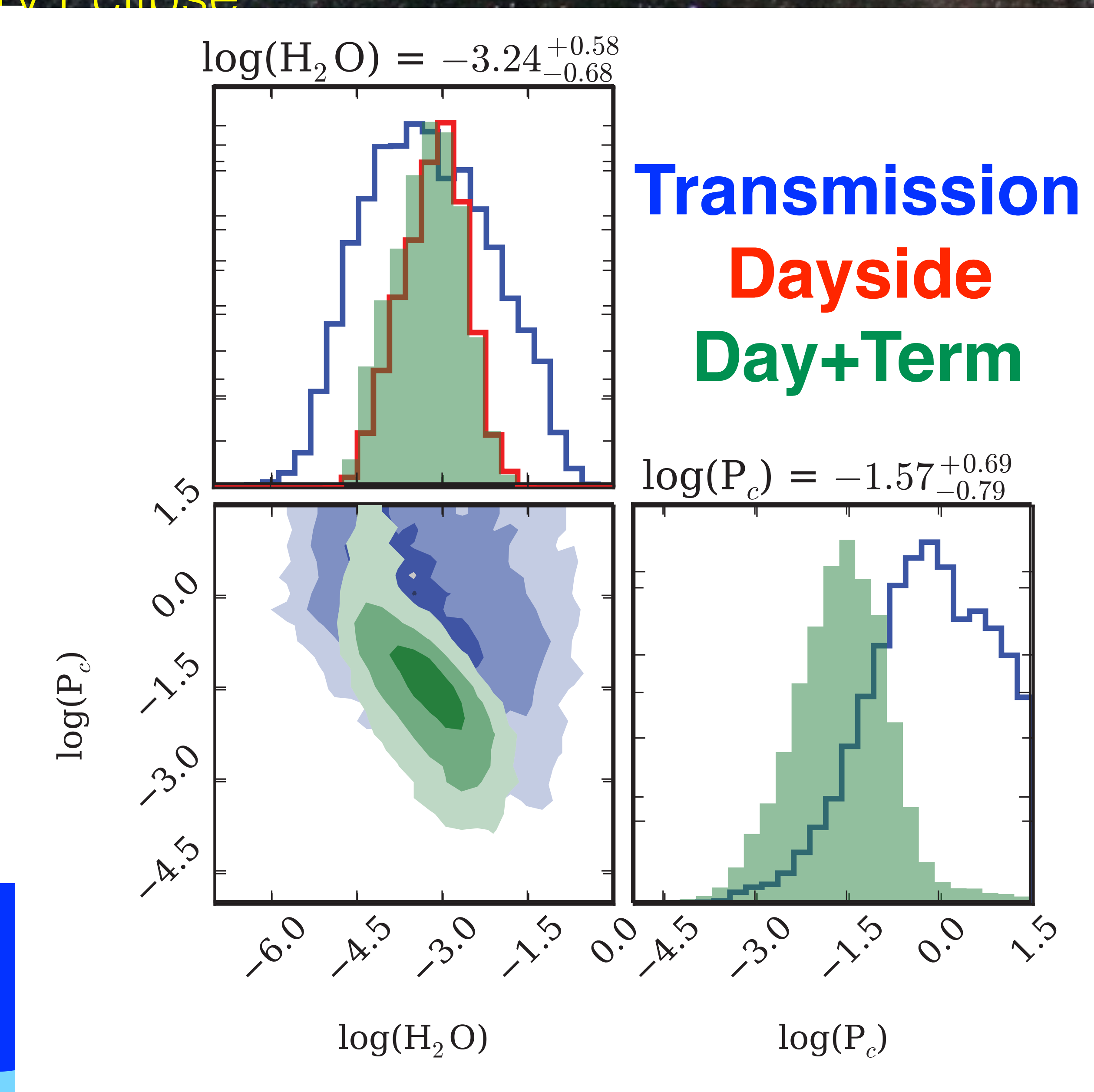
Feng+2016

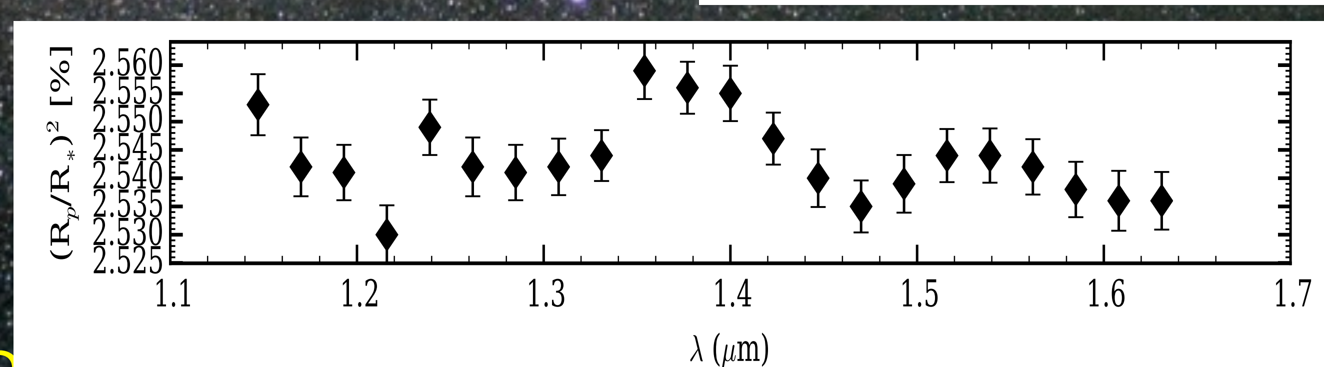
Simultaneous Transmission+Emission Fit

Secondary Eclipse



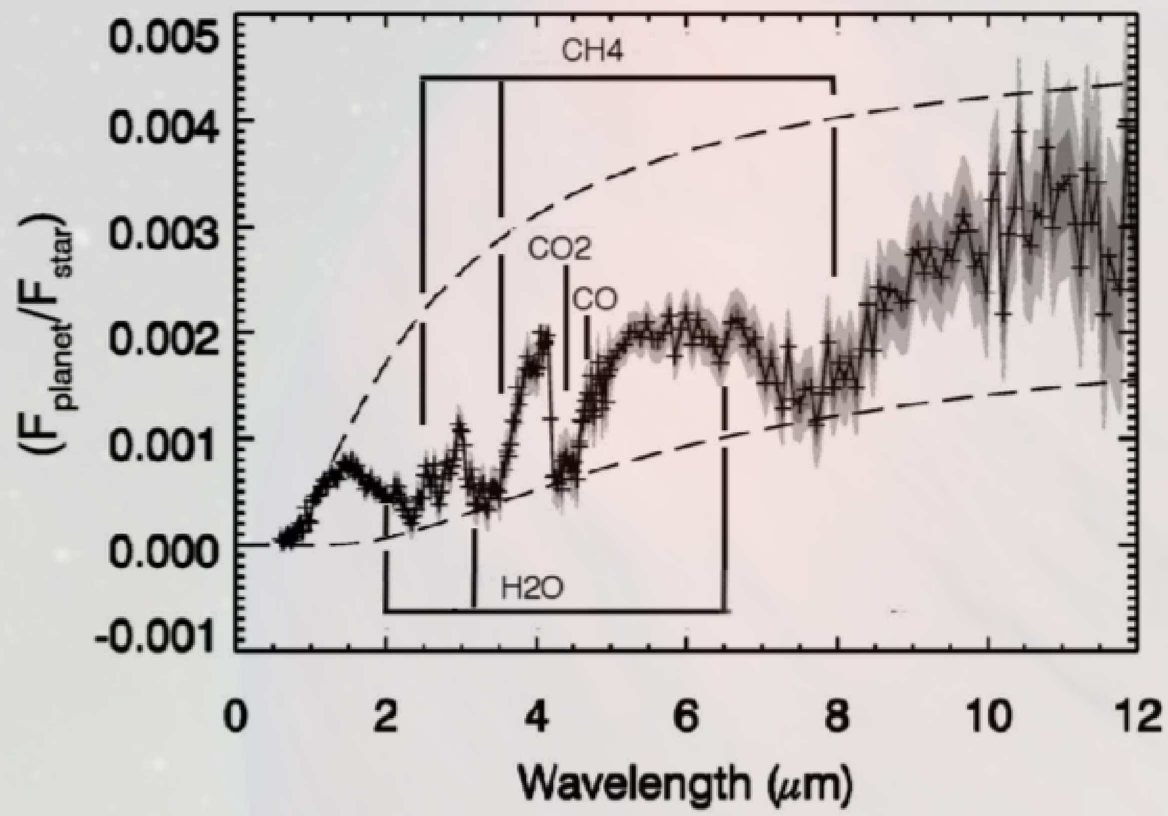


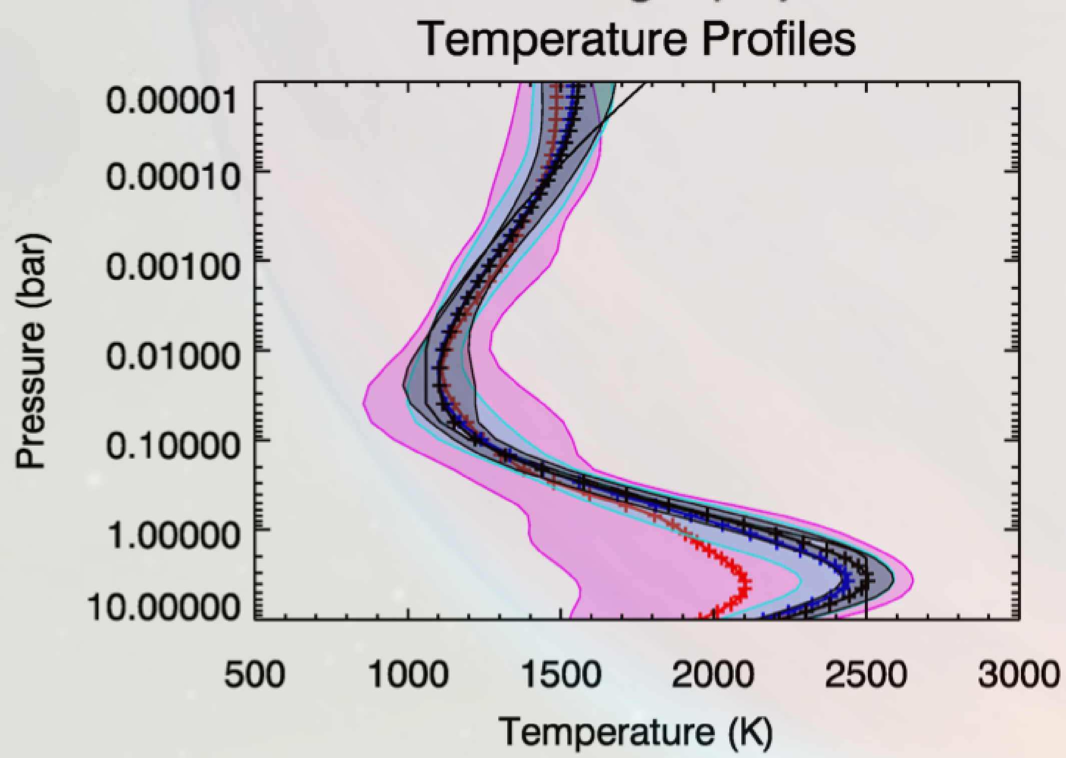


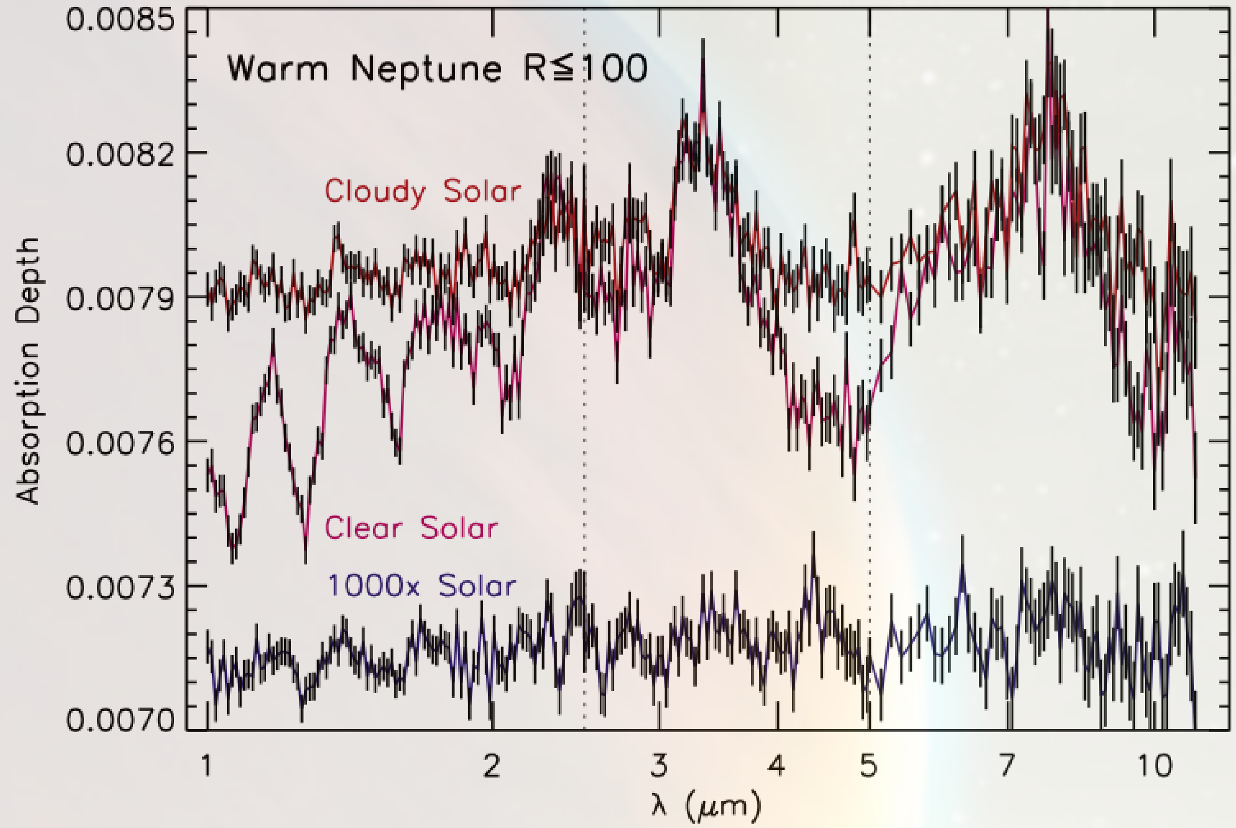


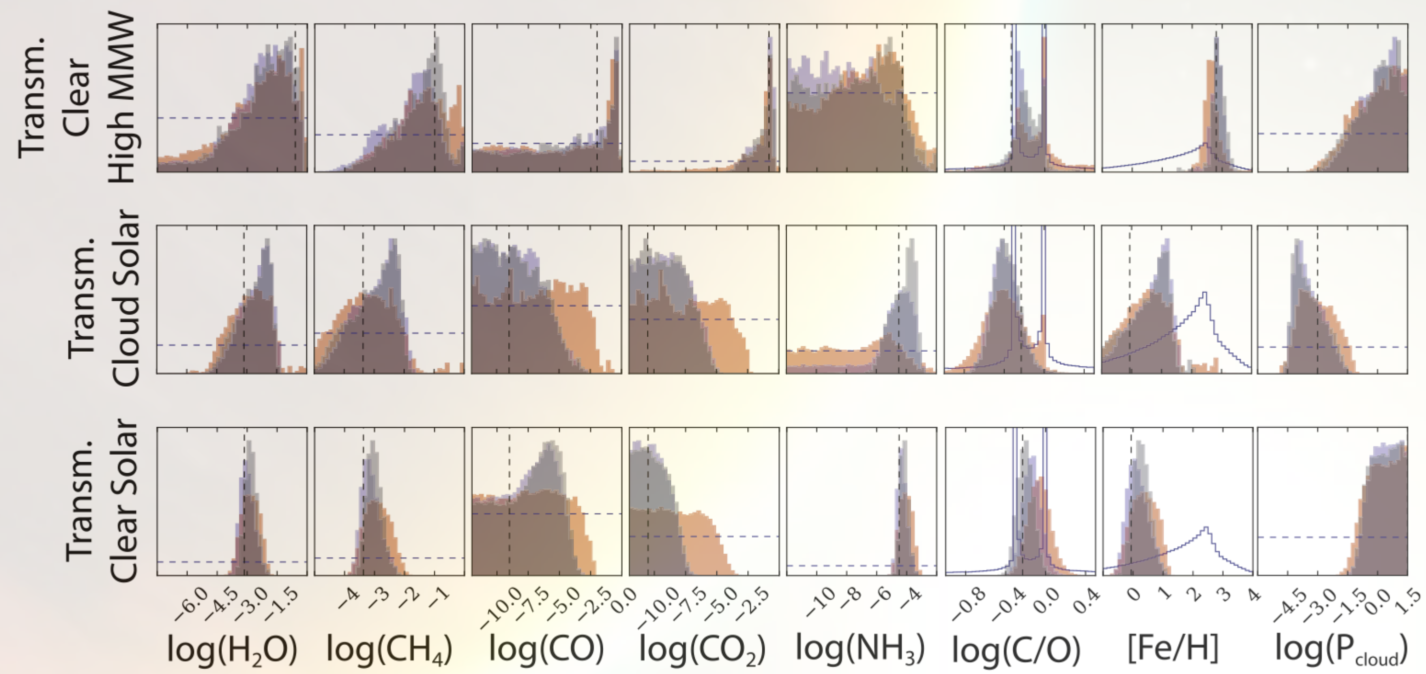
See also Griffith 2014 for preliminary test of this

Future Mission Planning









Barstow+2015

Greene+2016

Philosophies

- Results only as good as the robustness of the data—data not consistent with itself sometimes
- The hidden "prior"—Implicit assumption that forward model is "correct"—model selection helps with this (and is being done…)
- Human Choices matter :(
- Atmospheres are complex—much room for improved parameterizations

Backup

Stellar Uncertainties Matter

WASP-43 (K-star)   GJ436 (M-star)