ML JOURNAL CLUB

OPTIMISING EXOPLANET ATMOSPHERE RETRIEVAL USING UNSUPERVISED MACHINE LEARNING

JJC Hayes et al. (SPEARNET)

MNRAS Volume 494, Issue 3, May 2020
https://doi.org/10.1093/mnras/staa978

MANCHESTER
1824
The University of Manchester

@mosphericHayes

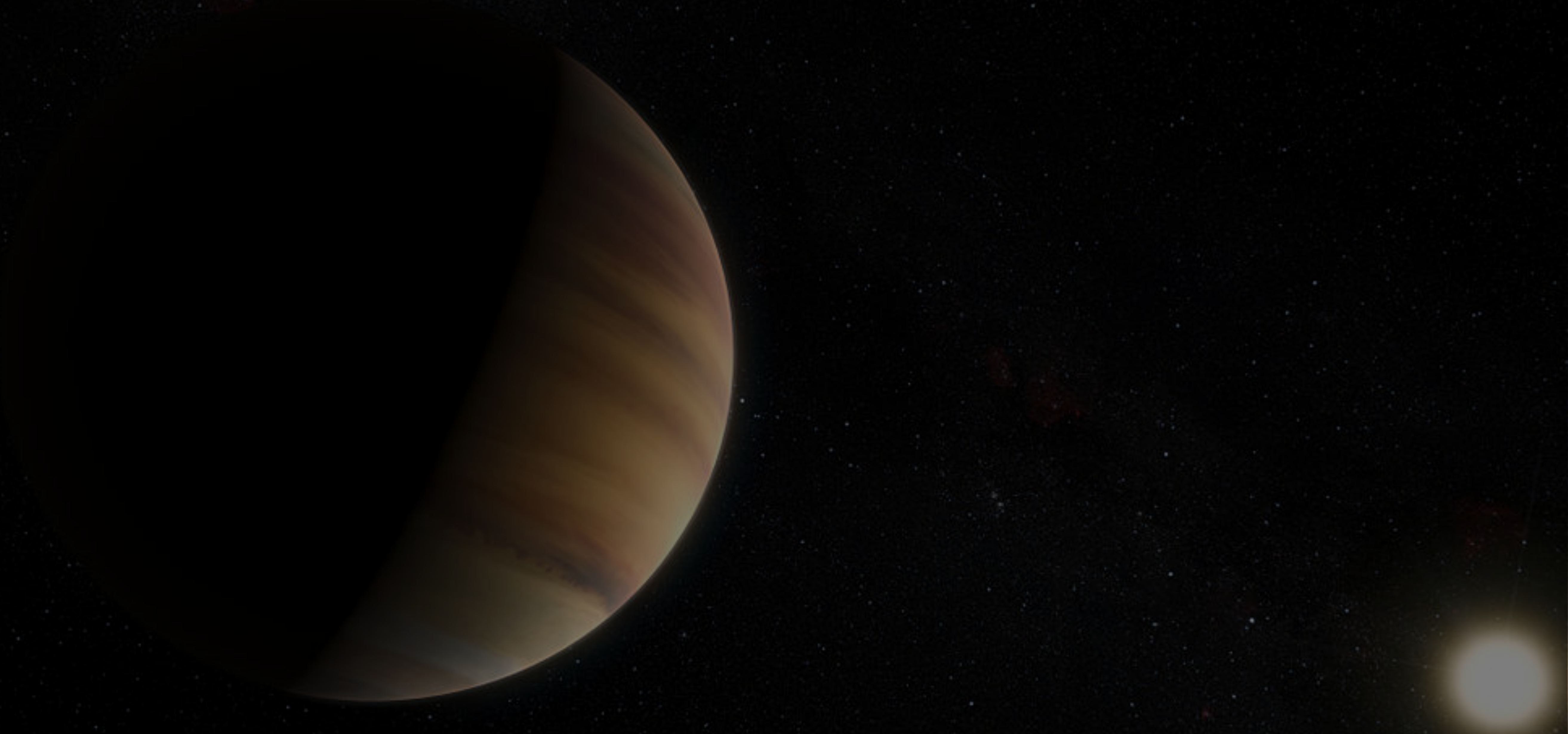
OVERVIEW

- Motivation
- Intro to transmission spectroscopy
- Current retrieval approaches
- Machine learning-informed priors in Bayesian retrieval

MOTIVATION

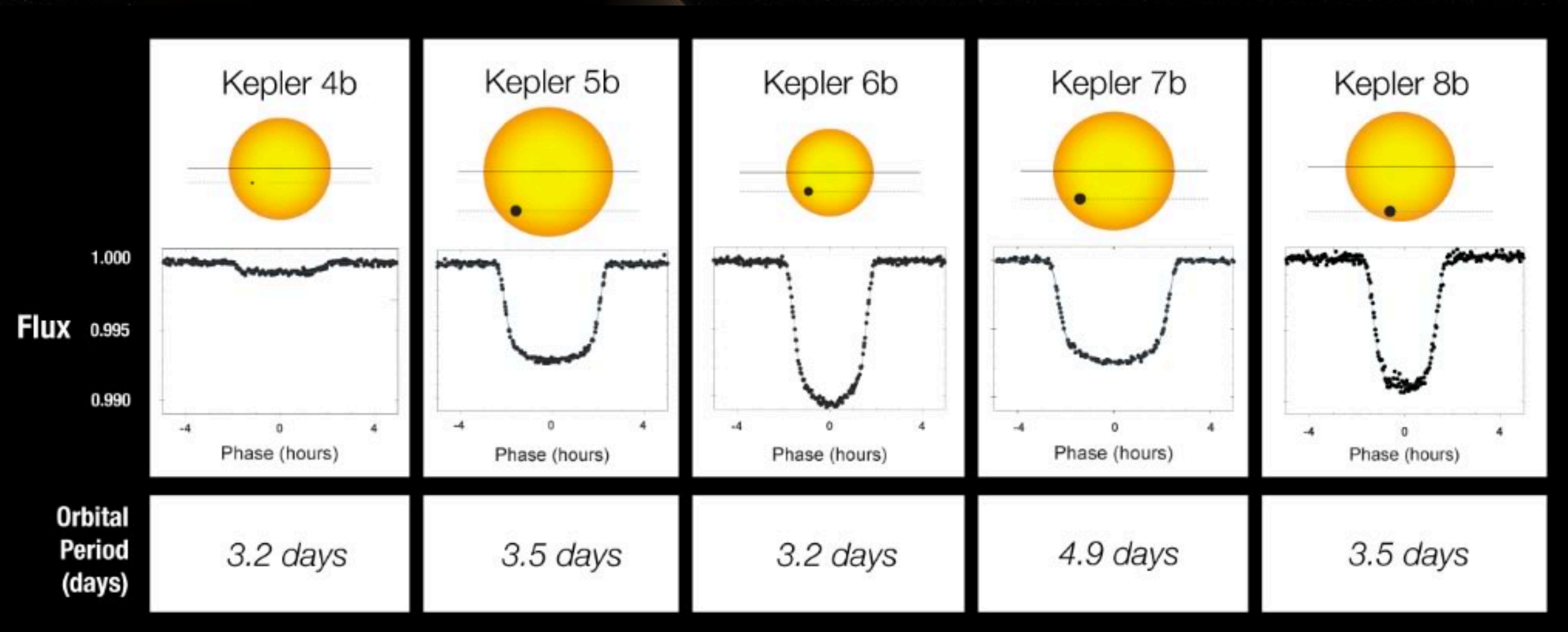
- Currently small number of viable targets for atmospheric study (~100s).
- TESS, NGTS, PLATO etc will raise this to >20,000
- Current retrieval methods are slow
- Can we use ML to improve this?

TRANSMISSION SPECTROSCOPY



TRANSMISSION SPECTROSCOPY

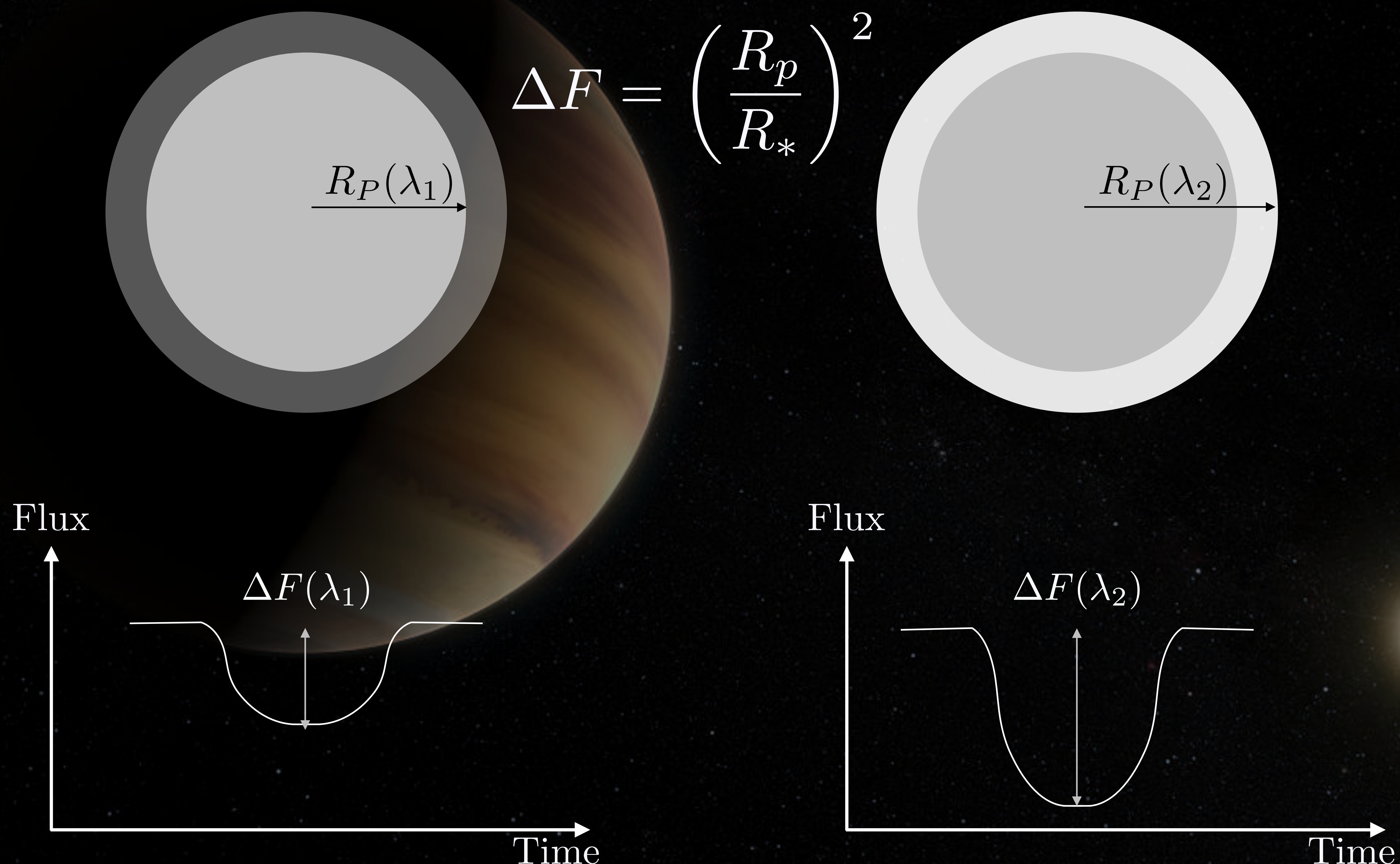
PLANETARY TRANSITS



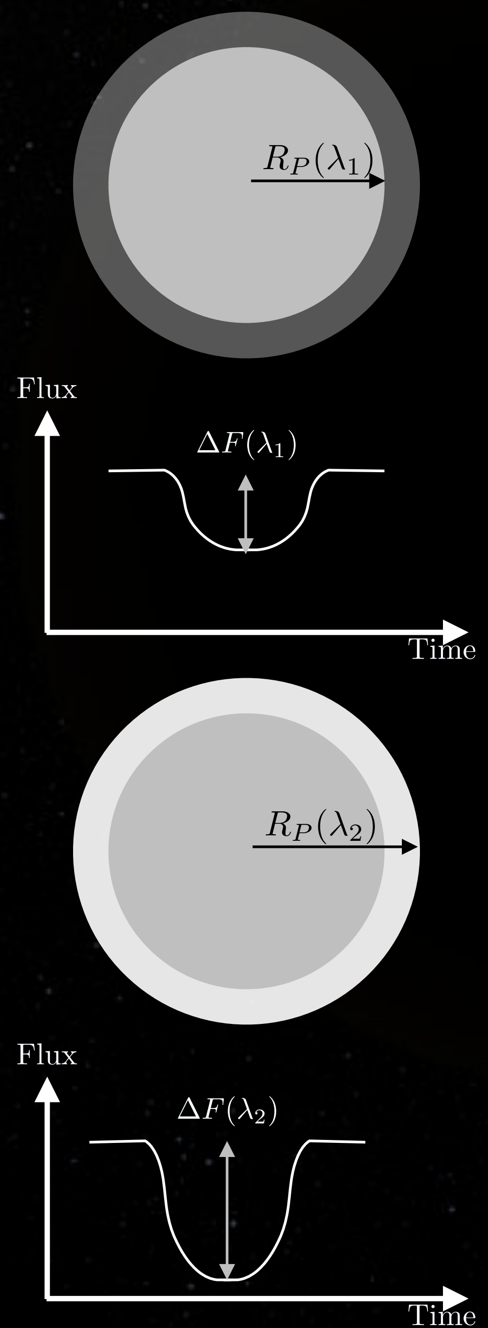
NASA

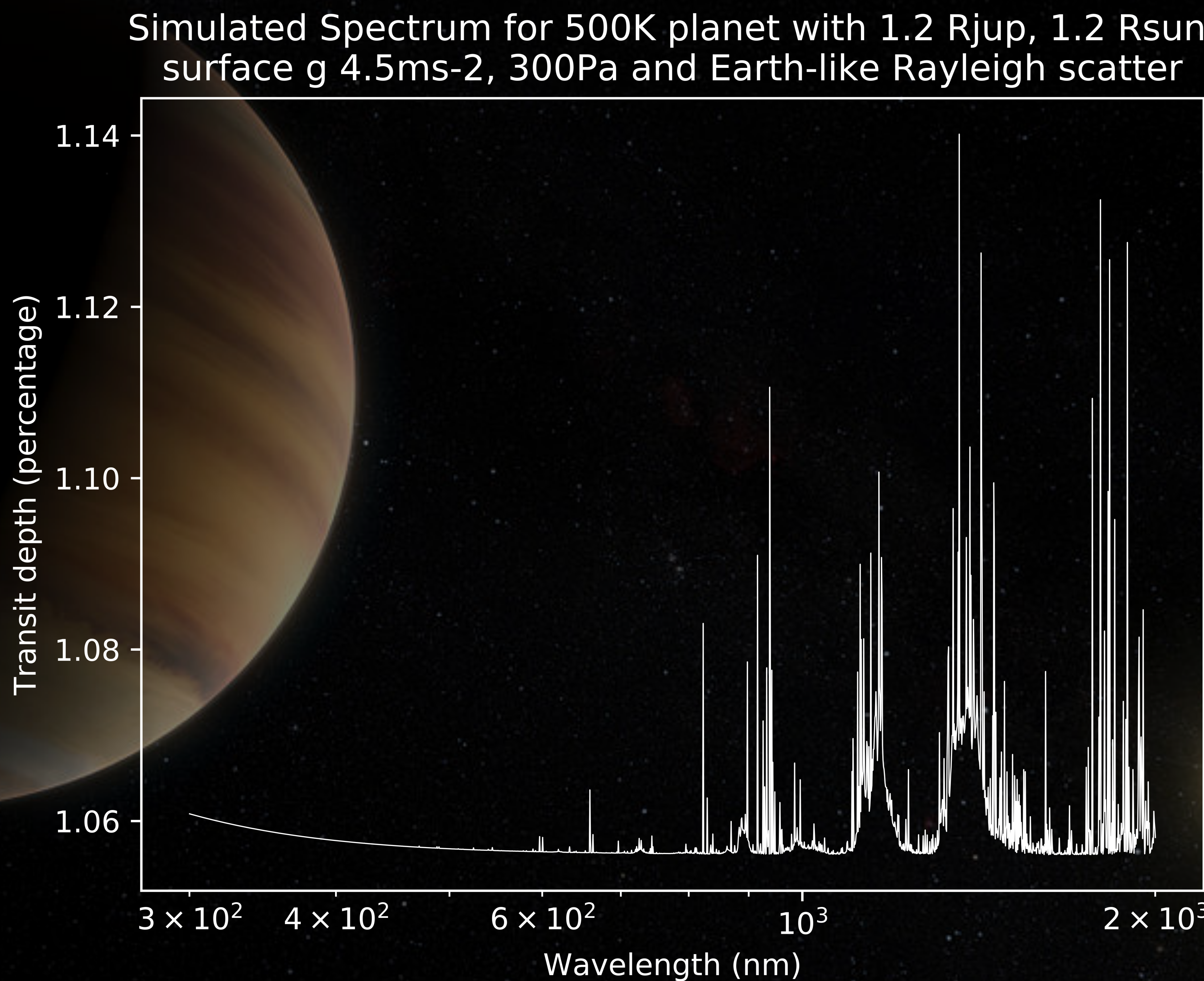
WAVELENGTH VARIATION

$$\Delta F = \left(\frac{R_p}{R_*}\right)^2$$



WAVELENGTH VARIATION





MANCHESTER 1824
The University of Manchester

@mosphericHayes

WHAT AFFECTS A SPECTRUM?

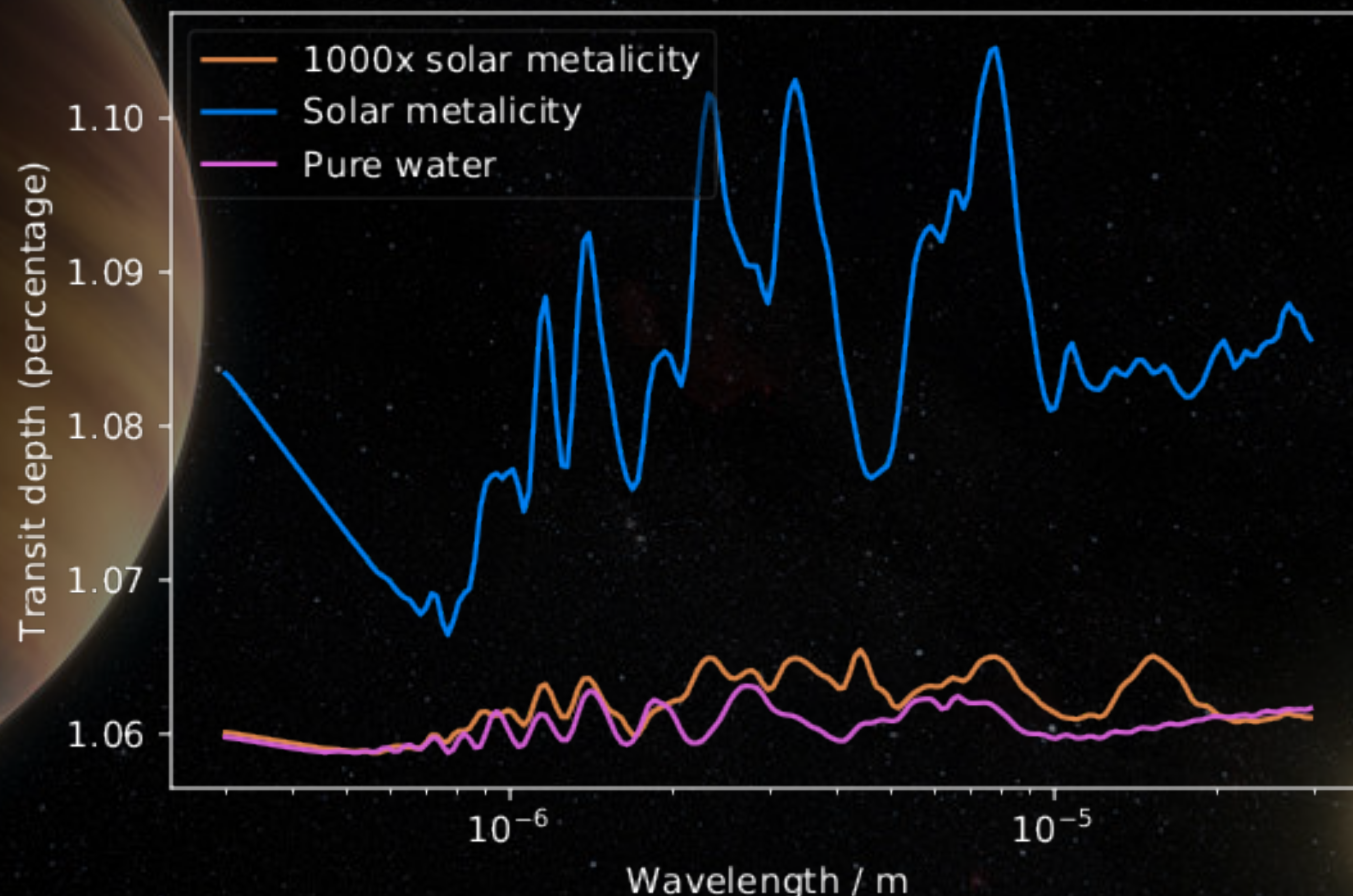
Basic parameters:

- Equation of state (EoS)
- Temperature
- Planet mass
- Planet Radius
- Star radius
- Presence and height of clouds
- Rayleigh scatter (atmospheric hazes)

WHAT AFFECTS A SPECTRUM?

Basic parameters:

- **Equation of state (EoS)**
- Temperature
- Planet mass
- Planet Radius
- Star radius
- Presence and height of clouds
- Rayleigh scatter (atmospheric hazes)

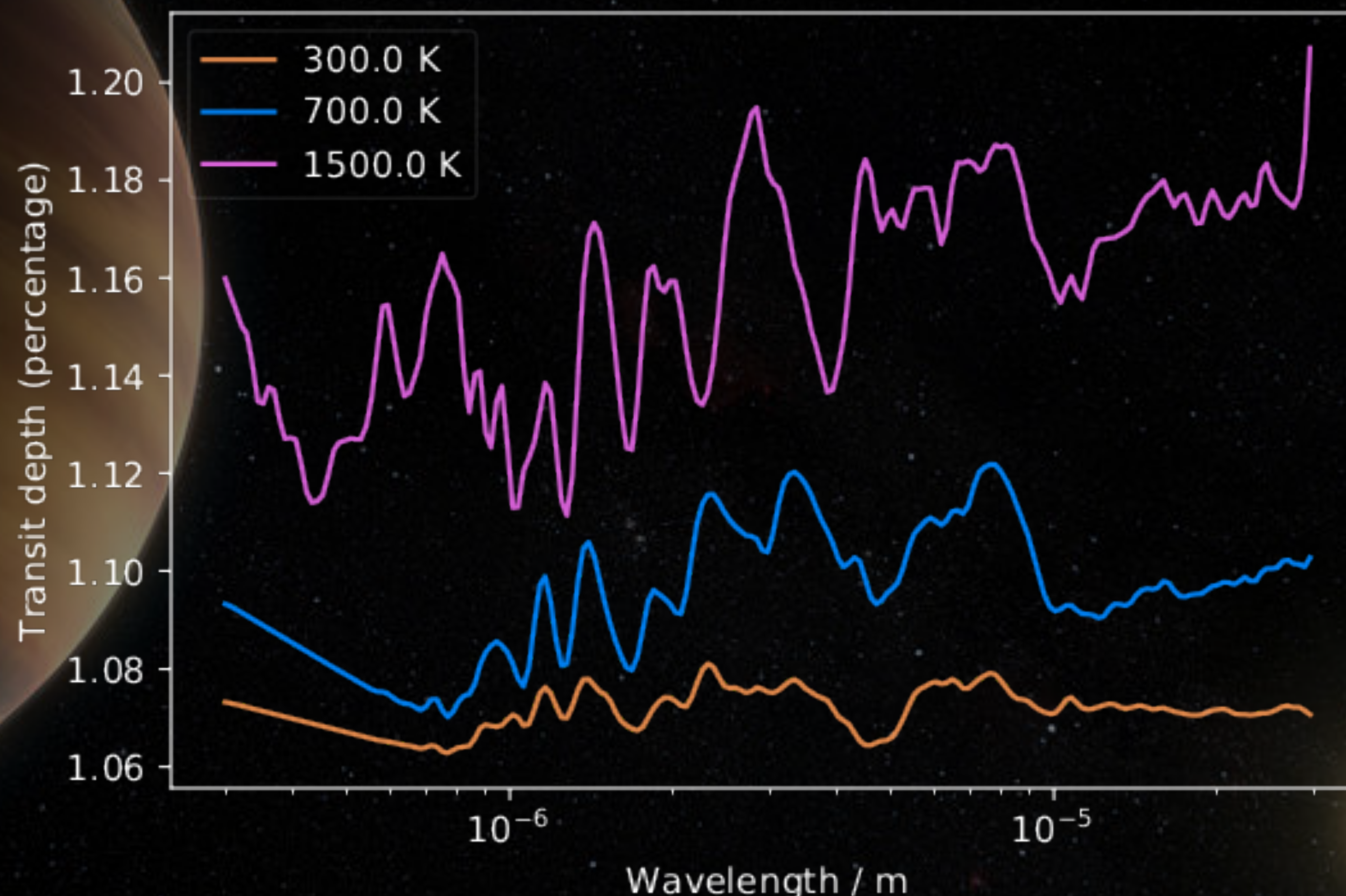


Default params:
EoS - Solar; T - 500K; g - 10ms^{-2}
R_p - 1 R_{Jup}; R_s - 1 R_{Sun};
P - No clouds; r -1

WHAT AFFECTS A SPECTRUM?

Basic parameters:

- Equation of state (EoS)
- **Temperature**
- Planet mass
- Planet Radius
- Star radius
- Presence and height of clouds
- Rayleigh scatter (atmospheric hazes)



Default params:

EoS - Solar; T - 500K; g - 10ms^{-2}

R_p - 1 R_{Jup}; R_s - 1 R_{Sun};

P - No clouds; r -1

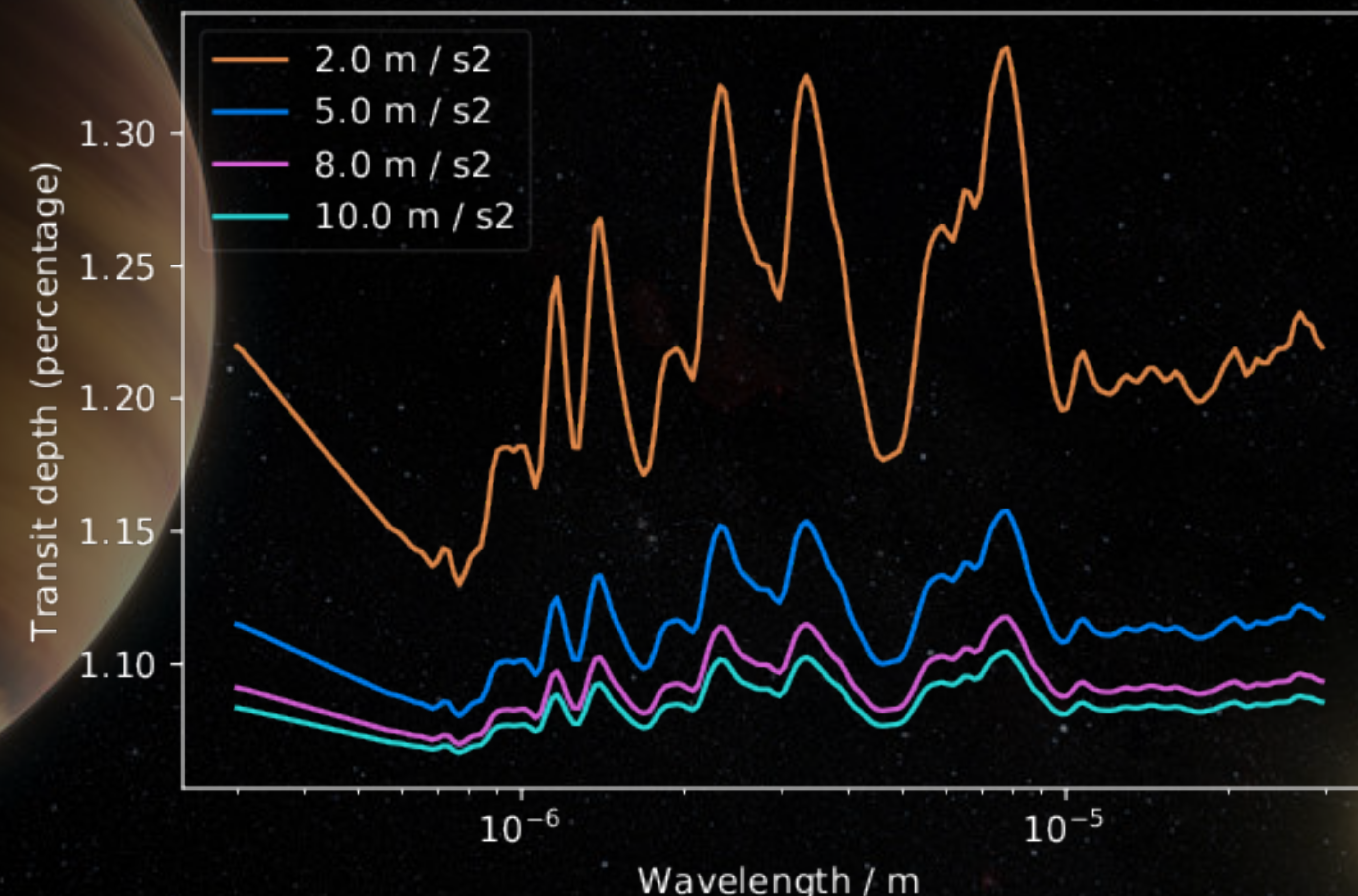
MANCHESTER 1824 The University of Manchester

@mosphericHayes

WHAT AFFECTS A SPECTRUM?

Basic parameters:

- Equation of state (EoS)
- Temperature
- **Planet mass**
- Planet Radius
- Star radius
- Presence and height of clouds
- Rayleigh scatter (atmospheric hazes)

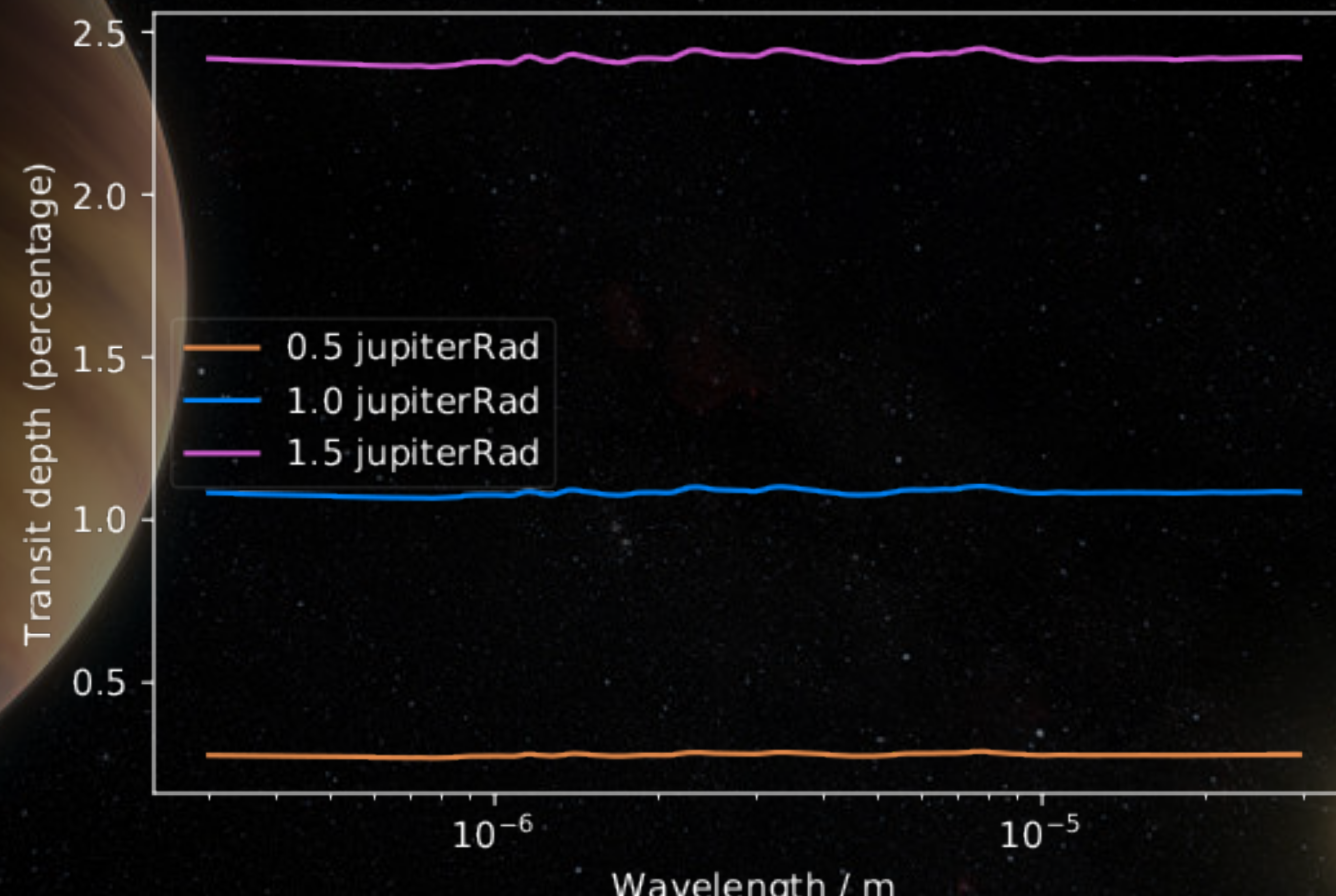


Default params:
EoS - Solar; T - 500K; g - $10ms^{-2}$
R_p - 1 R_{Jup}; R_s - 1 R_{Sun};
P - No clouds; r -1

WHAT AFFECTS A SPECTRUM?

Basic parameters:

- Equation of state (EoS)
- Temperature
- Planet mass
- **Planet Radius**
- Star radius
- Presence and height of clouds
- Rayleigh scatter (atmospheric hazes)

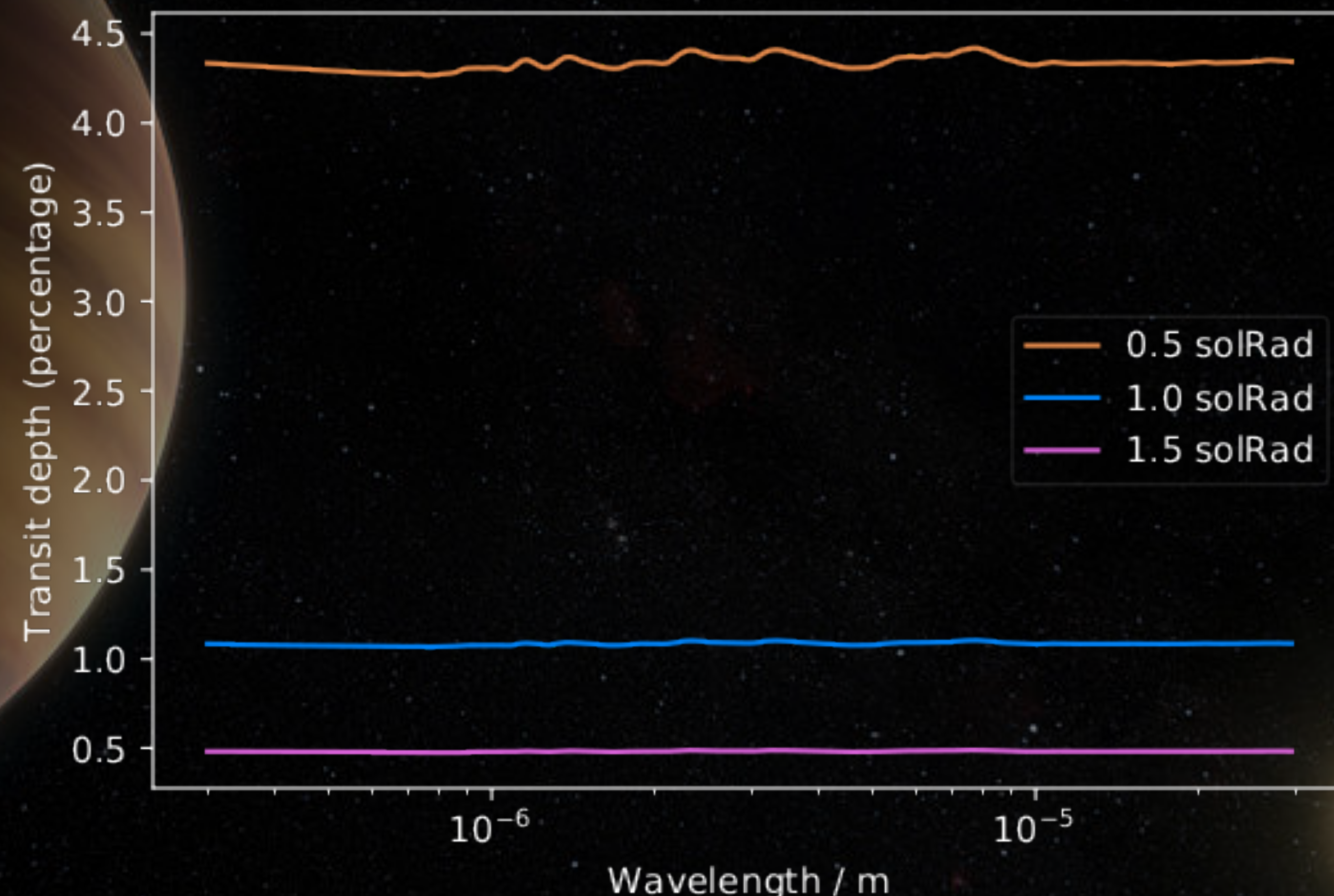


Default params:
EoS - Solar; T - 500K; g - 10ms^{-2}
R_p - 1 R_{Jup}; R_s - 1 R_{Sun};
P - No clouds; r -1

WHAT AFFECTS A SPECTRUM?

Basic parameters:

- Equation of state (EoS)
- Temperature
- Planet mass
- Planet Radius
- **Star radius**
- Presence and height of clouds
- Rayleigh scatter (atmospheric hazes)

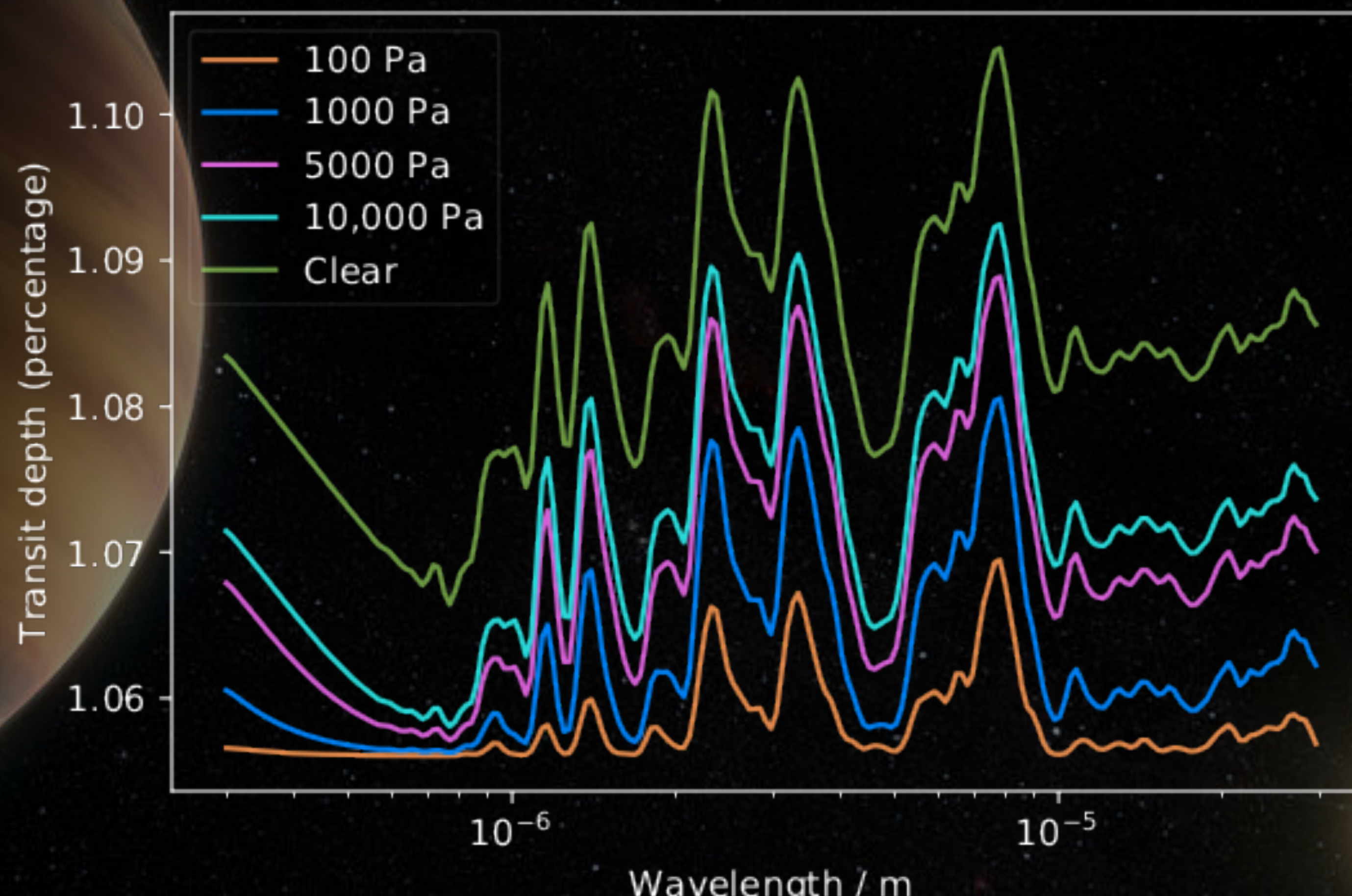


Default params:
EoS - Solar; T - 500K; g - $10ms^{-2}$
R_p - 1 R_{Jup}; R_s - 1 R_{Sun};
P - No clouds; r -1

WHAT AFFECTS A SPECTRUM?

Basic parameters:

- Equation of state (EoS)
- Temperature
- Planet mass
- Planet Radius
- Star radius
- **Presence and height of clouds**
- Rayleigh scatter (atmospheric hazes)

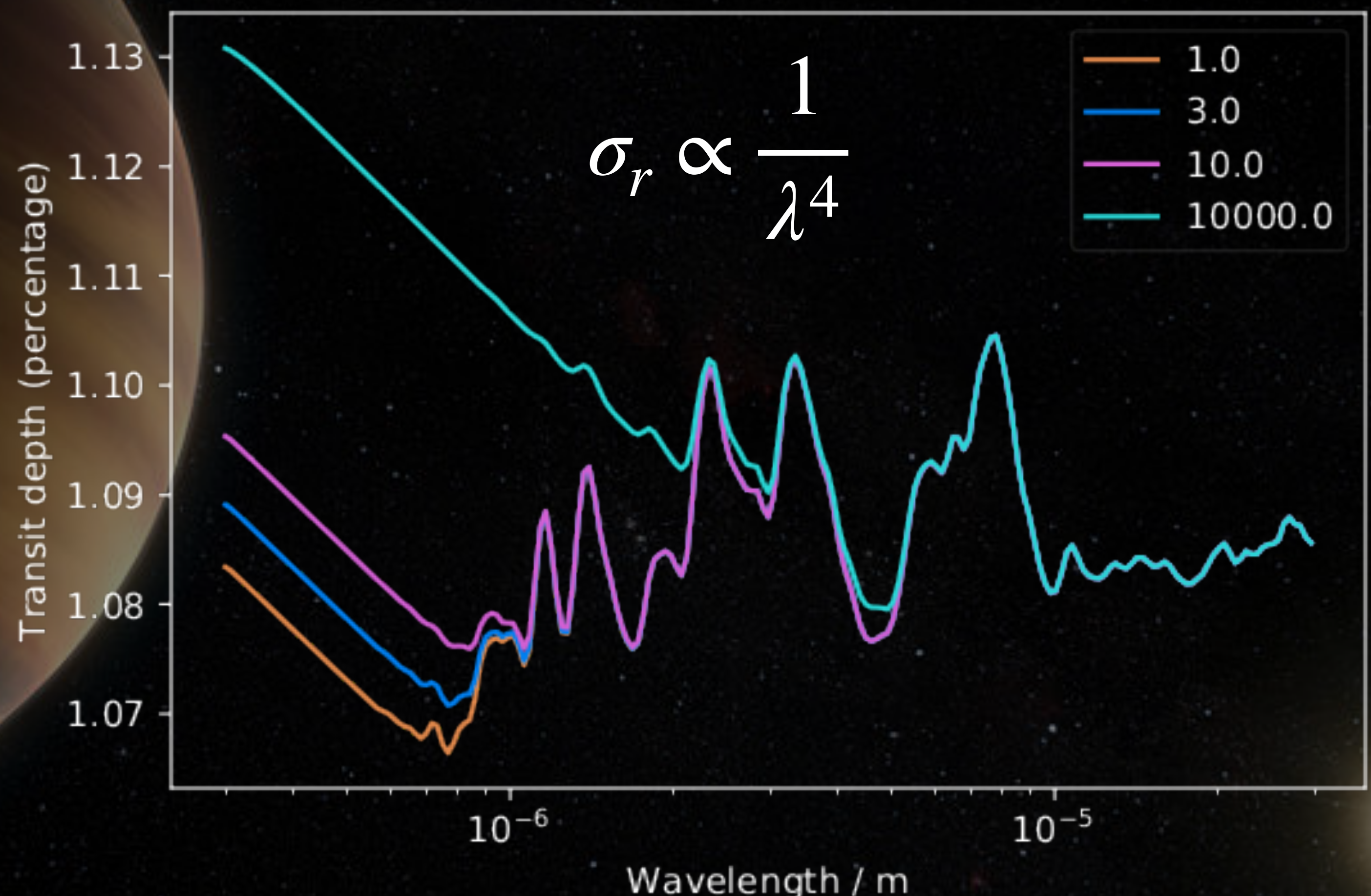


Default params:
EoS - Solar; T - 500K; g - 10ms^{-2}
R_p - 1 R_{Jup}; R_s - 1 R_{Sun};
P - No clouds; r -1

WHAT AFFECTS A SPECTRUM?

Basic parameters:

- Equation of state (EoS)
- Temperature
- Planet mass
- Planet Radius
- Star radius
- Presence and height of clouds
- **Rayleigh scatter (atmospheric hazes)**



Default params:

EoS - Solar; T - 500K; g - $10ms^{-2}$

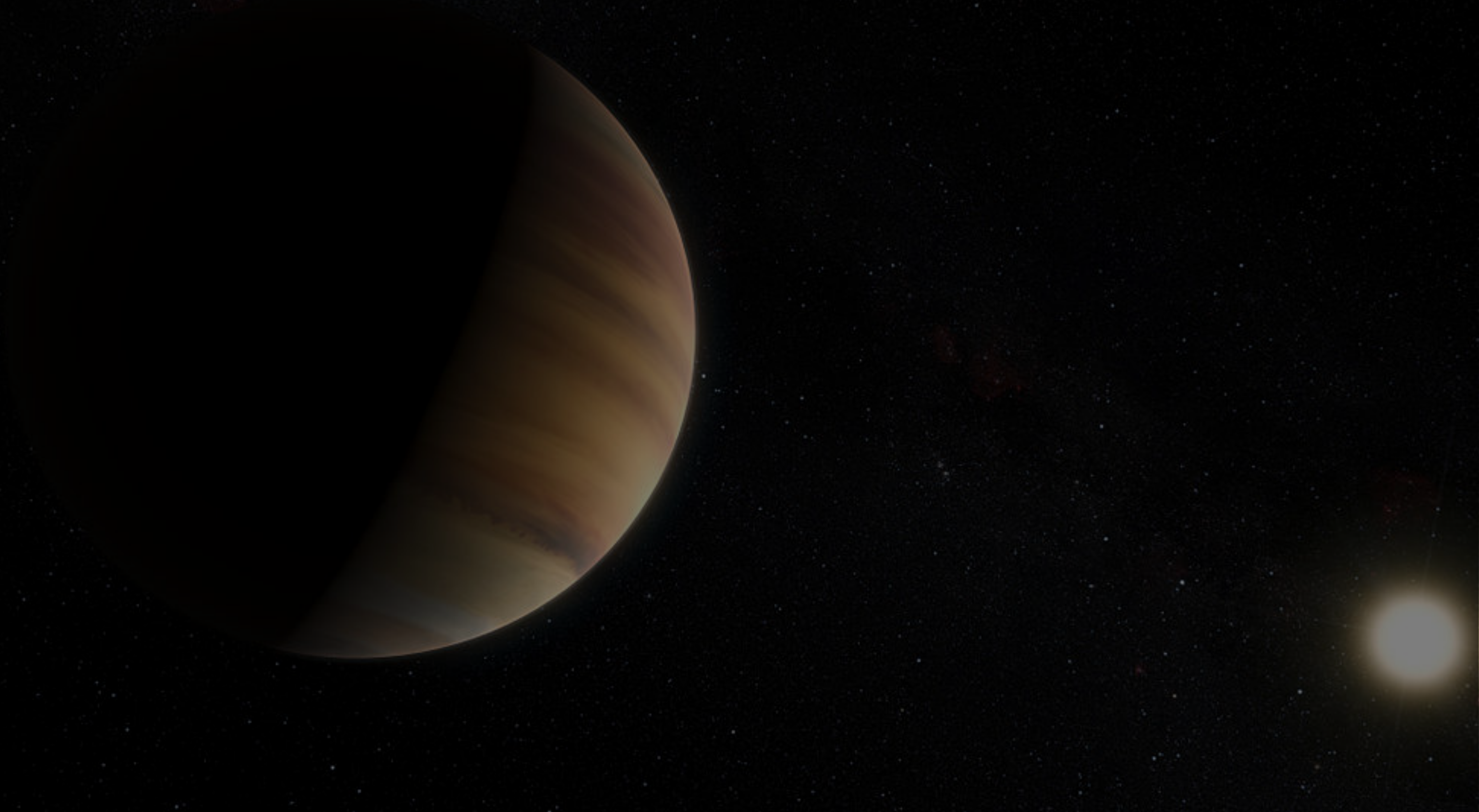
R_p - 1 R_{Jup}; R_s - 1 R_{Sun};

P - No clouds; r -1

MANCHESTER 1824 The University of Manchester

@mosphericHayes

ATMOSPHERIC RETRIEVAL



CURRENT METHODS

- Bayesian approaches - MCMC, nested sampling
  - \+ Sensitive to the physics. Can fully sample the parameter space. Established and fully supported.
  - \- Often slow. Requires input of priors.
- ML approaches - Random forests, Neural nets etc…
  - \+ Fast. Once trained, require very little/no maintenance
  - \- Usually agnostic to the physics, and can end up producing physically impossible results. Misclassification leads to incorrect results. Currently unproven in the field.

THE QUESTION:

We want to get the accuracy and completeness offered by Bayesian methods at an increased speed.

How can we use ML to improve existing Bayesian retrieval methods?

IMPROVING BAYESIAN APPROACHES

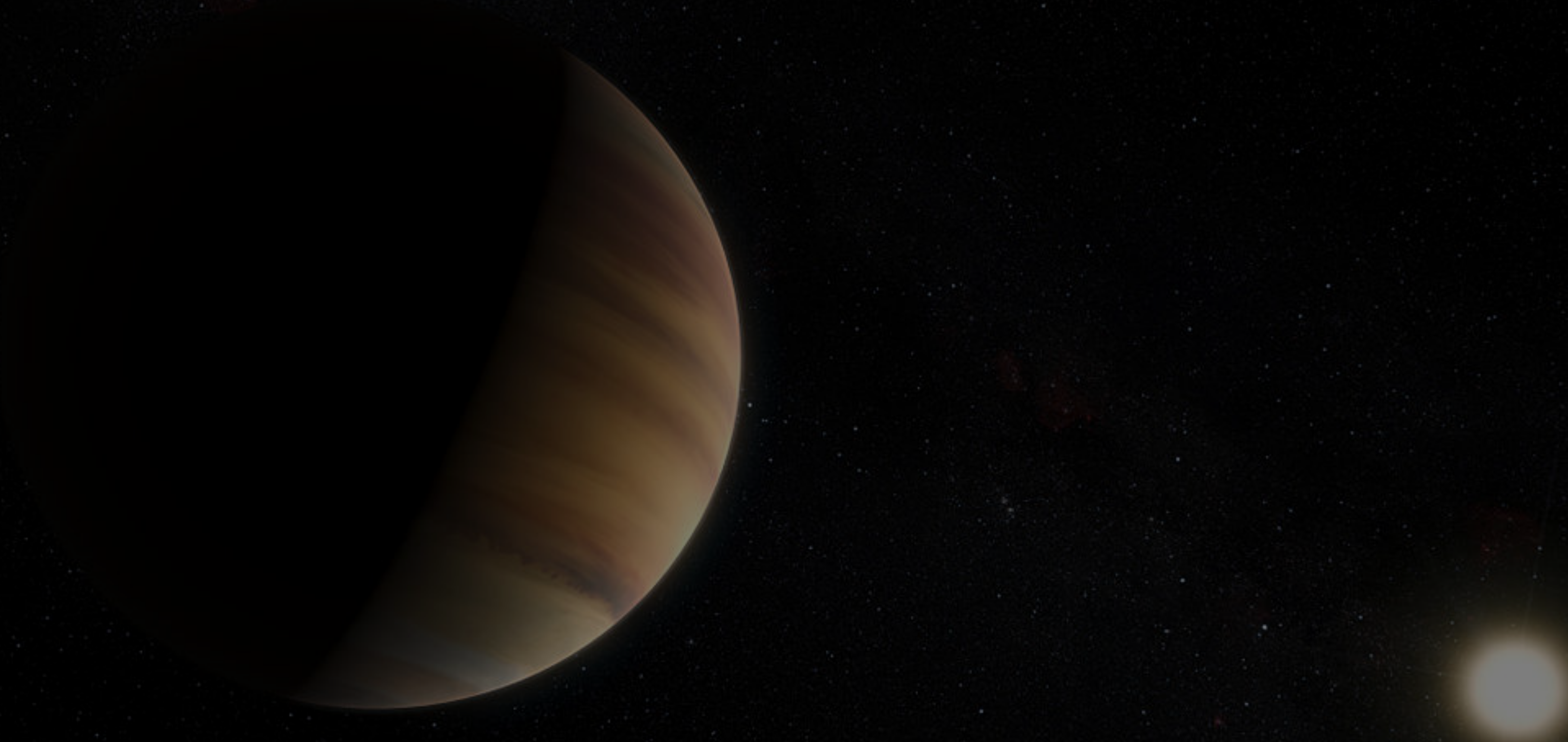
- Without rewriting algorithms, best place to start is with the priors.

- Smaller priors = more informed starting guess

  ➡ quicker convergence/meeting of stopping criteria

- **Use ML to produce informed priors quickly**

THE INFORMED PRIOR APPROACH



METHOD OVERVIEW

- Use a classifier to split simulated spectra into classes
- We can use the distributions from a class as priors

Unknown spectrum → RUN THROUGH CLASSIFIER → Priors!

Pick several 1000 random parameter sets

8D PARAMETER SPACE

- Star Radius R_*
- Planet mass M_p
- Planet radius R_p
- Atmosphere temperature T
- Metallicity $\log Z$
- CO ratio
- Cloud top pressure P
- Rayleigh scattering r

Pick several 1000 random parameter sets

Generate spectra for each parameter set

8D PARAMETER SPACE

ND SPECTRUM SPACE

Dimensionality depends on spectral resolution

$$\mathbf{s} = \left(\lambda_1, \lambda_2, \lambda_3 \ldots \lambda_{N_D}\right)$$

$$R = 100 \rightarrow N_D = 200$$

We want to cluster in spectrum space, but dimensionality is too high

Pick several 1000 random parameter sets

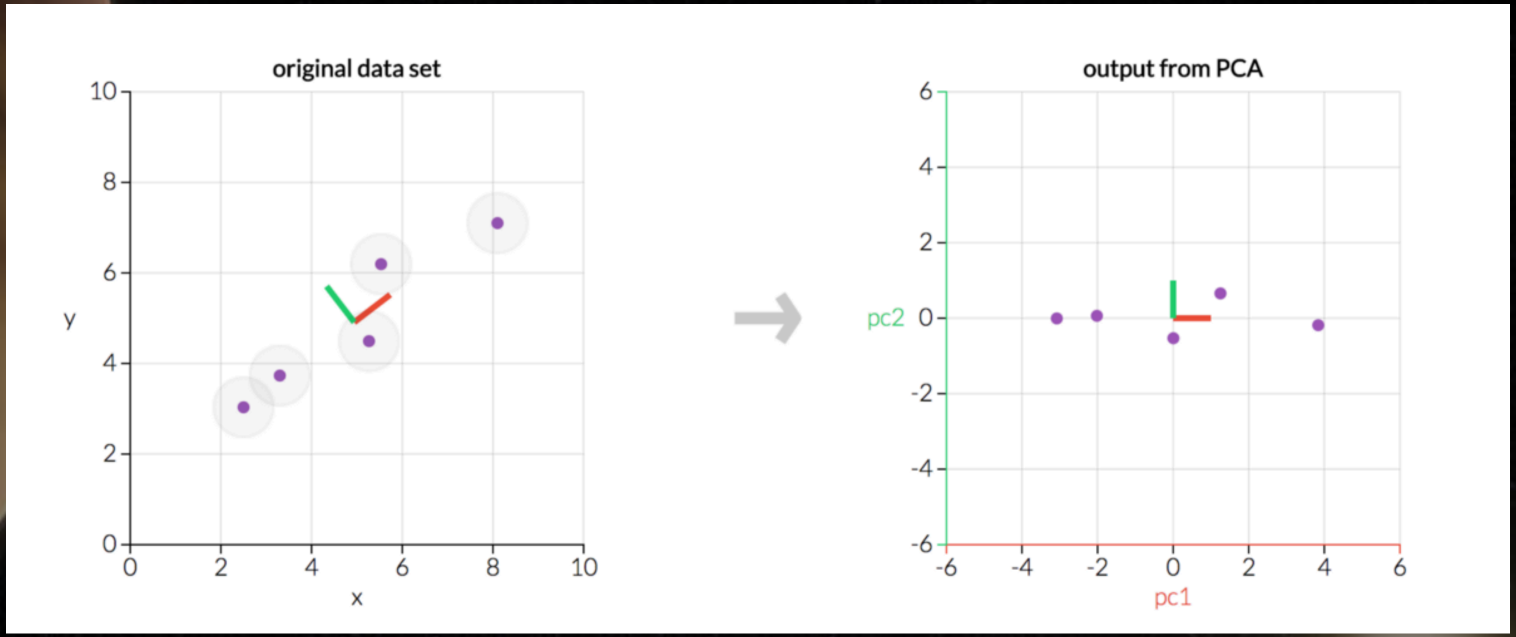
Generate spectra for each parameter set

Perform PCA and reduce dimensionality

8D PARAMETER SPACE

ND SPECTRUM SPACE

PCA REDUCED SPACE



@mosphericHayes

THE INFORMED PRIOR APPROACH

Pick several 1000 random parameter sets

Generate spectra for each parameter set

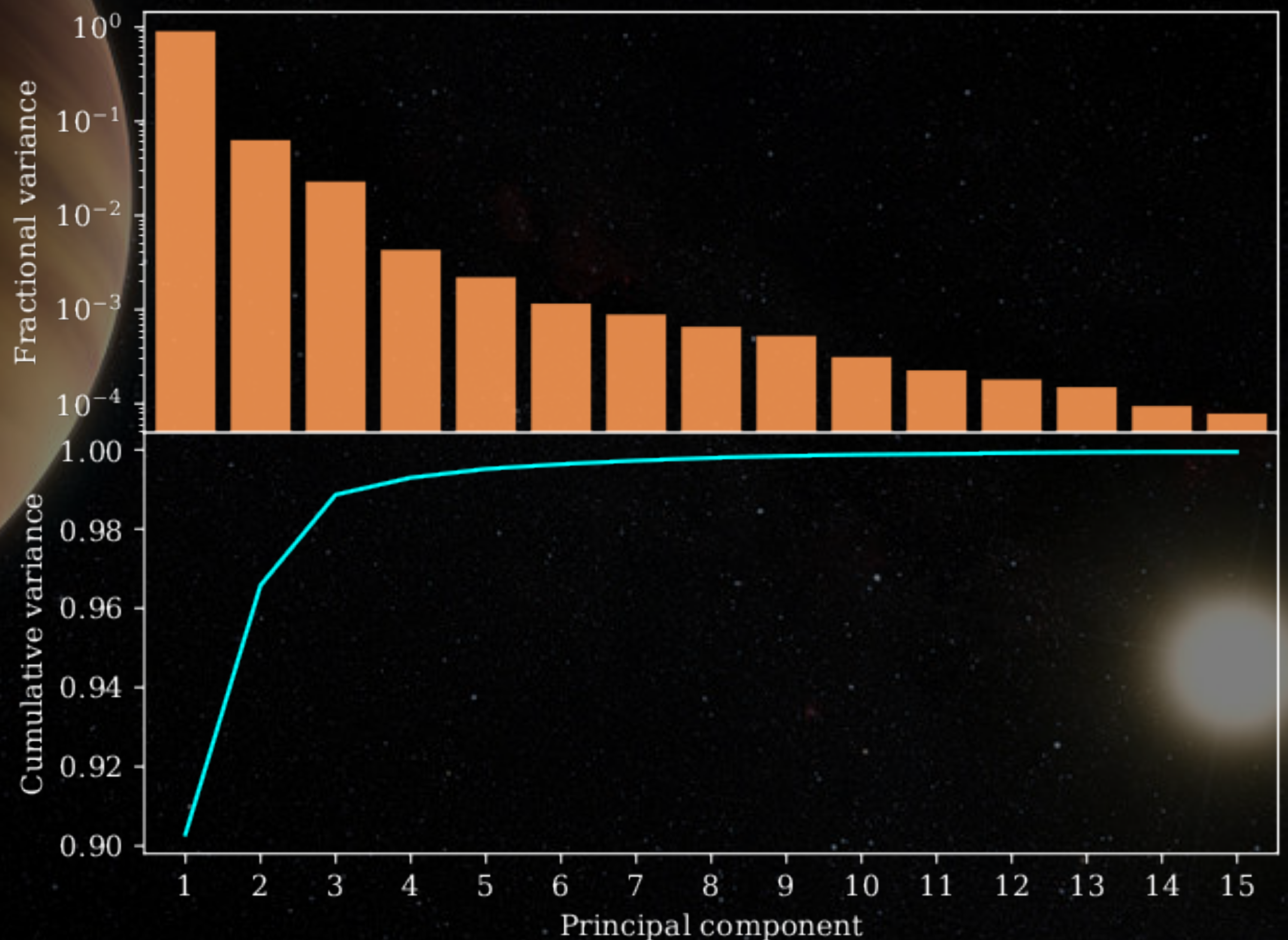
Perform PCA and reduce dimensionality

8D PARAMETER SPACE

ND SPECTRUM SPACE

PCA REDUCED SPACE

Hayes+SPEARNET, 2020



MANCHESTER 1824
The University of Manchester

@mosphericHayes

THE INFORMED PRIOR APPROACH

Pick several 1000 random parameter sets

Generate spectra for each parameter set

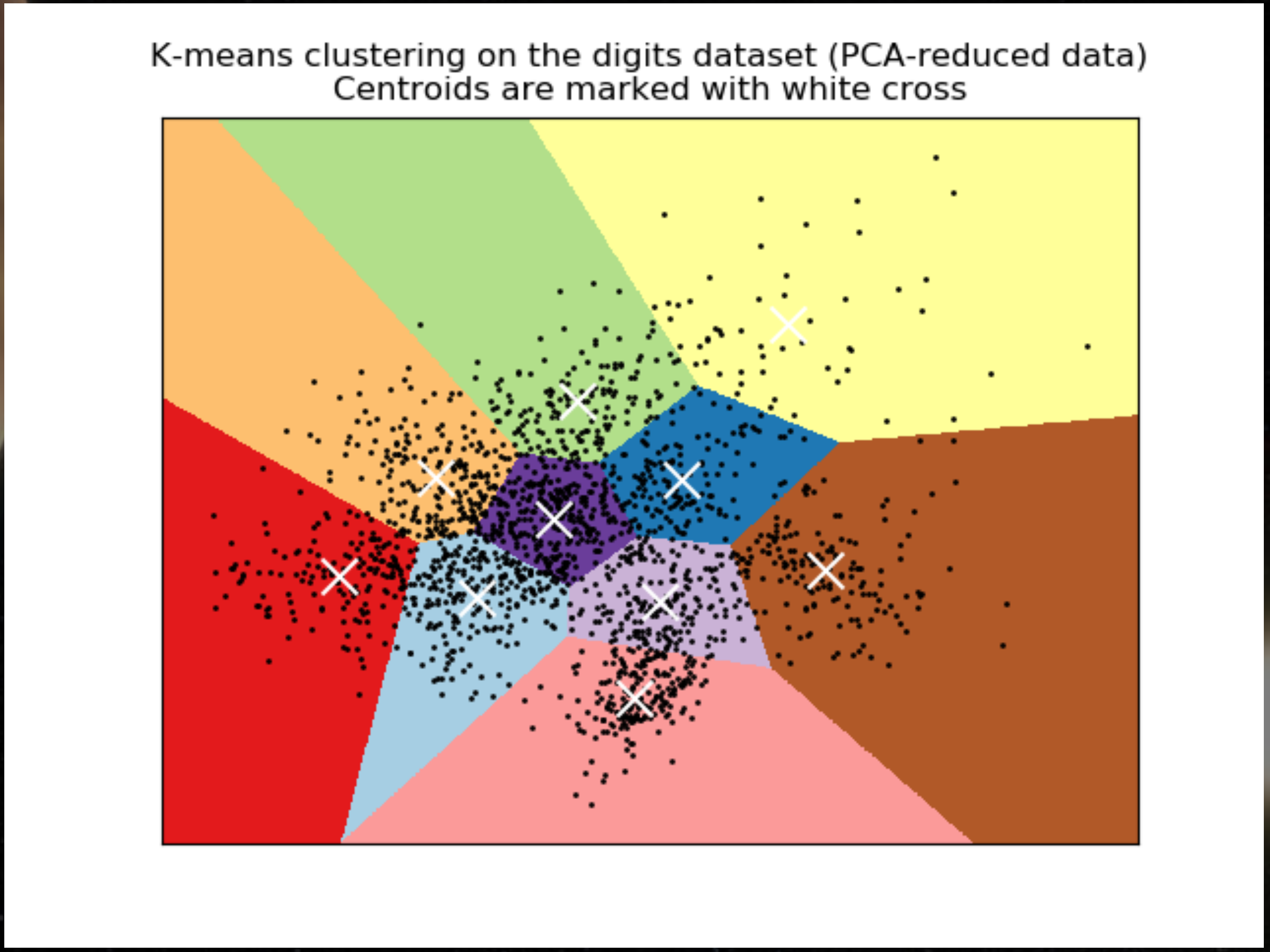
Perform PCA and reduce dimensionality

Train a k-means clustering classifier

8D PARAMETER SPACE

ND SPECTRUM SPACE

PCA REDUCED SPACE



THE INFORMED PRIOR APPROACH

Pick several 1000 random parameter sets

Generate spectra for each parameter set

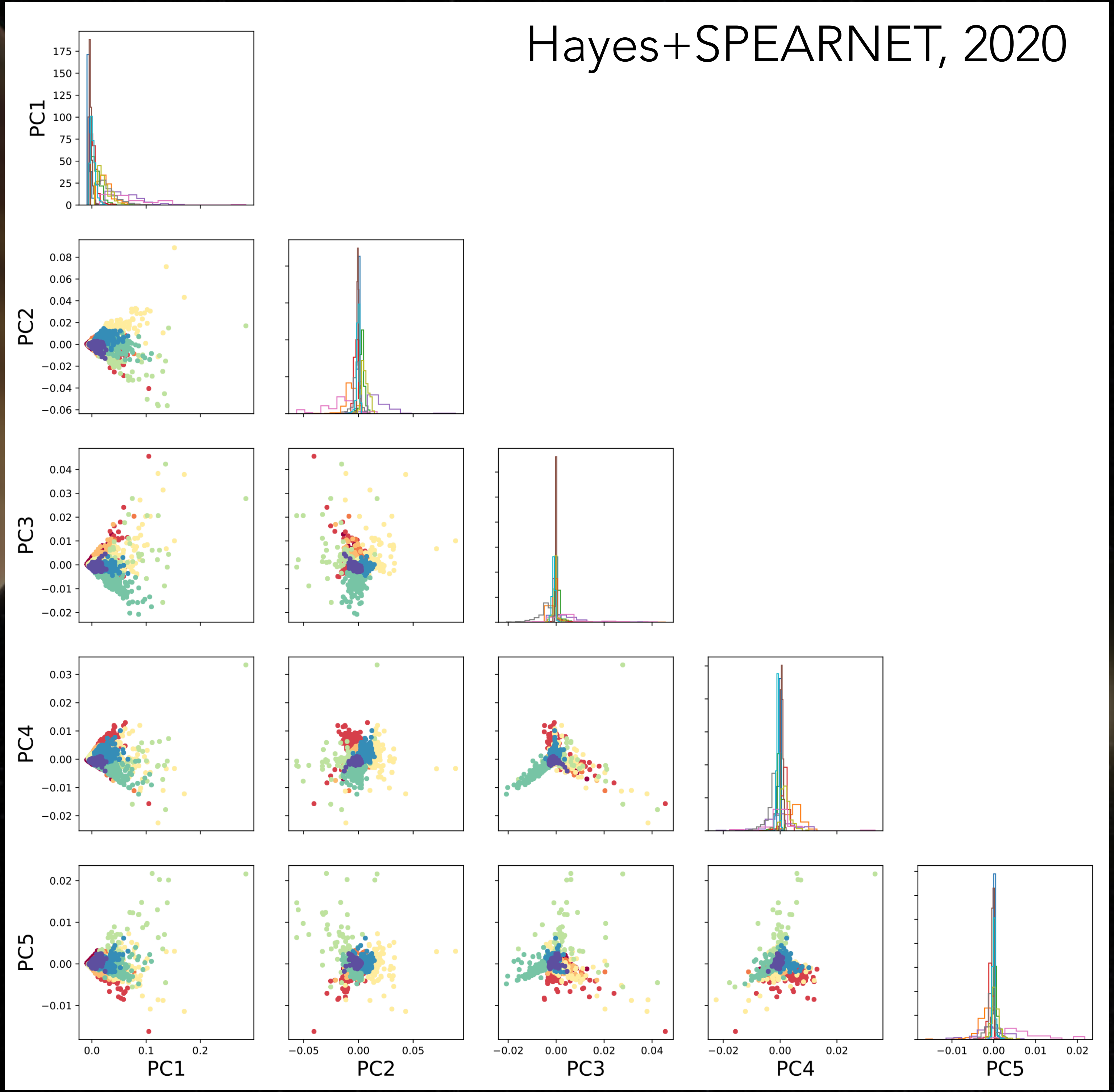
Perform PCA and reduce dimensionality

Train a k-means clustering classifier

8D PARAMETER SPACE

ND SPECTRUM SPACE

PCA REDUCED SPACE



MANCHESTER 1824
The University of Manchester

@mosphericHayes

THE INFORMED PRIOR APPROACH

8D PARAMETER SPACE

ND SPECTRUM SPACE

PCA REDUCED SPACE

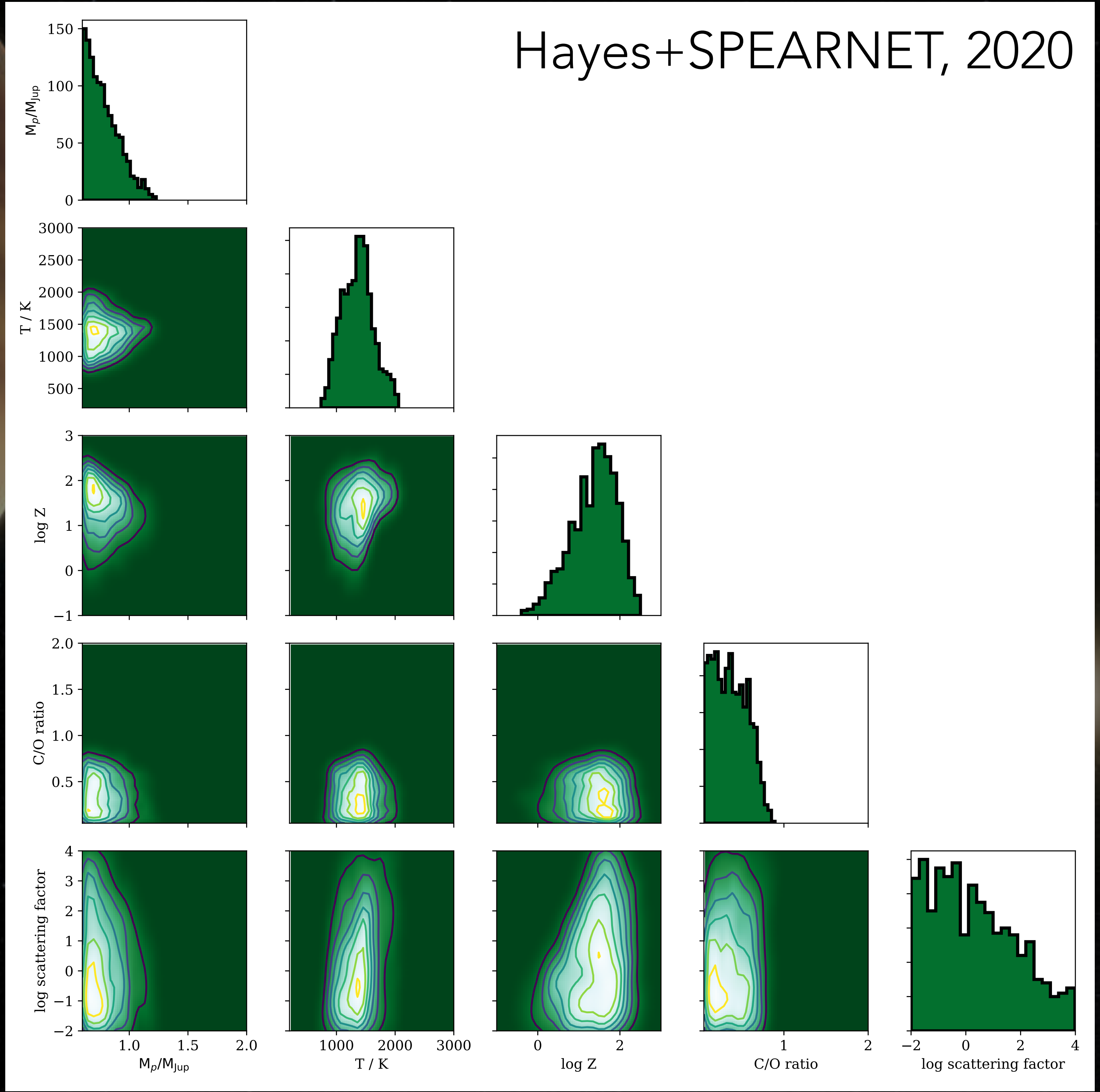
Pick several 1000 random parameter sets

Generate spectra for each parameter set

Perform PCA and reduce dimensionality

Train a k-means clustering classifier

Obtain parameter distributions for each cluster



PERFORMANCE TESTS

- Generate 240 new spectra and run nested sampling retrieval with uniform prior ('standard' method) with the informed prior ('classified' method)

- Investiate:

  - Speed (number of iterations)

  - Accuracy

  - Behaviour with misclassification

  - Effect of noise and resolution

PERFORMANCE TESTS - SPECTRA

- Full wavelength range
  - Resolutions 30, 100, 300
  - Noise at 1%, 10% dynamic range
- Observatory specific simulations:

Table 3. The wavelength ranges, spectral resolutions, and noise levels in ppm of the instruments for which spectra were generated and retrieval was run. *Twinkle* has three different wavelength bins, which are specified separately. In addition to these instruments being tested separately, a retrieval test was also run with a composite spectrum of (i) *HST*, *Twinkle*, and *JWST*-NIRSpec, and (ii) *HST* and FORS2, simulating the use of observations from multiple instruments.

Observatory	Low wavelength	High wavelength	Resolution	Noise level
HST (WFC3) (Kreidberg et al. 2014)	1.1 μm	1.7 μm	70	30 ppm
Twinkle (Edwards et al. 2019)	0.4 μm	1 μm	250	100 ppm
	1.3 μm	2.42 μm	250	100 ppm
	2.42 μm	4.5 μm	60	50 ppm
JWST-NIRSpec (Posselt et al. 2004)	0.6 μm	5.3 μm	100	30 ppm
FORS2 (GRIS600B and GRIS600RI grisms) (Nikolov et al. 2016)	0.411 μm	0.81 μm	60	240 ppm

To the paper!

PERFORMANCE TESTS - METRICS