EEG
EXOPLANETS
& EXOCLIMES
GROUP

FNSNF
FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION

b
UNIVERSITÄT
BERN

Supervised Machine Learning for Analysing Spectra of Exoplanetary Atmospheres

Chloe Fisher

10/07/2018, Spectroscopy of Exoplanets, Windsor

HELA

- Machine Learning – the buzz word!
- Different types of Machine Learning (Waldmann 2016).

- HELA – an open source random forest algorithm for analyzing exoplanet spectra.
- Marquez-Neila, Fisher, Sznitman & Heng (2018, Nature Astronomy Letters)

An Unexpected Collaboration

Medical Imaging

Pablo Marquez-Neila

Raphael Sznitman

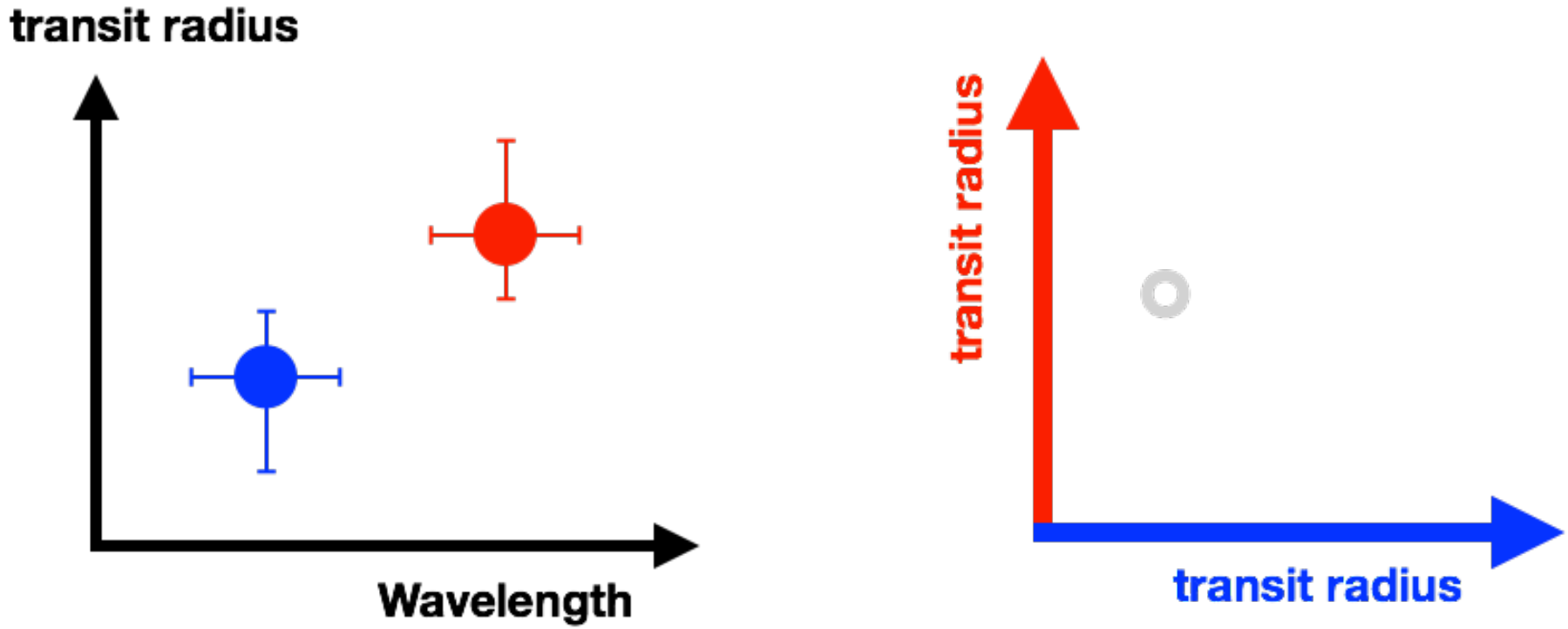
Astrophysics

Chloe Fisher

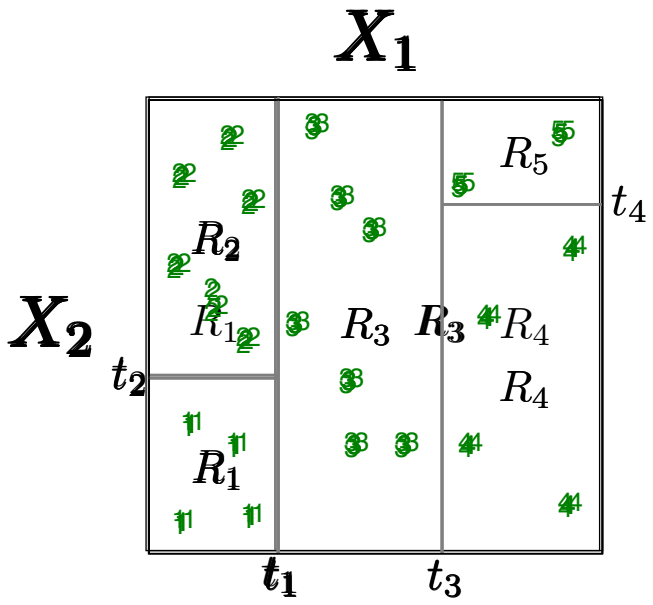
Kevin Heng

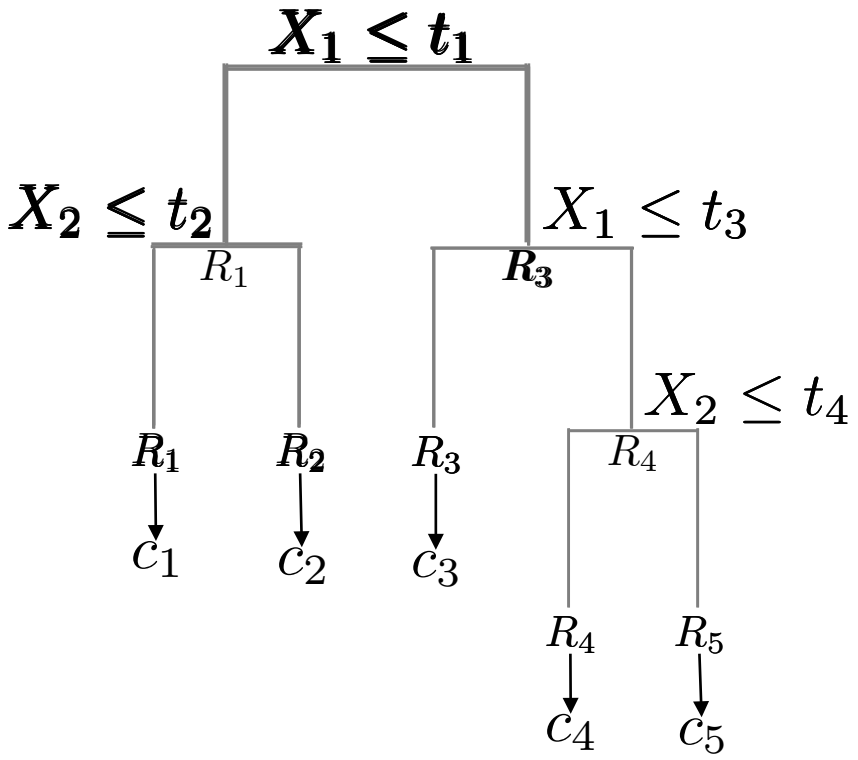
Chloe Fisher, Windsor 2018

A Single Regression Tree



A Single Regression Tree





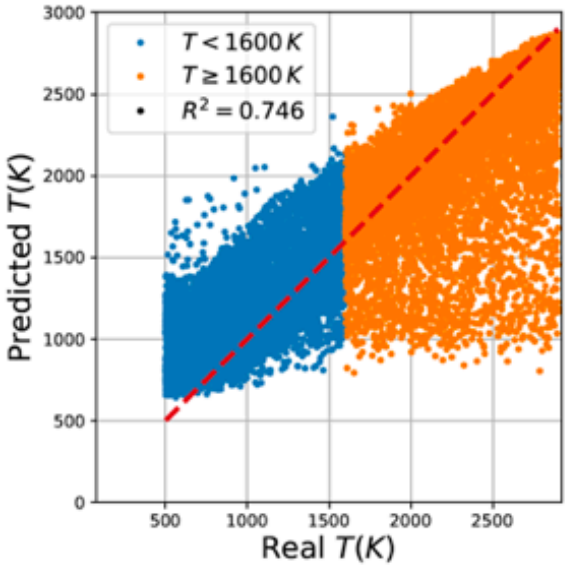
$$c_i = mean(\{y_n \mid x_n \in R_i\})$$

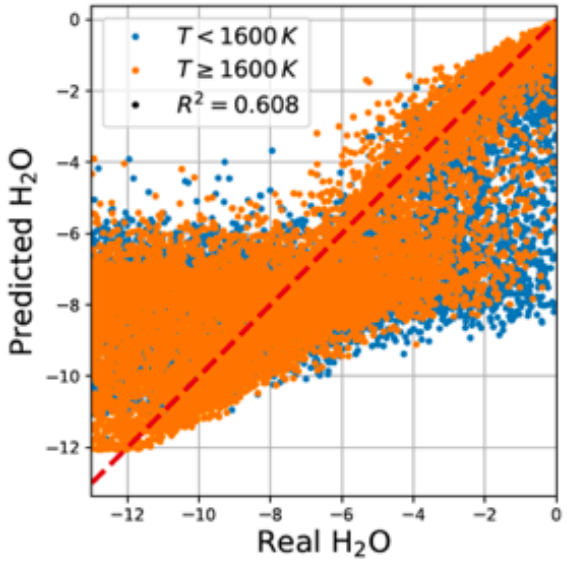
A Random Forest

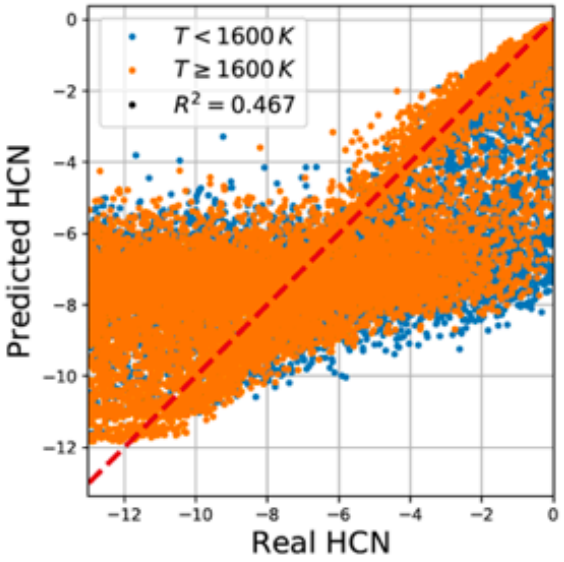
Every tree makes a ‘vote’ for the answer.

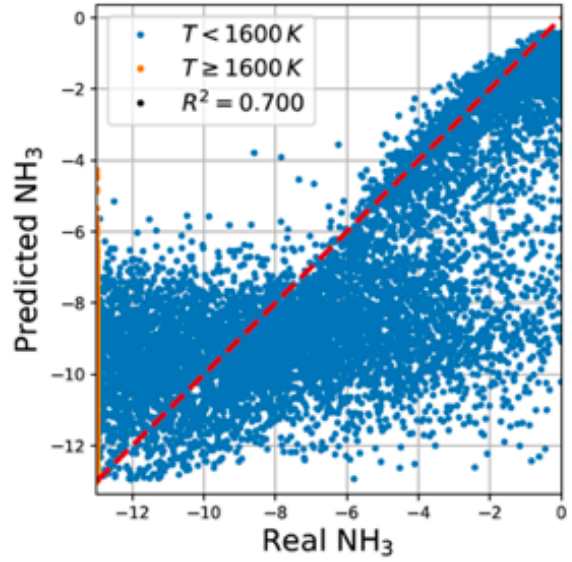
- Bias-Variance trade-off: Individual trees have low bias, but high variance.
- Combining the trees into a *Random Forest* gives you a model with low bias and low variance.
- ‘Random’ in several ways - bootstrapping and bagging.

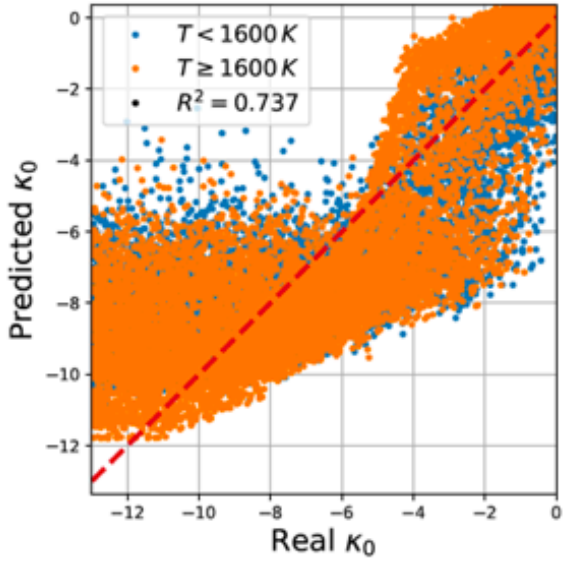
Predictions









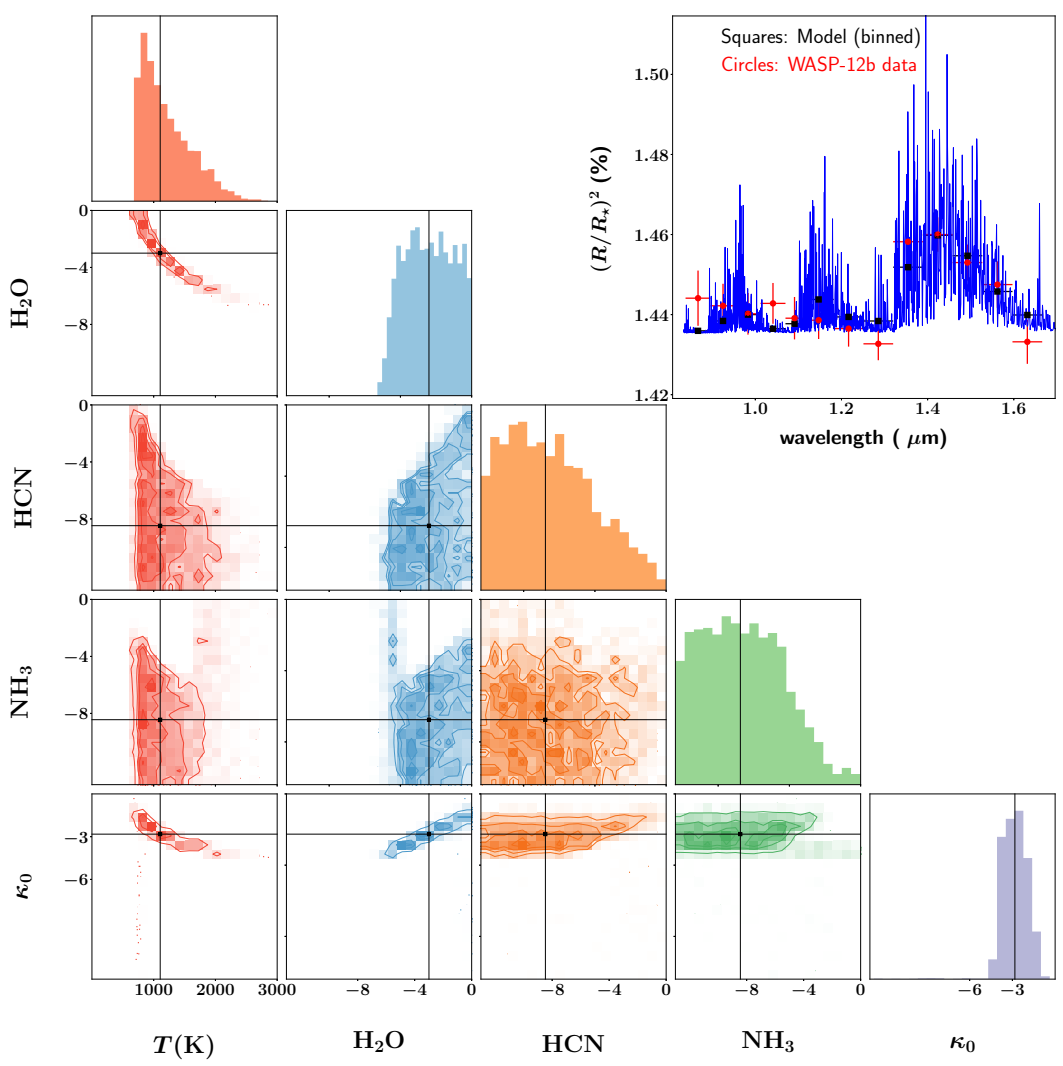


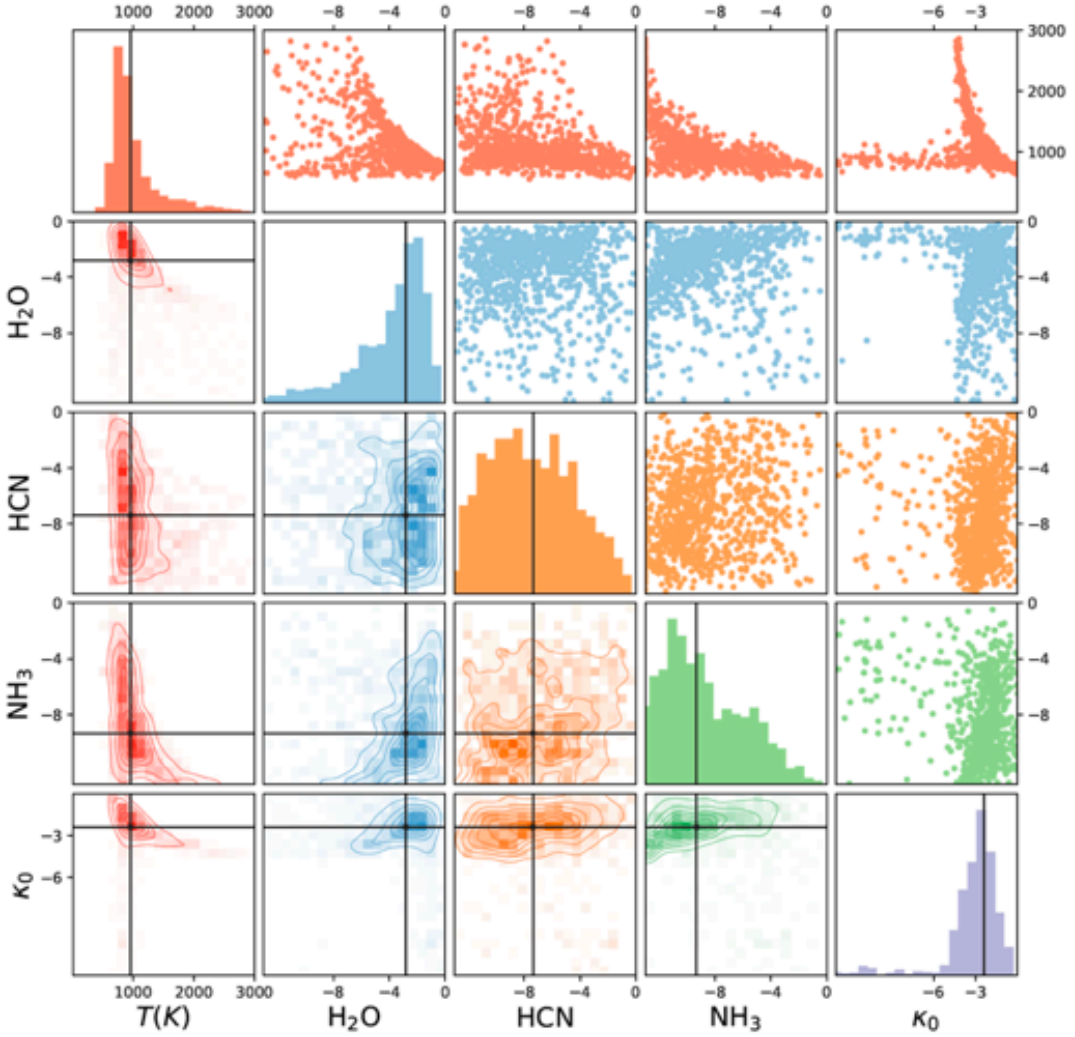
- Trained on 80,000 spectra.
- Tested on 20,000.
- Compare predicted with real.
- Results suggest a ‘detectability limit’.

Chloe Fisher, Windsor 2018

Marquez-Neila, Fisher, Sznitman & Heng (2018, Nature Astronomy Letters)

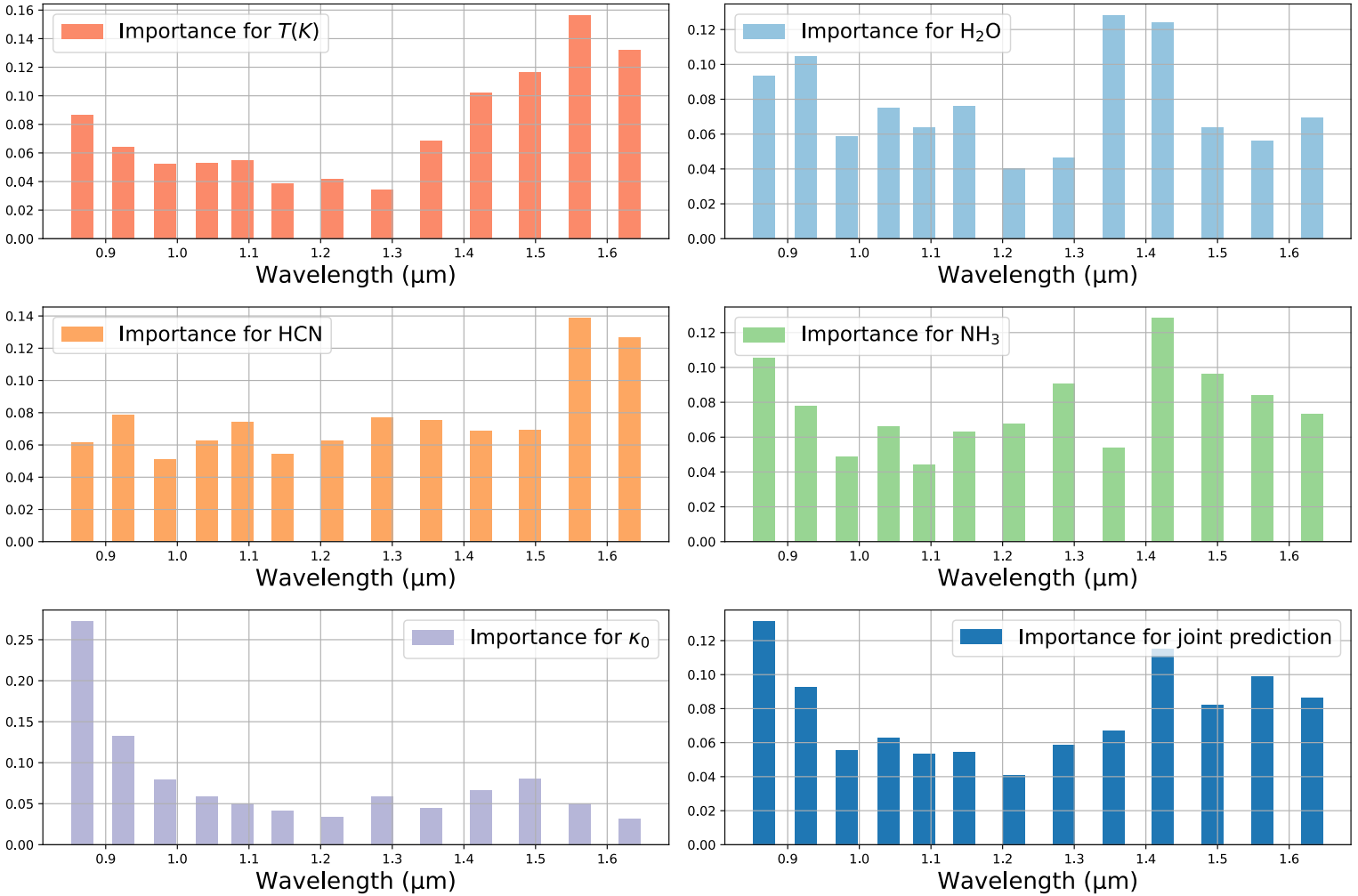
Nested Sampling vs Random Forest





Marquez-Neila, Fisher, Sznitman & Heng (2018, Nature Astronomy Letters)

Feature Importance



Marquez-Neila, Fisher, Sznitman & Heng (2018, Nature Astronomy Letters)

- Quantifies the information content of the wavebands.
- Goes beyond physical intuition.
- Could be used when applying for telescope time, or even when proposing new telescopes.

The Future

- Use it to analyze the current WFC3 planets.
- Use it to analyze Brown Dwarfs using existing model grids.
- Develop it for use with High-Resolution data.
- Prepare it for use with James Webb data….eventually!