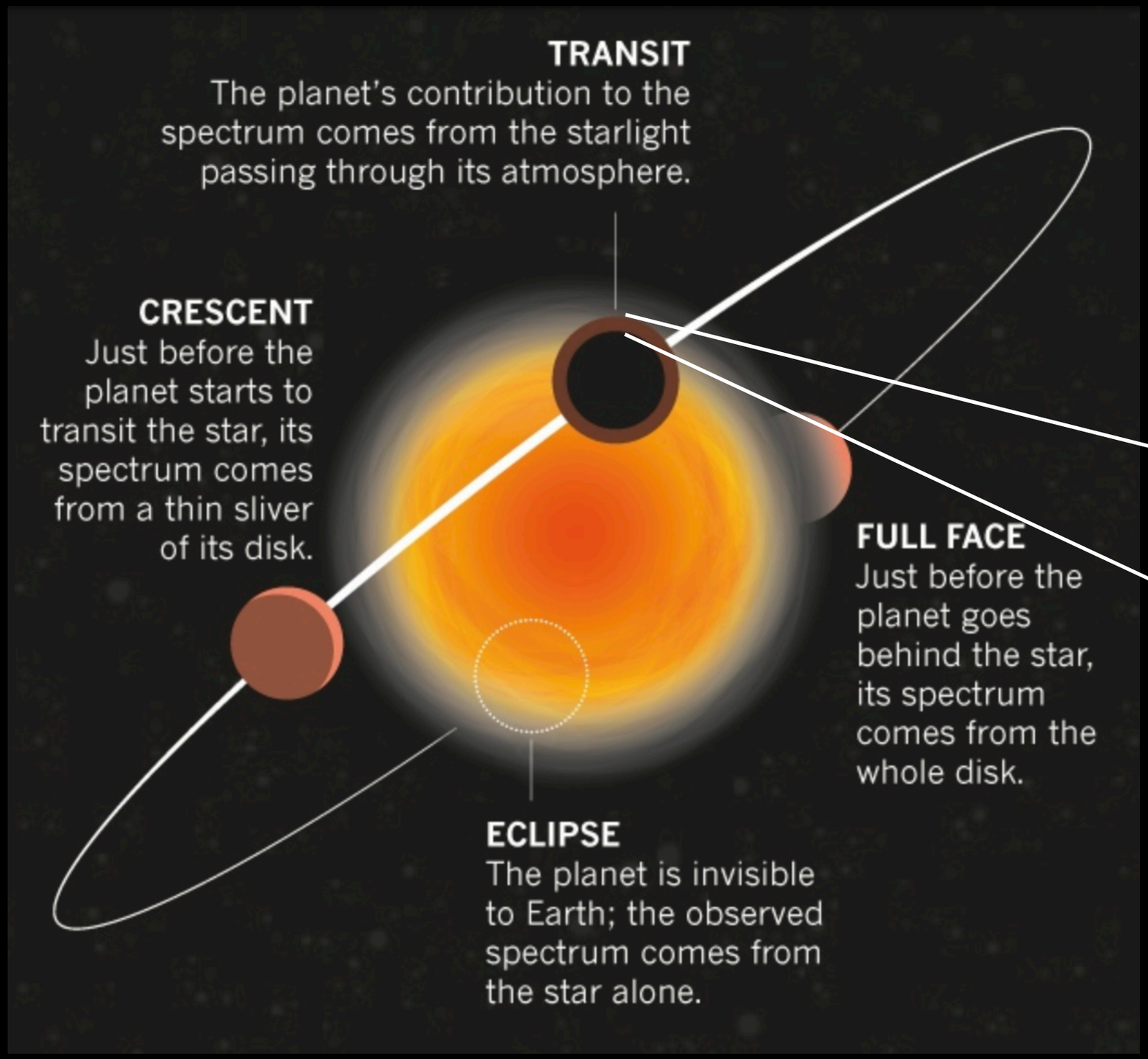
Data Analysis in the era of
Strategies for efficient population studies

Angelos Tsiaras

Arcetri Fellow
Osservatorio Astrofisico di Arcetri

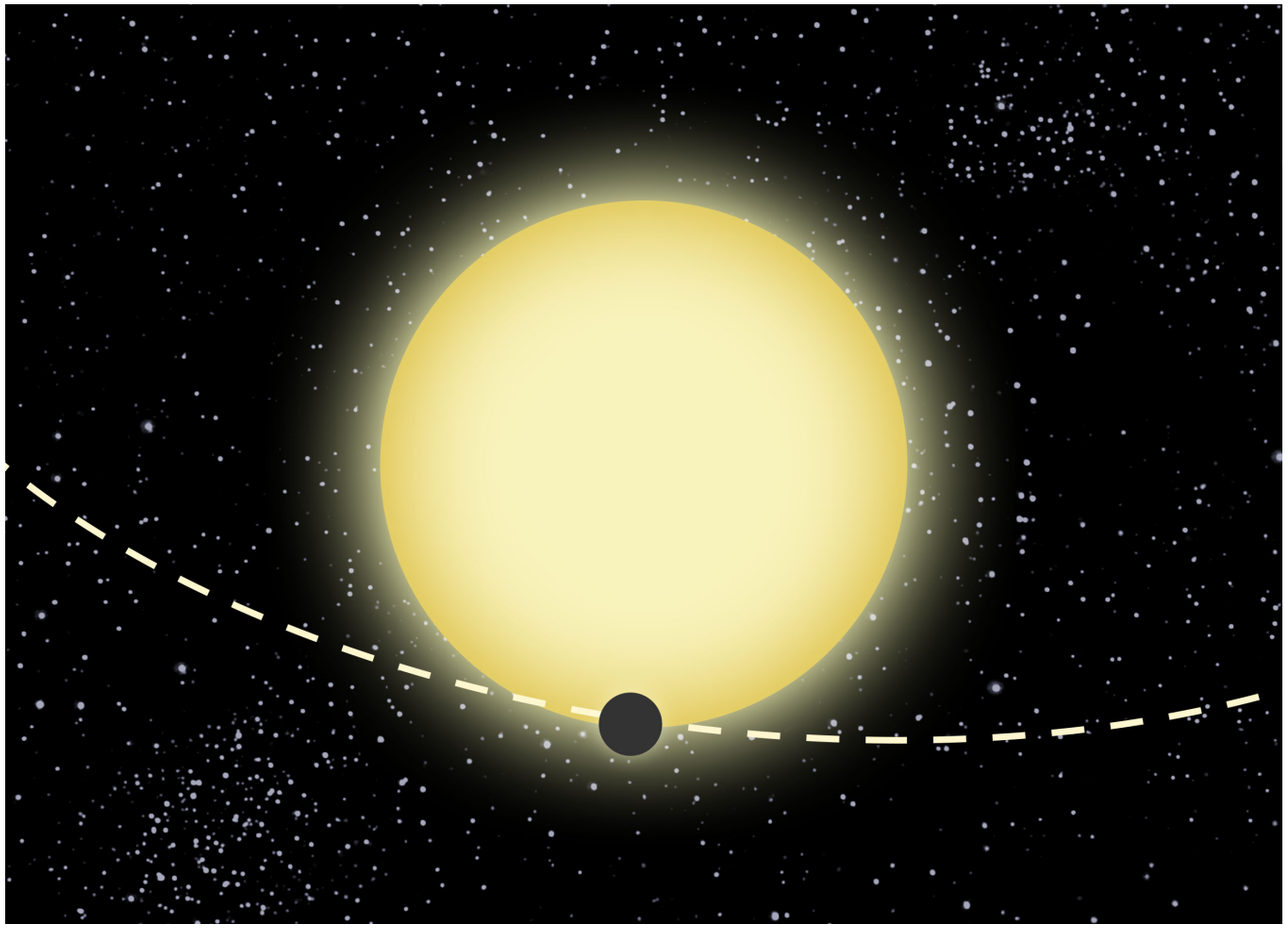
OSSERVATORIO ASTROFISICO DI ARCETRI

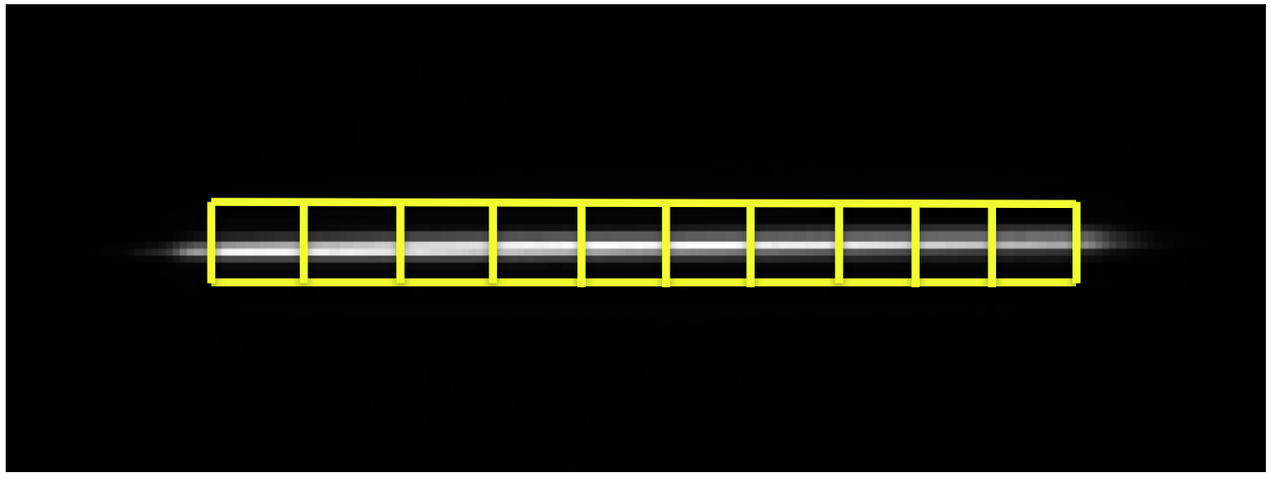
Transmission and emission spectroscopy

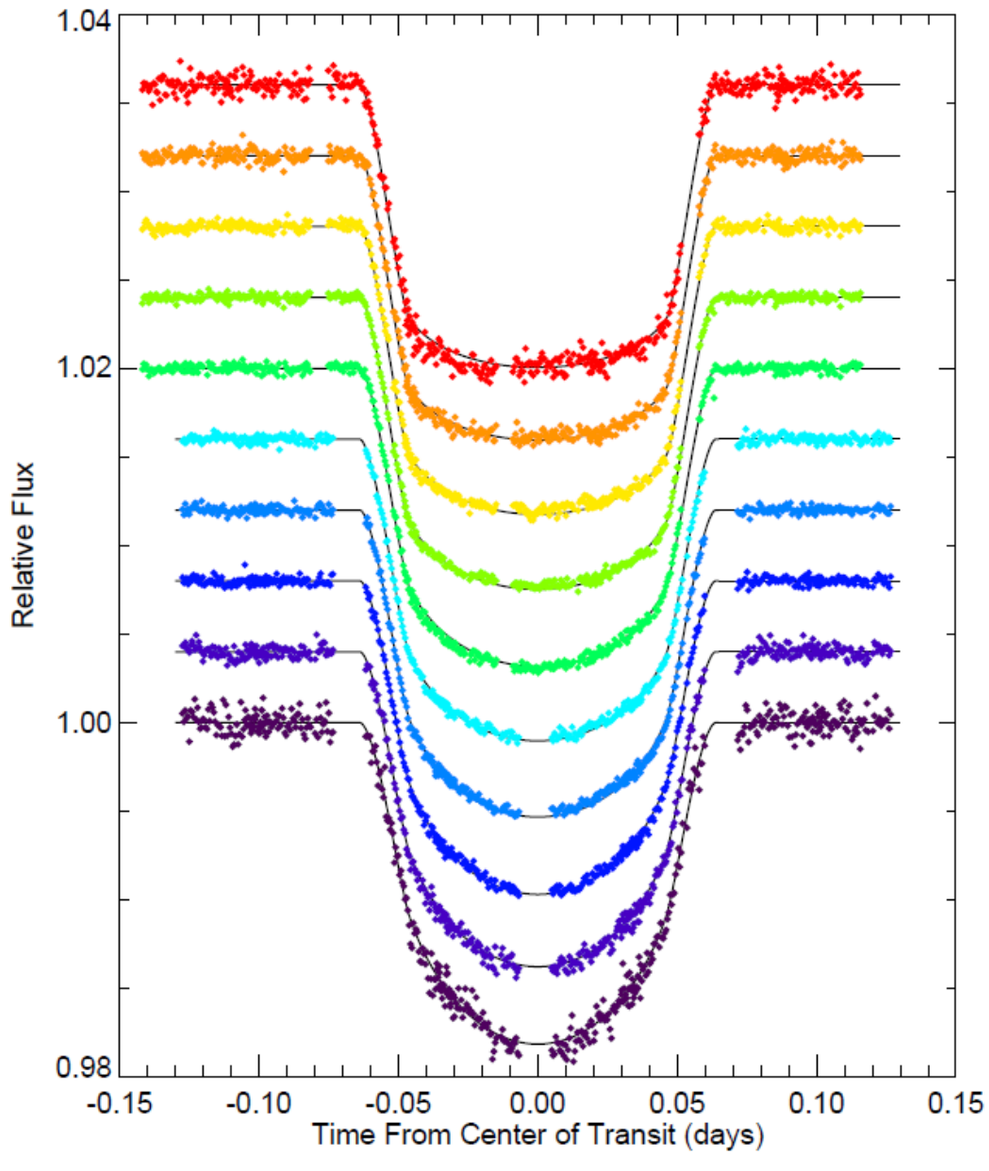


Hecht 2006

Transmission spectroscopy







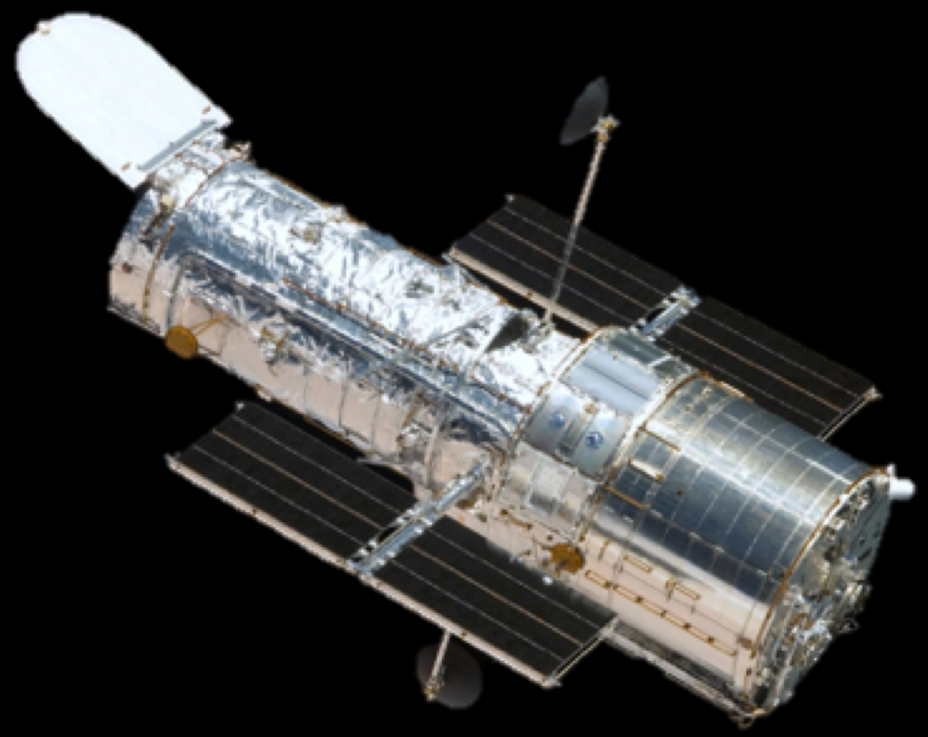
Knutson et al. 2007

The legacy of HST/WFC3:
Towards population studies of exoplanet atmospheres

WIDE-FIELD CAMERA 3 - WFC3
installed in 2009 during HST Service Mission 4

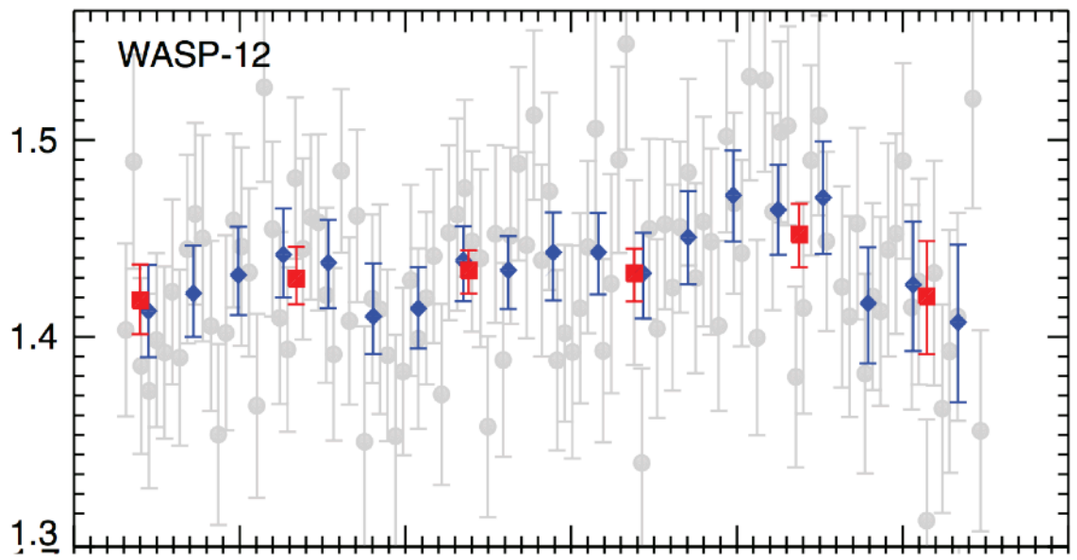
Step 1 - High S/N observations

Wavelength

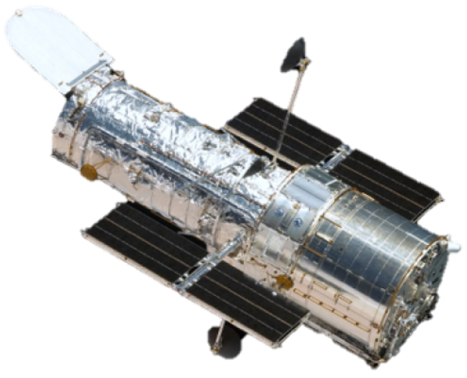


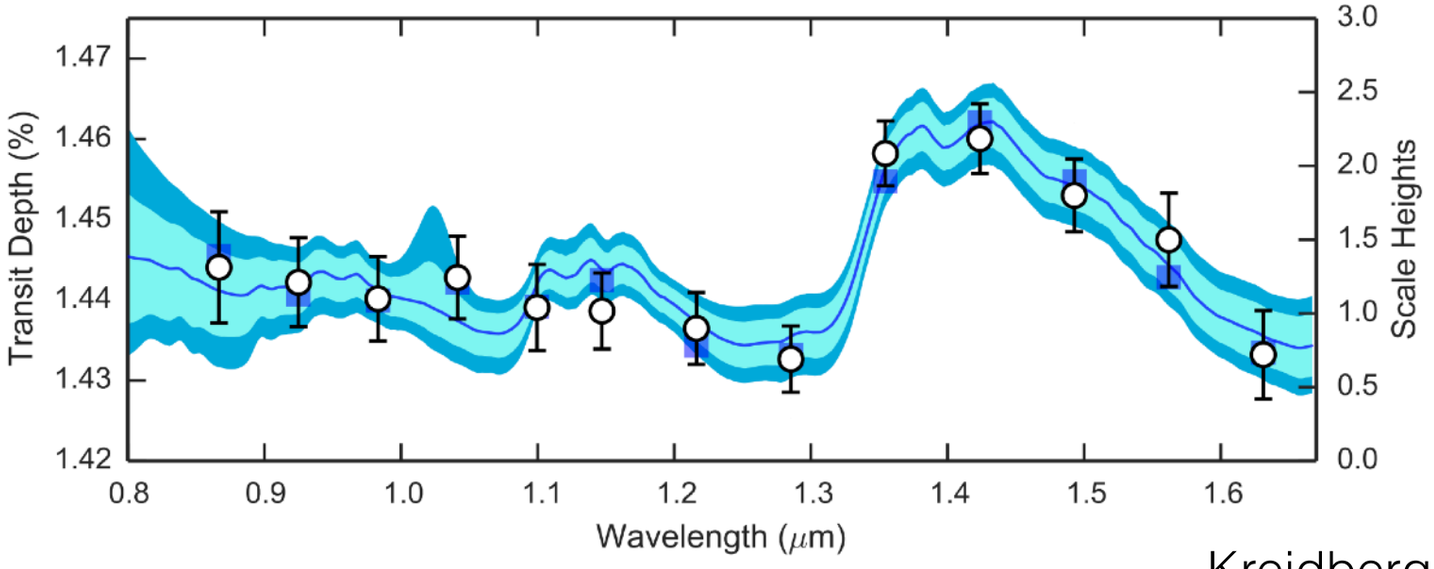
Step 1 - High S/N observations

WASP-12 b



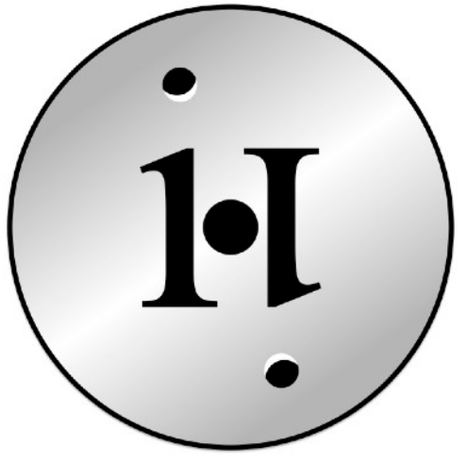
Mandel et al. 2013





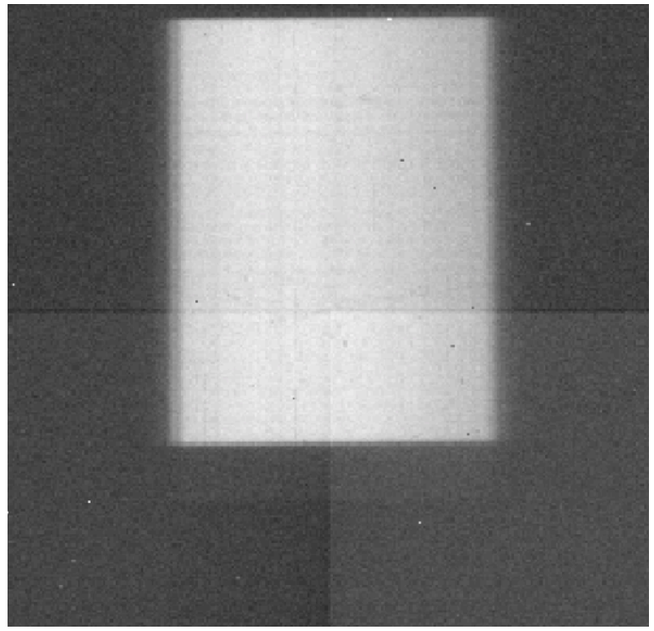
Kreidberg et al. 2015

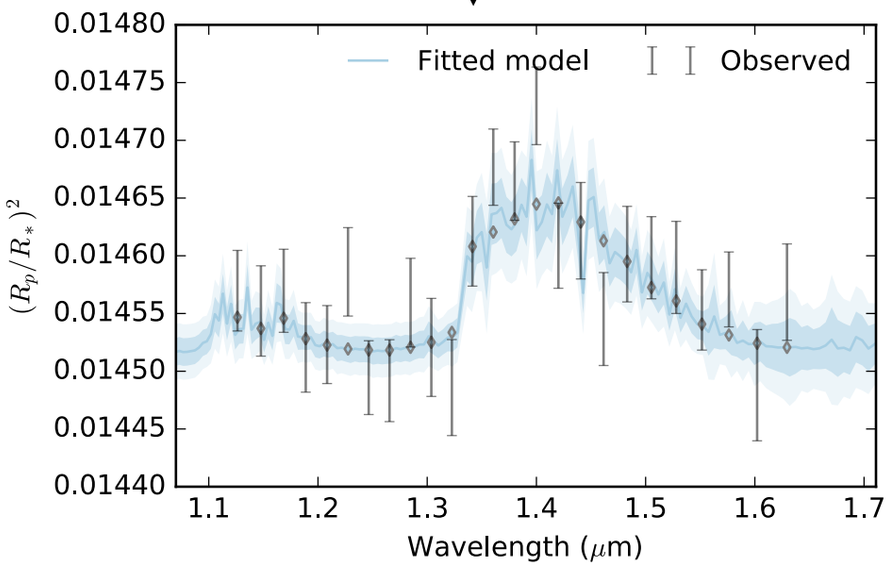
Step 2 - *A pipeline*



IRACLIS

- ☑ Specialised for WFC3 scanning
- ☑ Specialised for exoplanets
- ☑ Reliable and fast
- ☑ Automated end-to-end analysis
- ☑ Open source



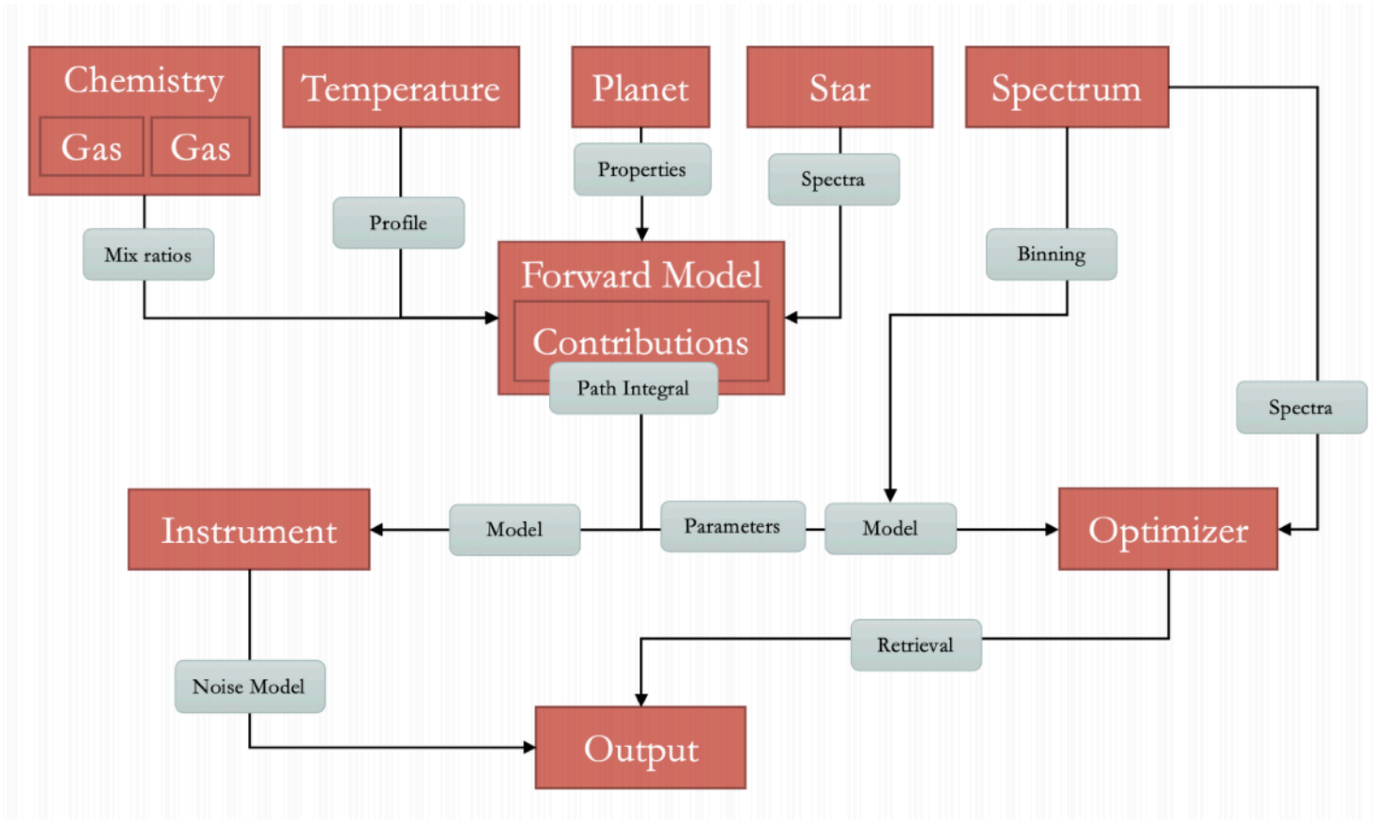


Tsiaras 2017 - PhD thesis

Step 3 - A retrieval

Tau-Rex 3

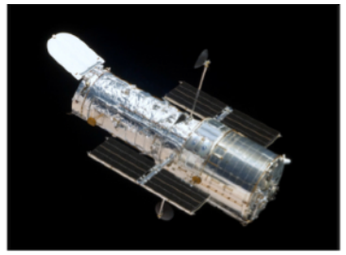
- ☑ Fully Bayesian
- ☑ Flexible modeling
- ☑ Reliable and fast
- ☑ Automated end-to-end analysis
- ☑ Open source

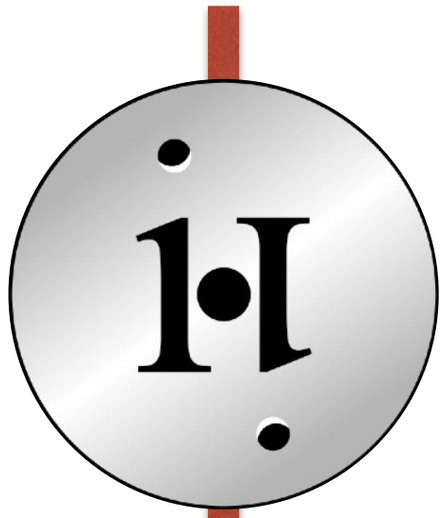


Al Refaie et al. 2019

Step 4 - OPEN SCIENCE

The new standard in data analysis



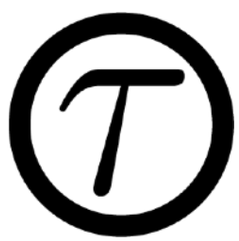


spectrum

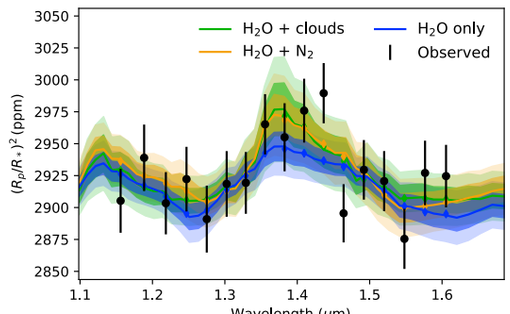
archive.stsci.edu

github.com/ucl-exoplanets

exomol.com



ExoMol



First catalogue of exoplanet atmospheres

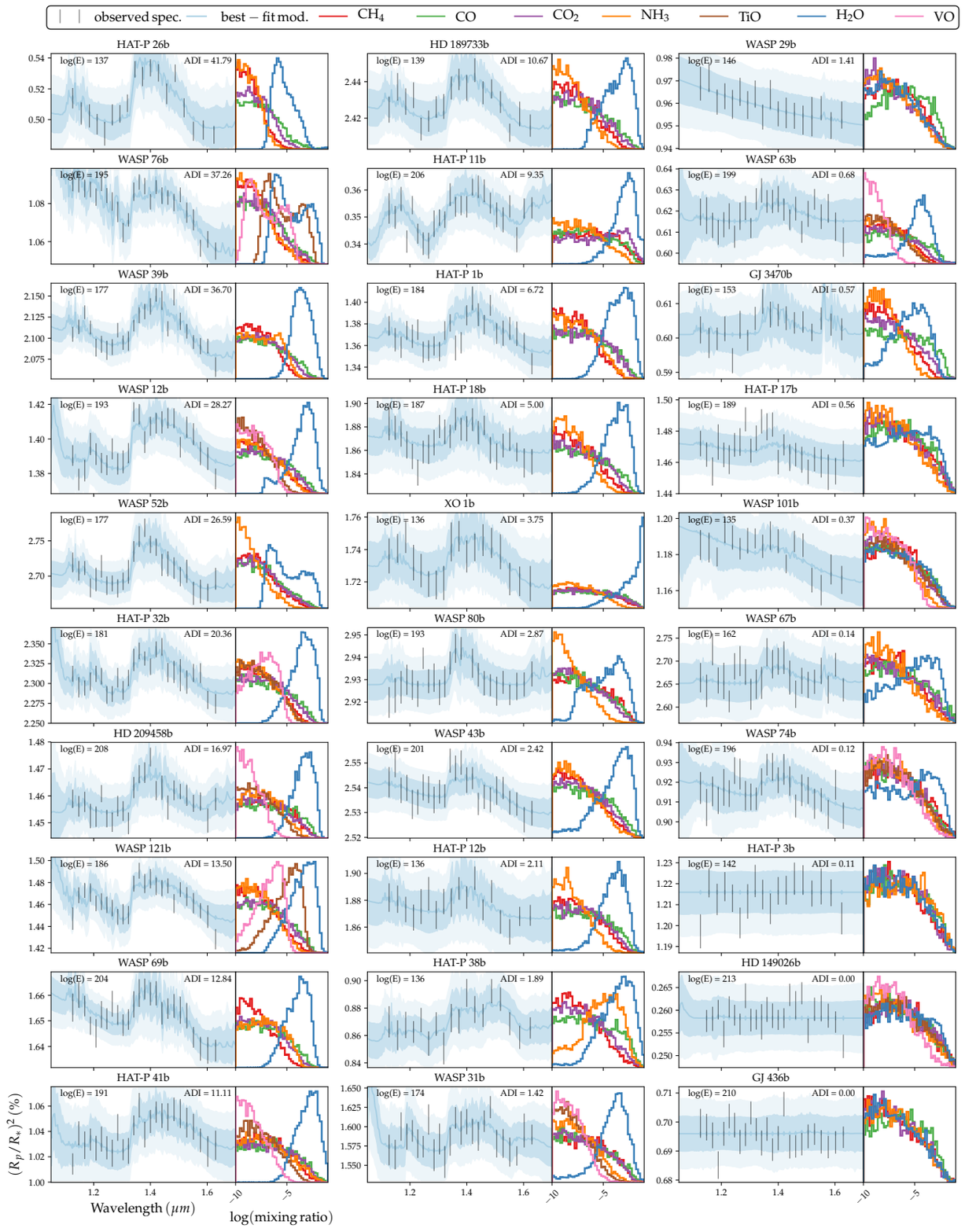
- **50** Datasets analysed
  largest data volume ever

- **30** exoplanet atmospheres
  largest sample ever

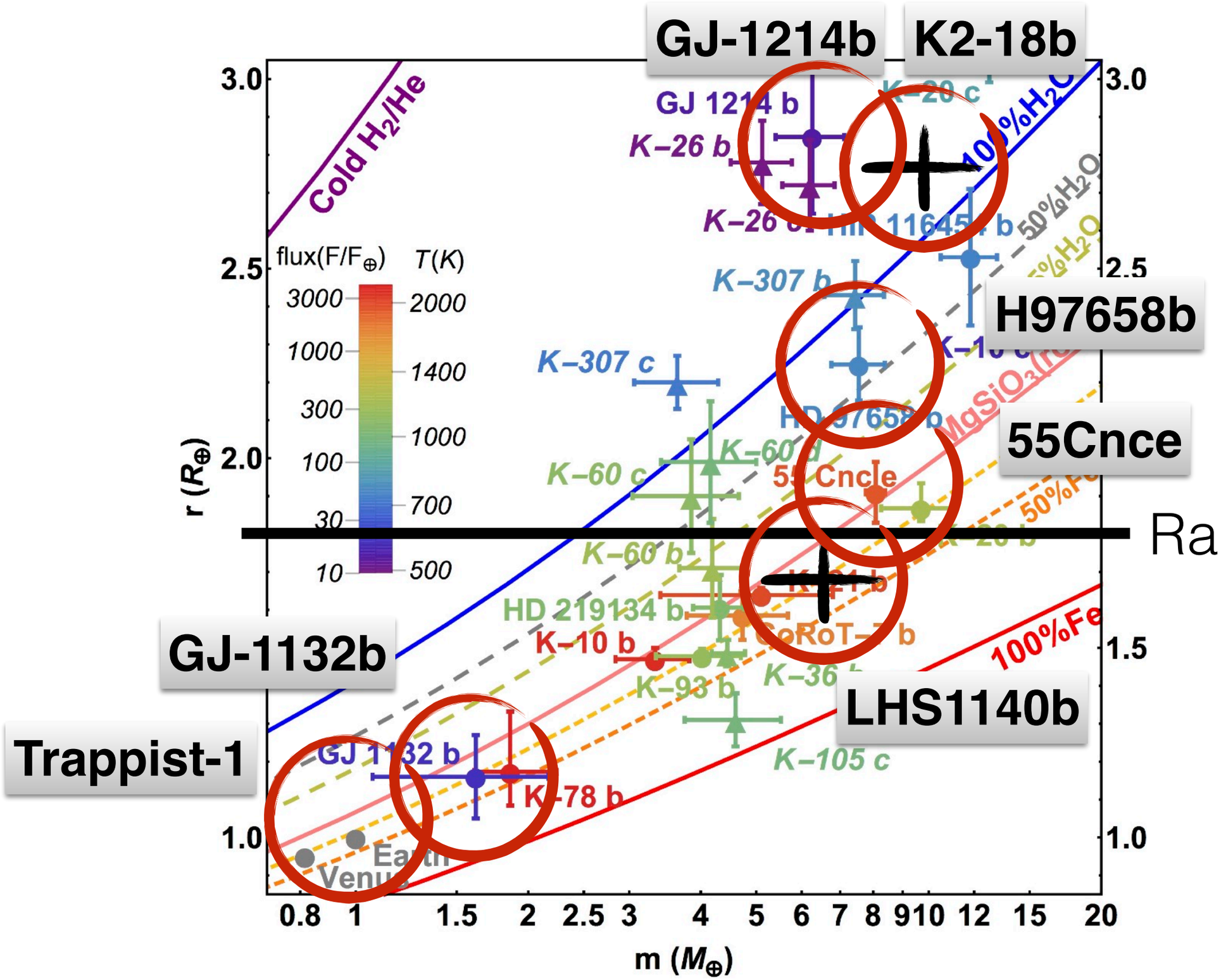
- **16** clear detections with **H2O**

Tsiaras et al. 2018

- **And many more in the years that followed**



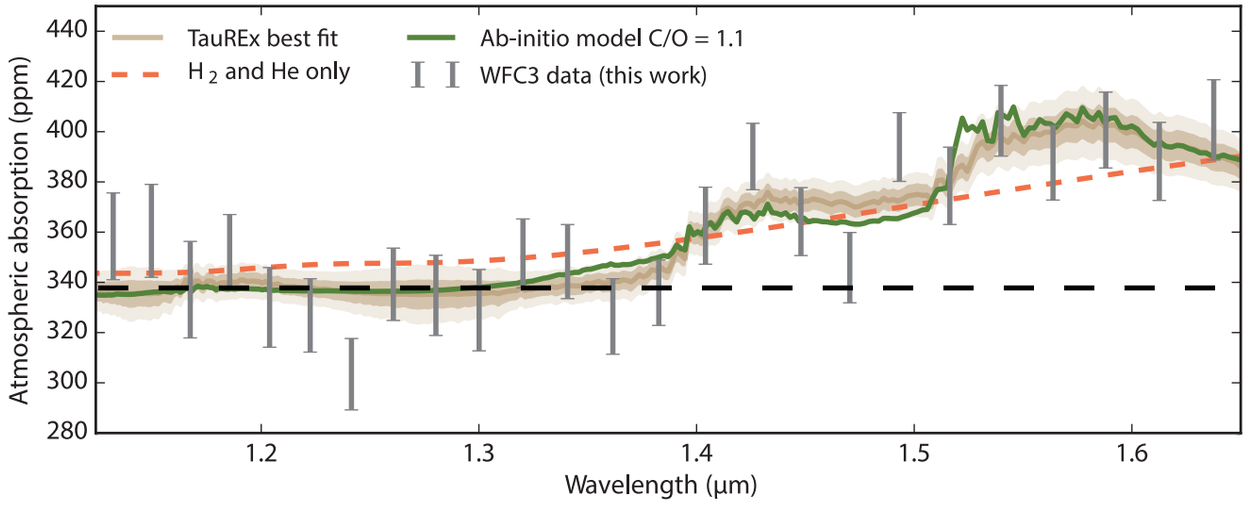
First observations of Super-Earths



First detections of super-Earth atmospheres

55 Cancri e

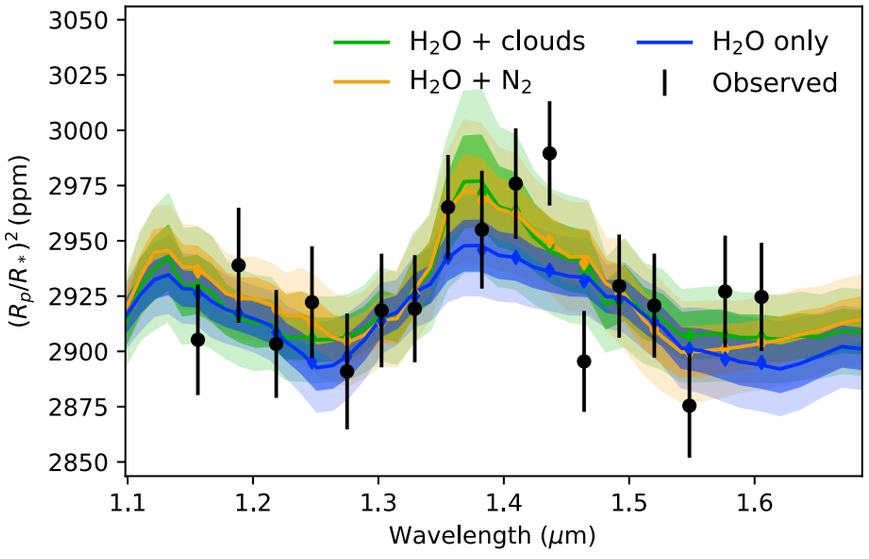
- +2000 K Temperature
- A light atmosphere (high C/O?)



Tsiaras et al. 2016

K2-18 b

- ~300 K Temperature
- H2O detection

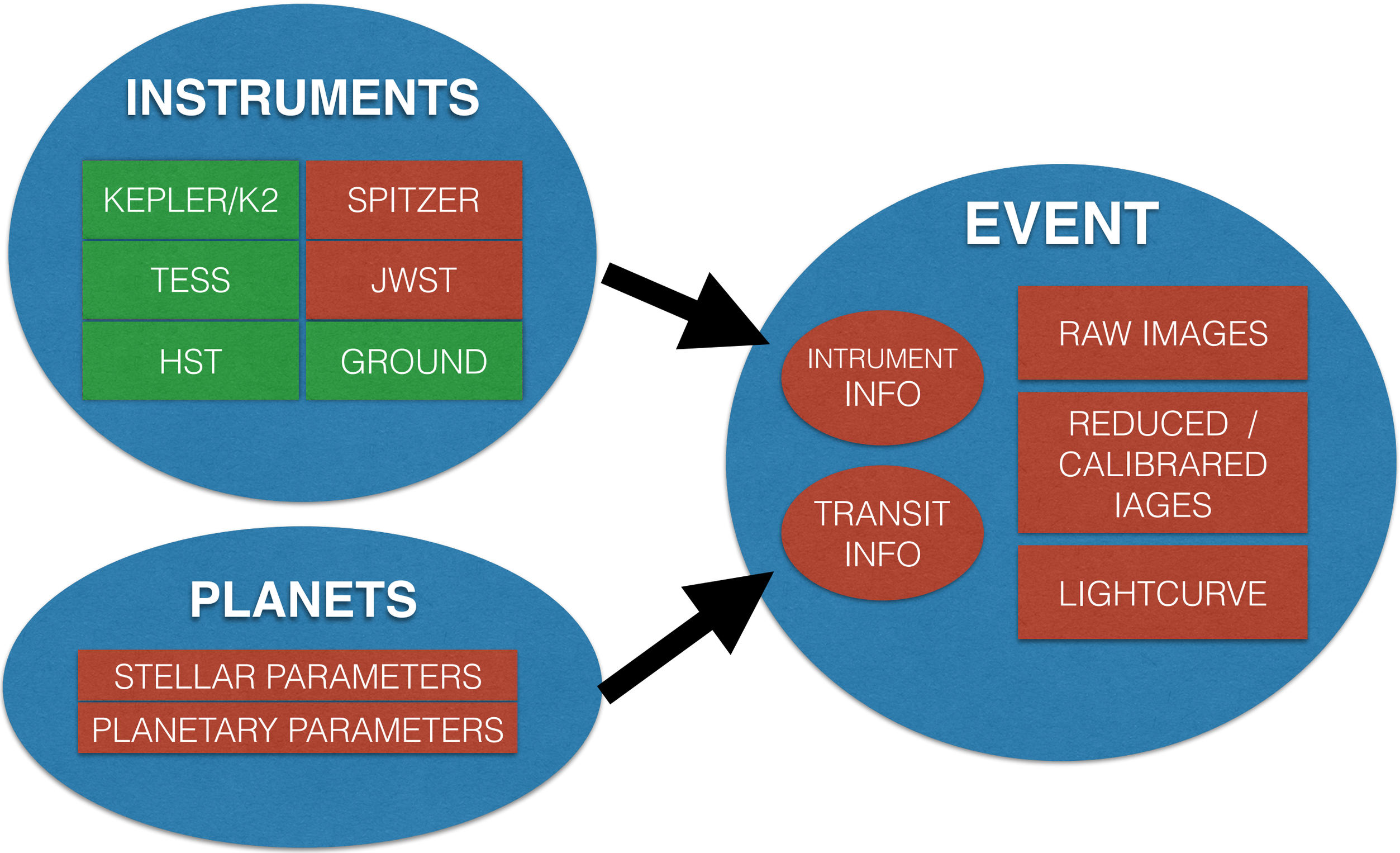


Tsiaras et al. 2019 - Nature Astronomy

The challenges…

1. Database
   - **Easy access to past observations**
   - **Multi-level products**
2. Follow-up
   - **Robust parameters to select targets and to schedule observations**
3. Pipeline
   - **Multiple instruments**
   - **Multiple approaches to de-trending**
   - **HPC capabilities**

Database


INSTRUMENTS
KEPLER/K2
SPITZER
TESS
JWST
HST
GROUND
EVENT
INTRUMENT INFO
TRANSIT INFO
RAW IMAGES
REDUCED / CALIBRARED IAGES
LIGHTCURVE
PLANETS
STELLAR PARAMETERS
PLANETARY PARAMETERS

Database: 600 known planets

HST
~600 transits
and growing…

Kepler/K2
~6000 transits

TESS
~6000 transits
and growing…

Ground
~4000 transits
and growing…

Follow-up: the ExoClock project

www.exoclock.space

Space
- Kepler/K2
- TESS
- HST
- Spitzer
- CHEOPS

Ground
- ExoClock Network
- ETD
- Exoplanet Watch
- TelescopeLive
- LCO

Literature

Reliable catalogue of transit parameters

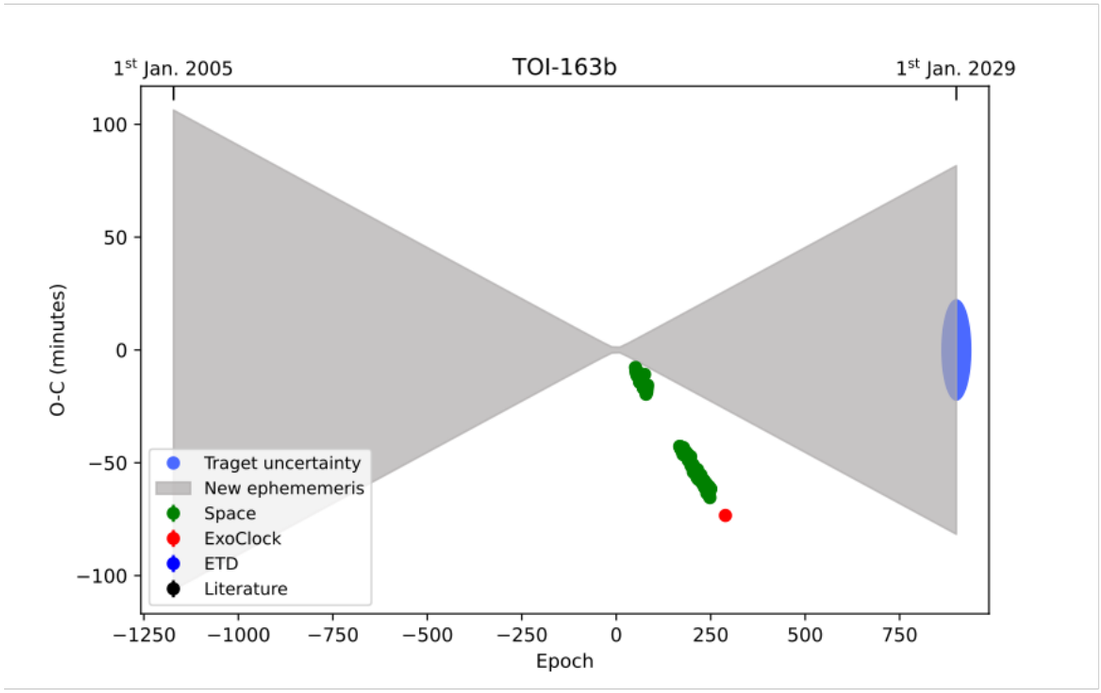
- Ephemeris
- Depth
- Duration

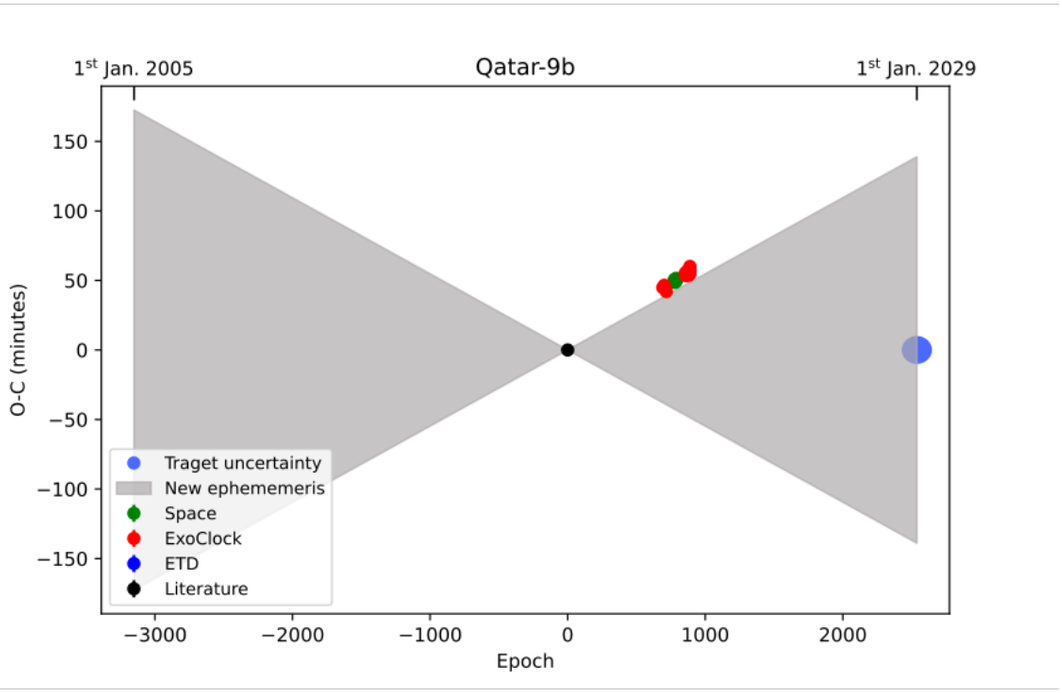
Kokori et al. 2021, 2022

Follow-up: fixing parameters…

Initial parameters
25% of the planets need to be fixed

Ephemerides
45% of the planets need to be fixed





Follow-up: the ExoClock project

500 participants from 35 countries

4000 observations for 350 planet

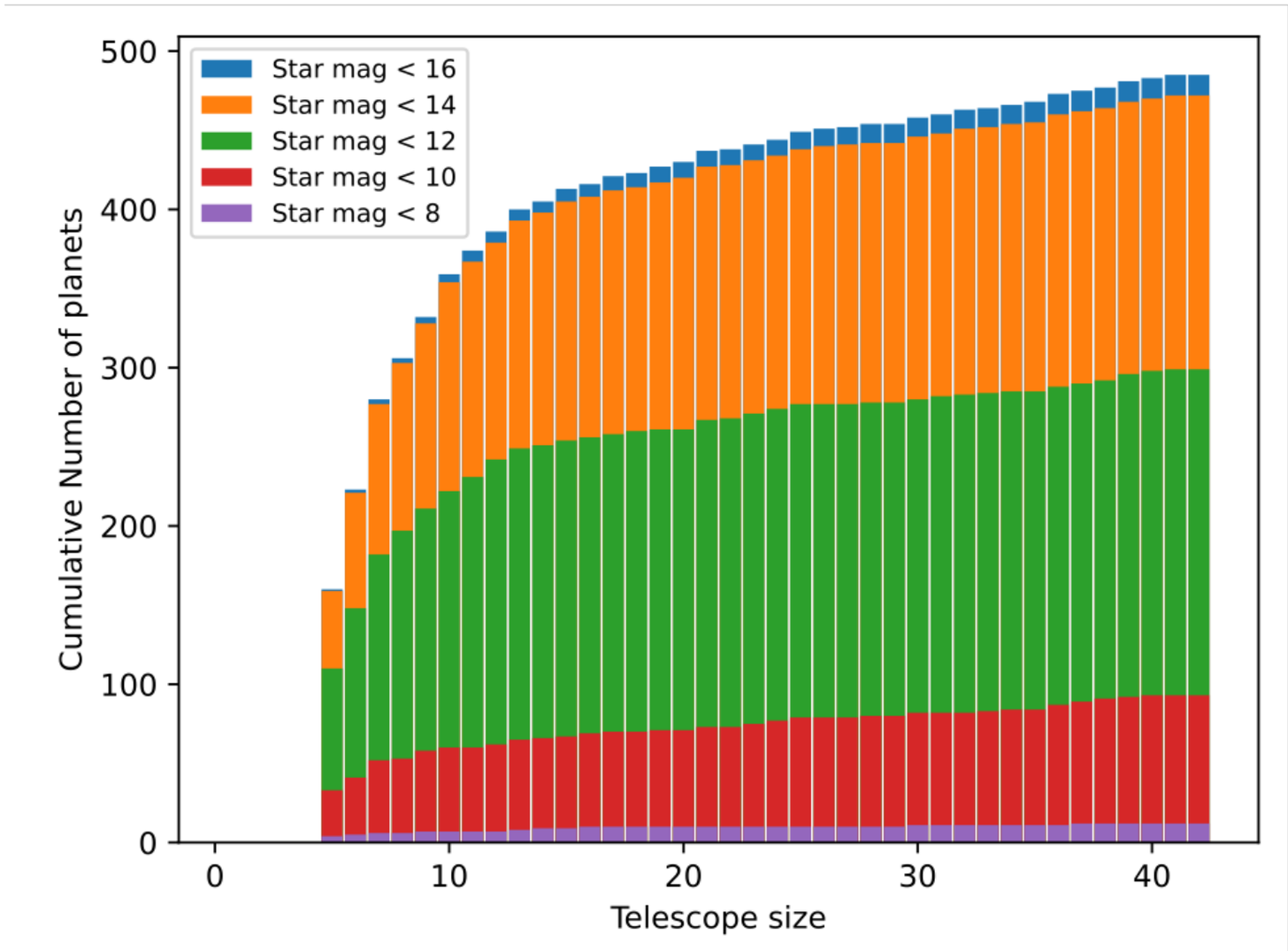
2 Data Releases

- Create a global community
- Involve and train citizens scientist
- Reach out many communities

Kokori et al. 2021, 2022

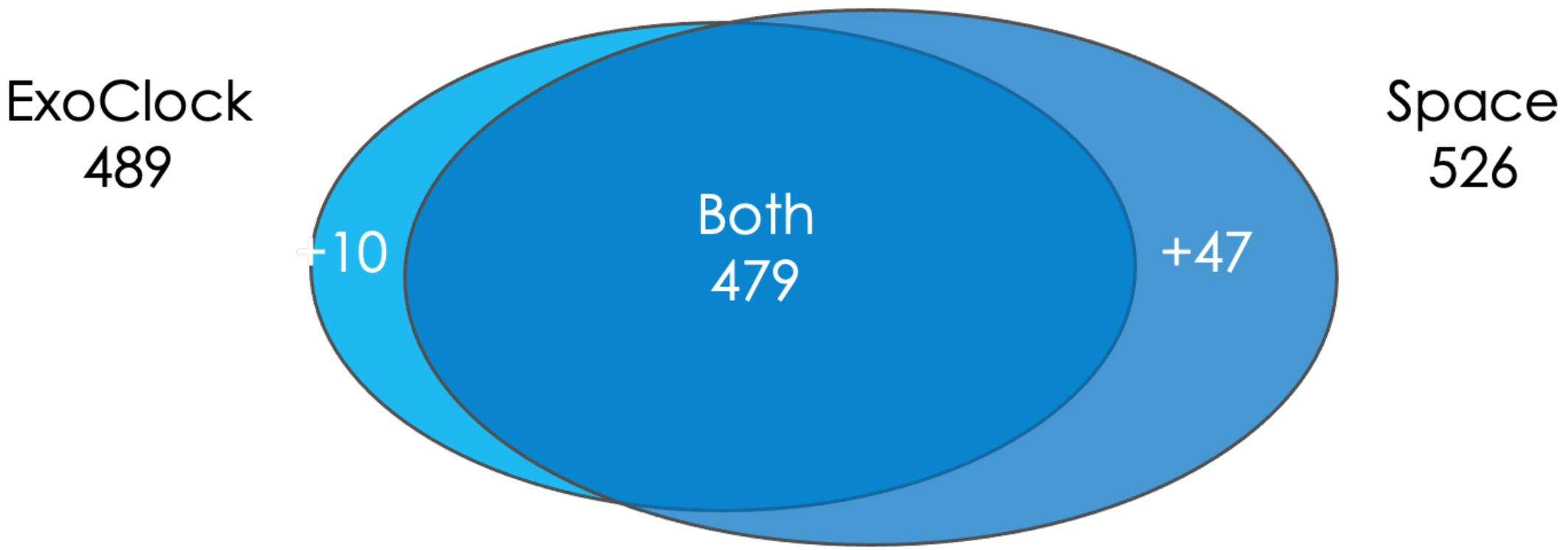
The power of small telescopes

70% of the targets can be observed by telescopes < 20 inches



Space and Ground

Target coverage - achievable – 94%
(ignoring TESS gaps or long-period planets)



Pipeline - multiple instruments

REDUCTION

- Non-linearity
- Bias
- Dark
- Flat
- Persistance
- Gain

BACKGROUND

- Sky
- Cosmic Rays

CALIBRATION

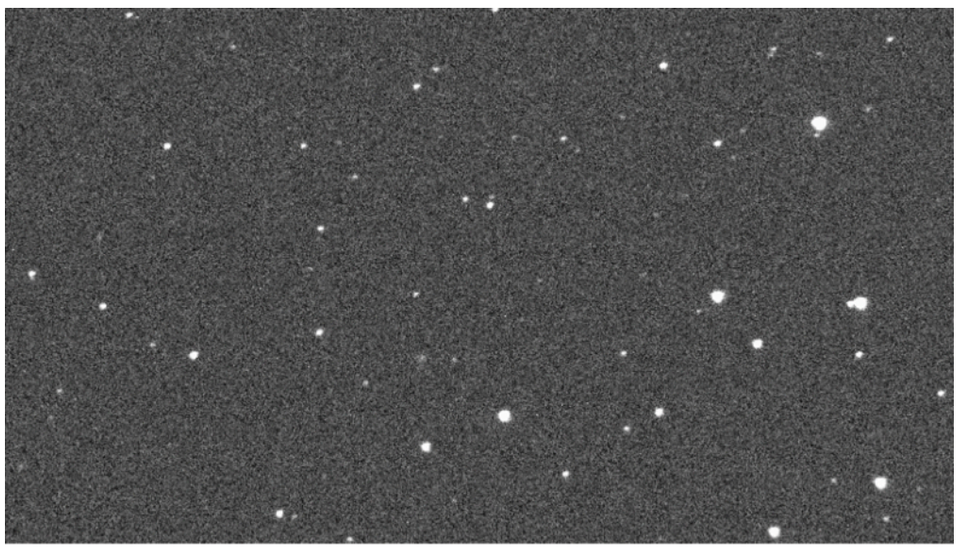
- Drifts
- Rotation

Extraction

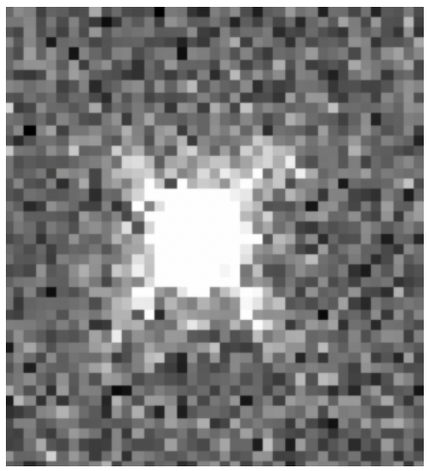
- Optimal Aperture
- Fractional pixels

Pipeline - multiple instruments

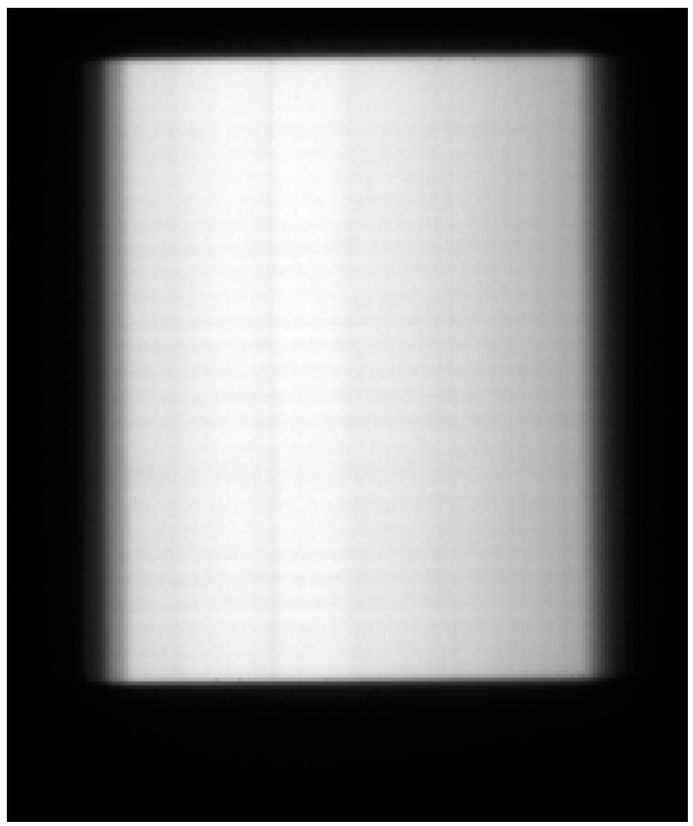
Ground telescope
Photometry



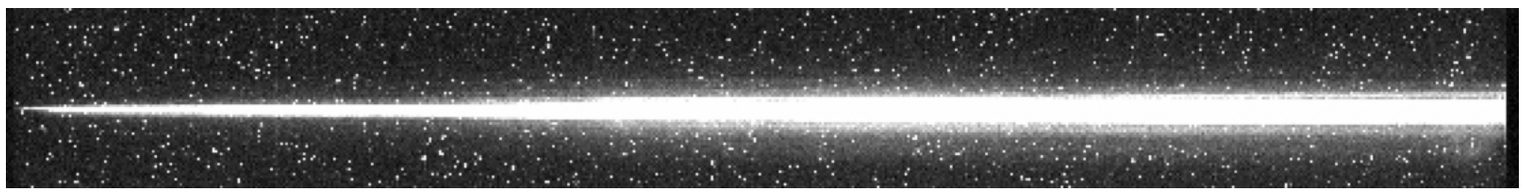
HST/WFC3
Photometry



HST/WFC3
Spectroscopy



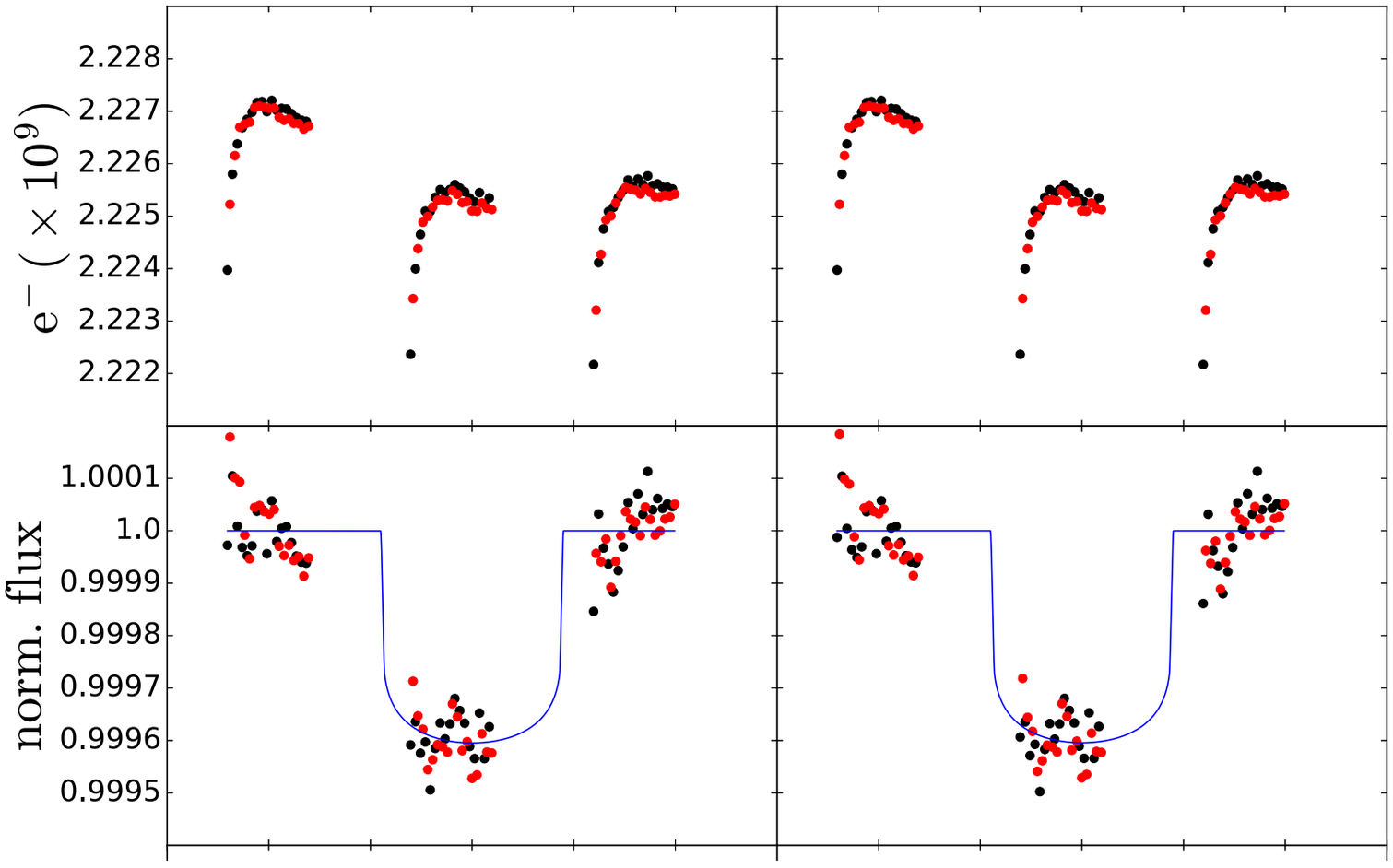
HST/STIS - Spectroscopy



ARIEL- Spectroscopy

Pipeline - multiple approaches to de-trending

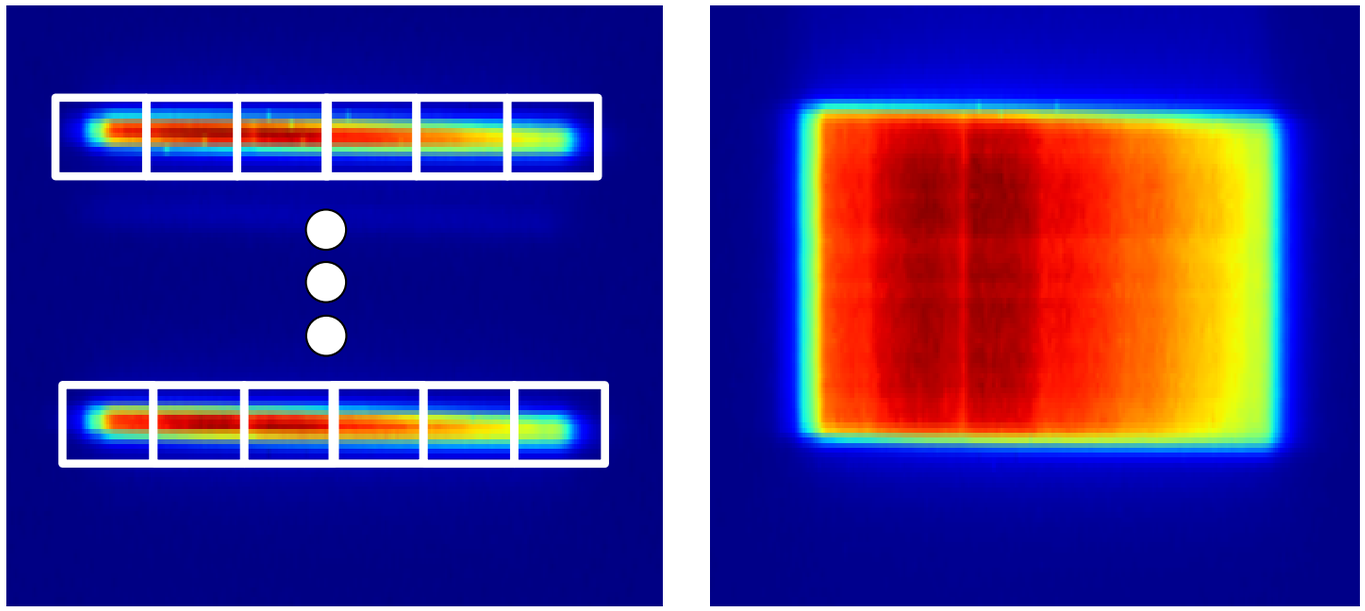
Classic parametrised de-trending on HST

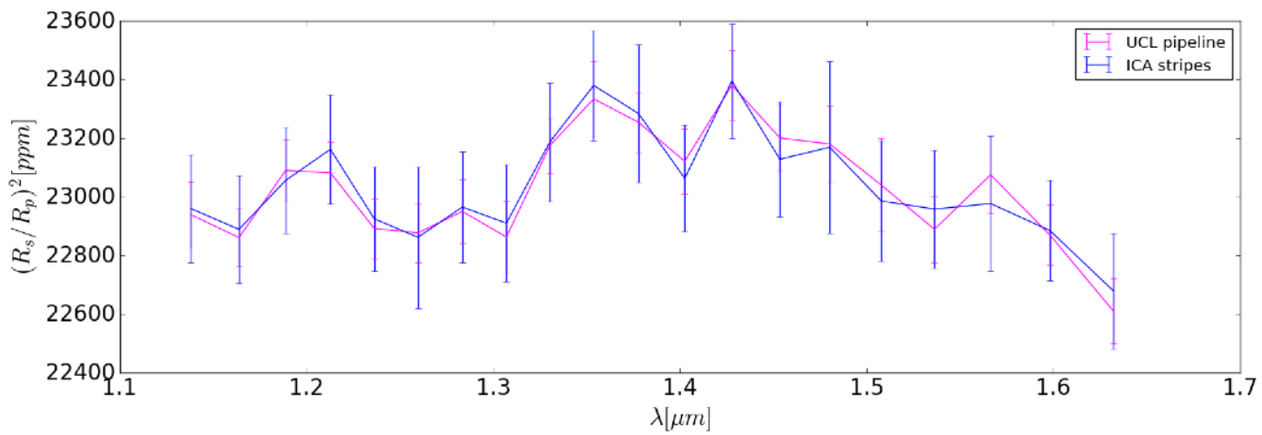


Tsiaras et al. 2016

Pipeline - multiple approaches to de-trending

ICA applied on HST

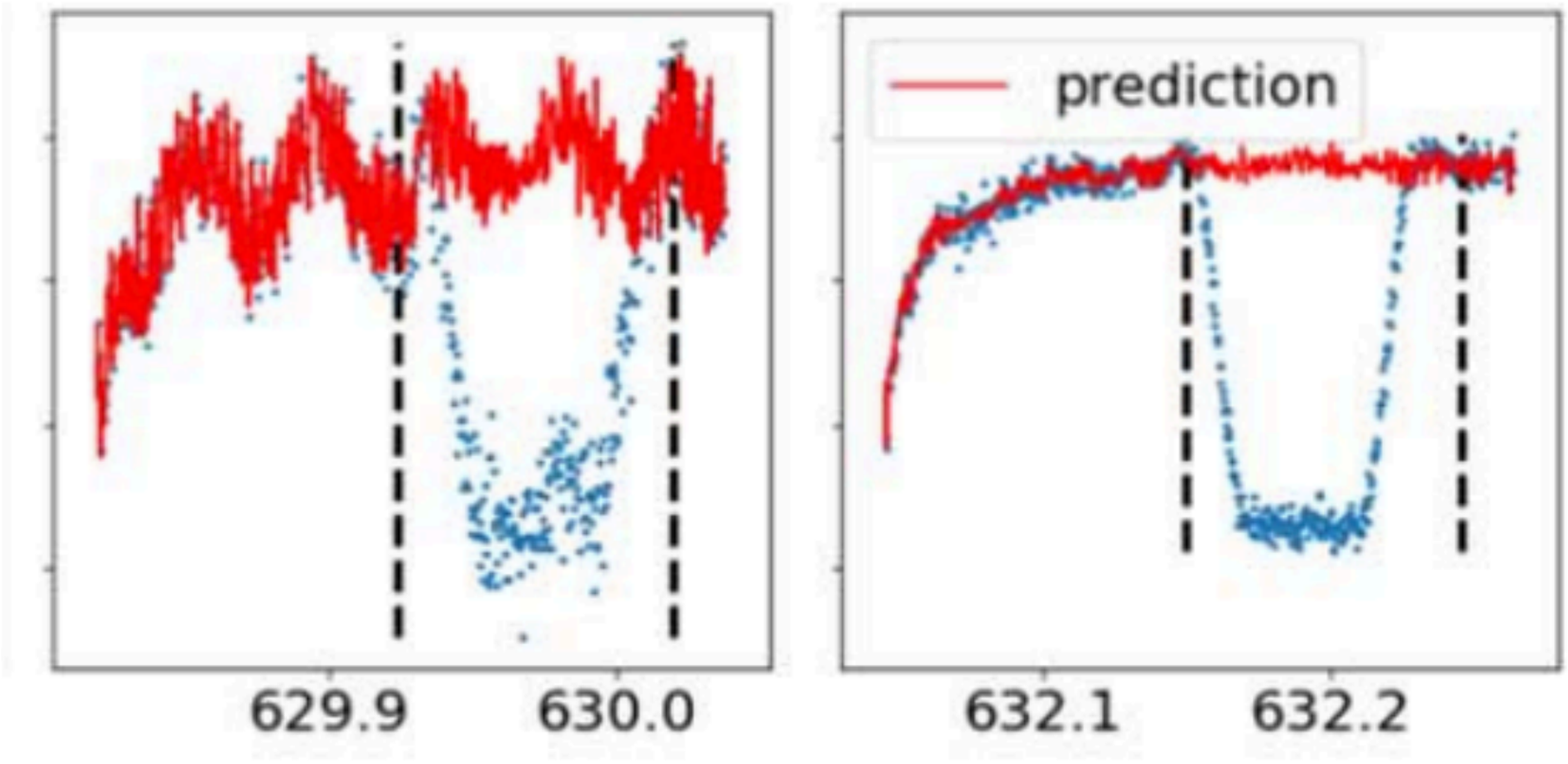




Damiano et al. 2017

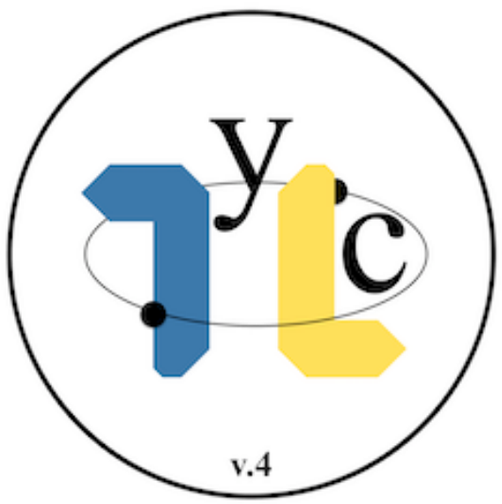
Pipeline - multiple approaches to de-trending

LSTM networks applied on Spitzer

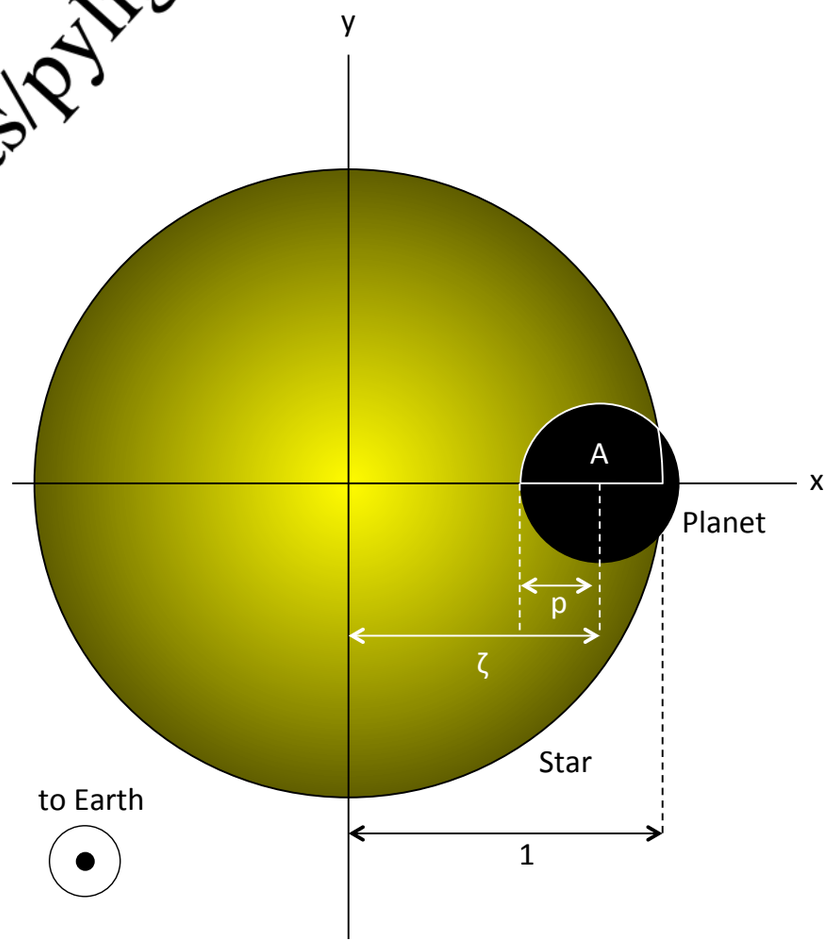


Morvan et al. 2020

Pipeline - HPC capabilities



check on GitHub
github.com/ucl-exoplanets/pylightcurve



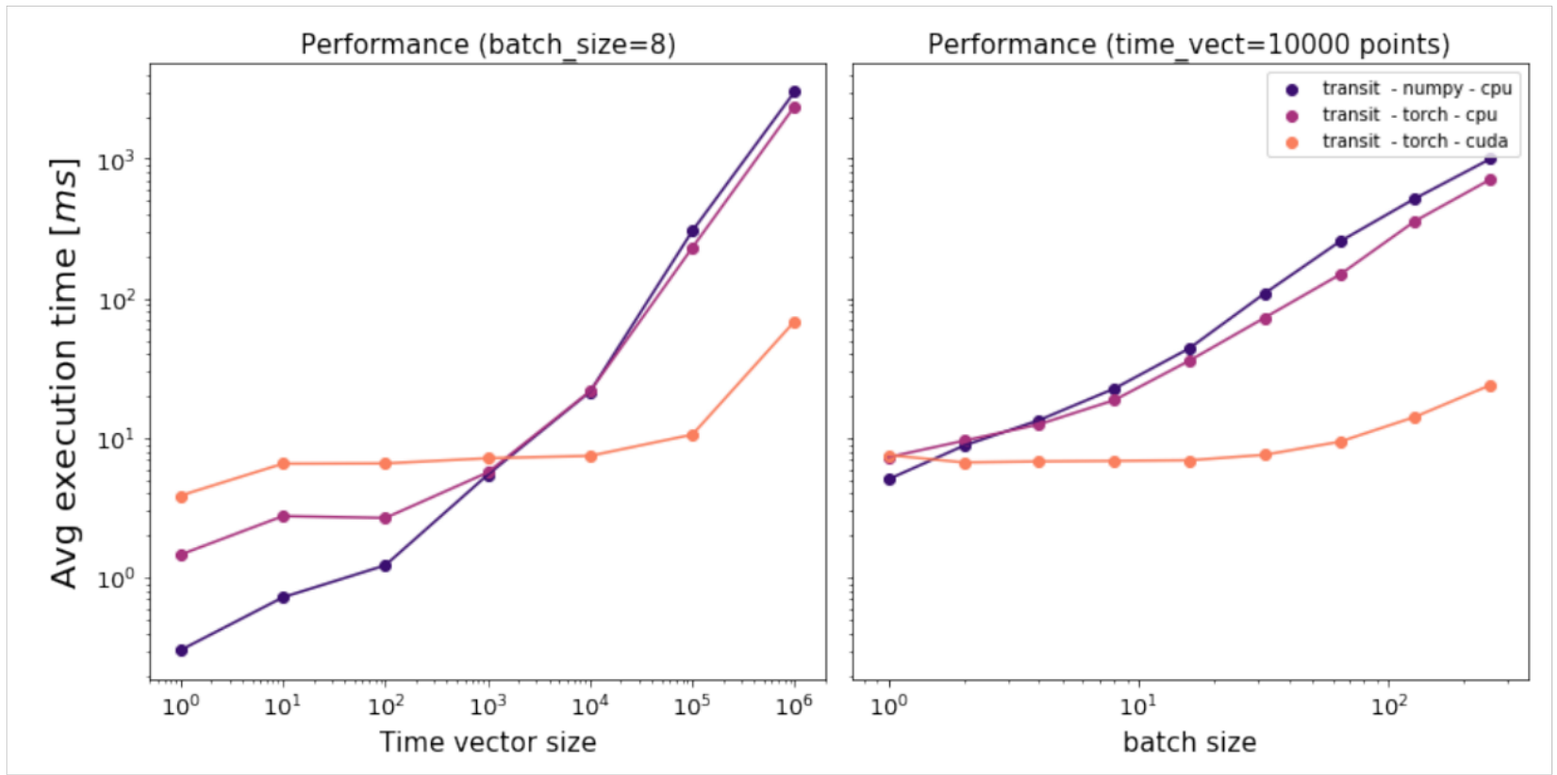
Comparison with Batman

Array-size	Speed increase
2500	x35
5000	x25
10000	x15
20000	x5

Pipeline - HPC capabilities

check on GitHub
github.com/ucl-exoplanets/pylightcurve-torch

- GPU compatibility
- Real large-scale applications
- Deep learning applications



PyLIghtcurve-torch package: Morvan+ 2020

Population studies:
a **simple** concept, a **challenging** plan

1. Prepare an **open, multi-level database**.

2. **Follow up** the targets and **fix the parameters**.

3. Prepare an **open, stand-alone data analysis framework**.