

# The ARIEL Data Analysis Challenge

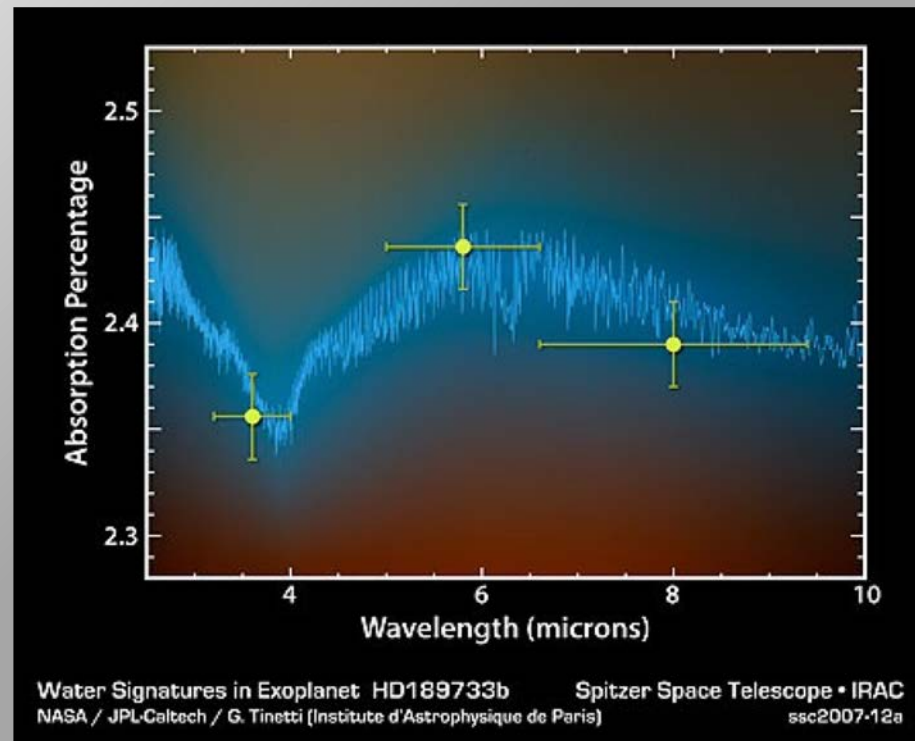
Angelos Tsiaras  
UCL CSED



Horizon 2020

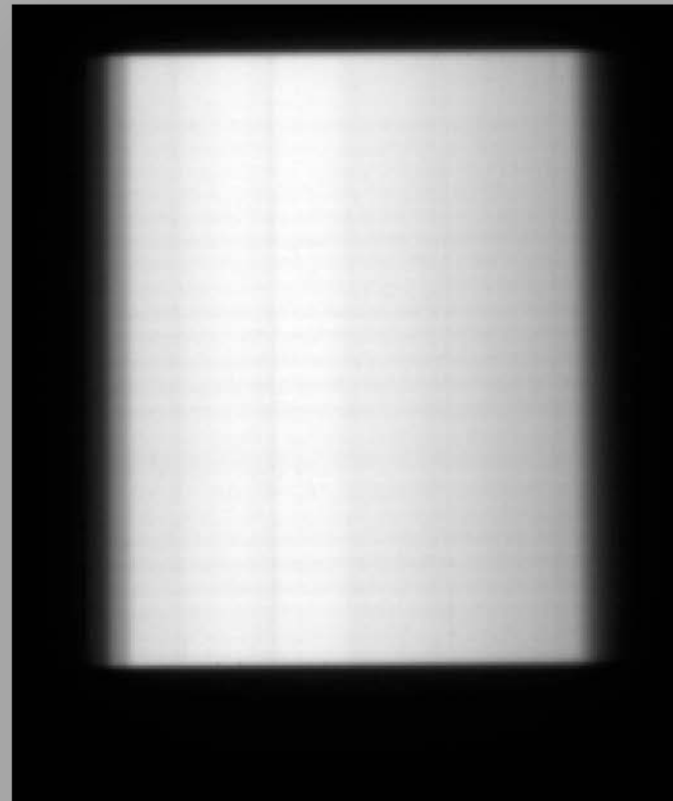
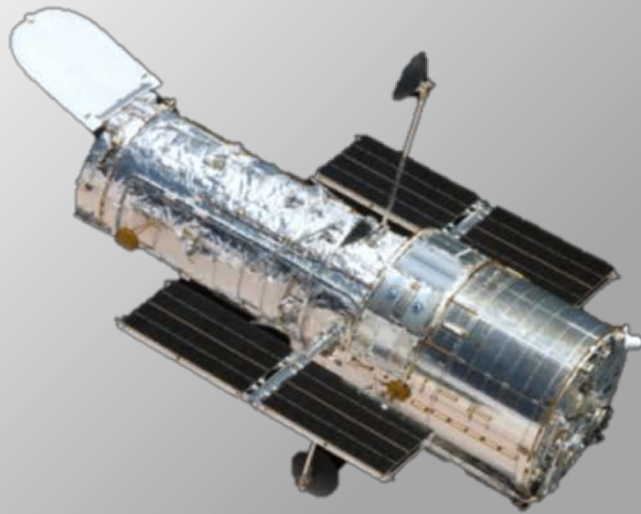


# The era of exoplanet characterisation

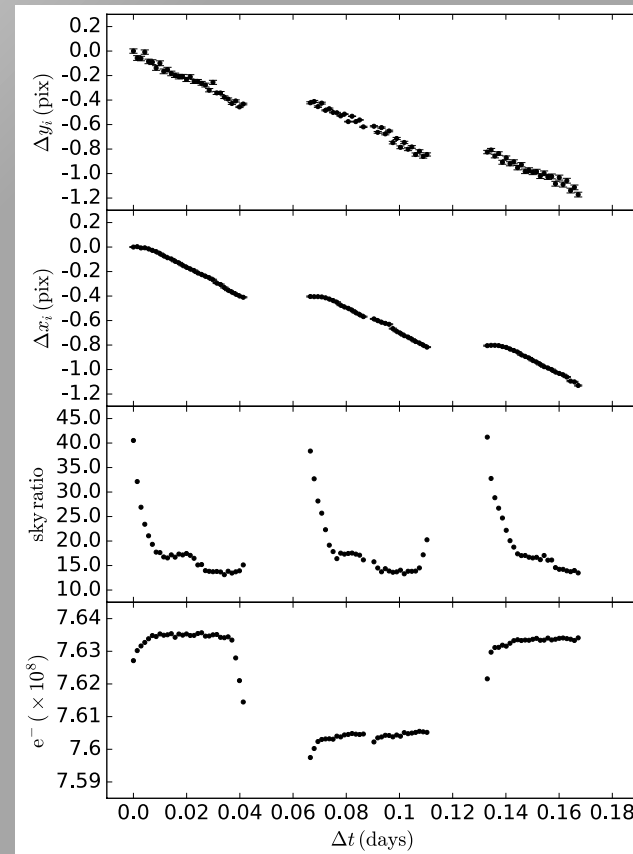
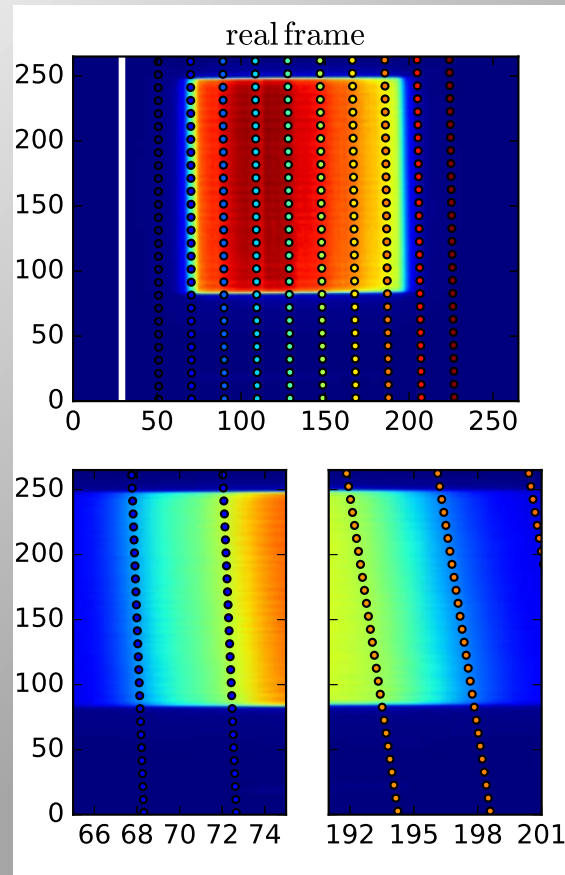


Tinetti et al. 2007

# HST / WFC3 – spatial scanning



# HST / WFC3 – spatial scanning



# HST / WFC3 – key components

- Data simulator
- Automated pipeline
- Lots of trials...

# ARIEL Data products

**Automatic Pipeline Processing Steps**  
Level 0 ---> Level 2

**Level 0 data**

- Decompression
- Unpacking
- Metadata extraction

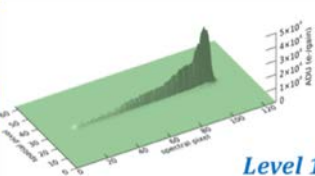
**Level 1 data**

- Flag saturated pixels
- ADU to e<sup>-</sup> unit conversion
- Detector non-linearity correction
- Pixel cross-talk correction
- Dark current subtraction
- Wavelength assignment
- Pixel QE disuniformity flatfielding
- Offset (bias variation) correction
- Detector persistence correction
- Background subtraction
- e<sup>-</sup> to e<sup>-</sup>/s unit conversion
- Bad pixel correction
- Inter-exposure pointing jitter correction

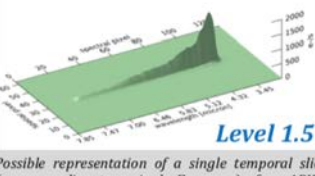
**Level 1.5 data**

- Aperture masking (spectroscopic channels)
- Extract 1-D spectrum
- Aperture masking (photometric channels)
- Spectral lightcurve binning
- e<sup>-</sup>/s to flux conversion

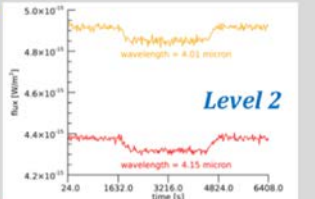
**Level 2 data**



*Possible representation of one temporal slice (i.e.: one Science Frame) of an ARIEL Data Level 1 Data Product, for the spectroscopy channel AIRS CH0. Signal units (z axis) are in ADU, but, since the image is produced from simulated data (without many instrumental features), and given the currently unknown gain values, the reported quantities are in units of e<sup>-</sup>/gain.*



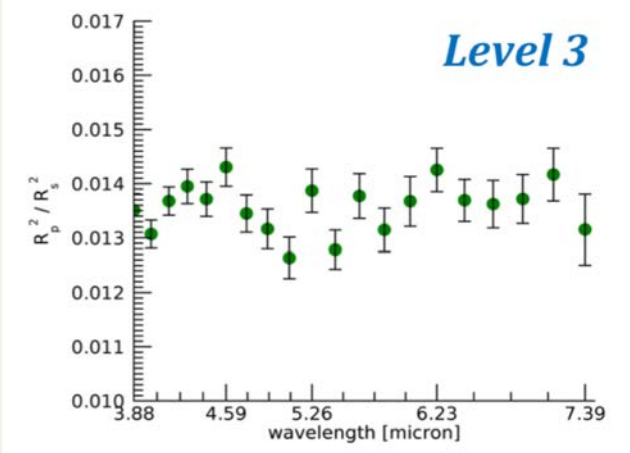
*Possible representation of a single temporal slice (corresponding to a single Exposure) of an ARIEL Level 1.5 Data Product (without many instrumental features), for the spectroscopy channel AIRS CH0.*



*Possible representation of ARIEL Level 2 Data Product for the spectroscopy channel AIRS CH0, showing the binned light curves.*

**Customised Data Processing Steps:**  
Level 2 – Level 3

**Level 3**



*Possible representation of ARIEL Level 3 Data Products.*

- Instrumental systematics correction (detrrending)
- Stellar activity correction 1 (lightcurves)
- Lightcurves transit models fitting
- Stellar activity correction 2 (spectra)

**Level 3 data**

# ARIEL Data products

**Automatic Pipeline Processing Steps**  
Level 0 ---> Level 2

**Level 0 data**

- Decompression
- Unpacking
- Metadata extraction

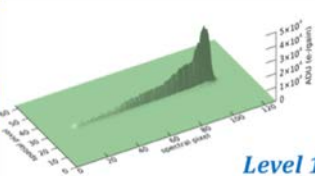
**Level 1 data**

- Flag saturated pixels
- ADU to e<sup>-</sup> unit conversion
- Detector non-linearity correction
- Pixel cross-talk correction
- Dark current subtraction
- Wavelength assignment
- Pixel QE disuniformity flatfielding
- Offset (bias variation) correction
- Detector persistence correction
- Background subtraction
- e<sup>-</sup> to e<sup>-</sup>/s unit conversion
- Bad pixel correction
- Inter-exposure pointing jitter correction

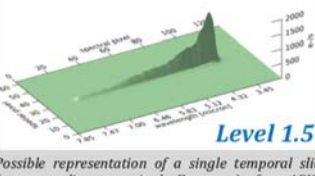
**Level 1.5 data**

- Aperture masking (spectroscopic channels)
- Extract 1-D spectrum
- Aperture masking (photometric channels)
- Spectral lightcurve binning
- e<sup>-</sup>/s to flux conversion

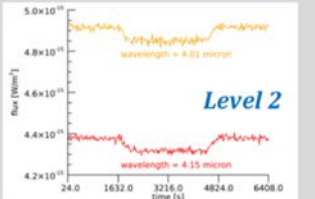
**Level 2 data**



*Possible representation of one temporal slice (i.e.: one Science Frame) of an ARIEL Data Level 1 Data Product, for the spectroscopy channel AIRS CH0. Signal units (z axis) are in ADU, but, since the image is produced from simulated data (without many instrumental features), and given the currently unknown gain values, the reported quantities are in units of e<sup>-</sup>/gain.*



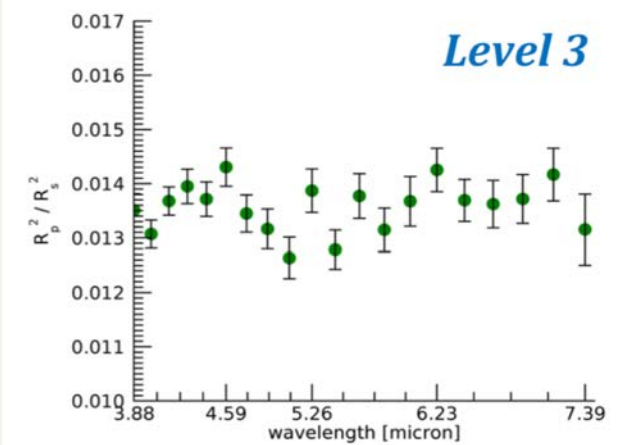
*Possible representation of a single temporal slice (corresponding to a single Exposure) of an ARIEL Level 1.5 Data Product (without many instrumental features), for the spectroscopy channel AIRS CH0.*



*Possible representation of ARIEL Level 2 Data Product for the spectroscopy channel AIRS CH0, showing the binned light curves.*

**Customised Data Processing Steps:**  
Level 2 – Level 3

**Level 3**



*Possible representation of ARIEL Level 3 Data Products.*

- Instrumental systematics correction (detrrending)
- Stellar activity correction 1 (lightcurves)
- Lightcurves transit models fitting
- Stellar activity correction 2 (spectra)

**Level 3 data**

# Data analysis challenge

- Produce end-to-end simulations for a couple of ARIEL observations
- Optimise the data-analysis process to deal with current systematics
- Identify parts of the data-analysis process that may introduce noise and develop a data-analysis strategy
- Evaluate the uncertainty introduced by fitting strategy
- Evaluate the level of noise correlation across wavelength bins



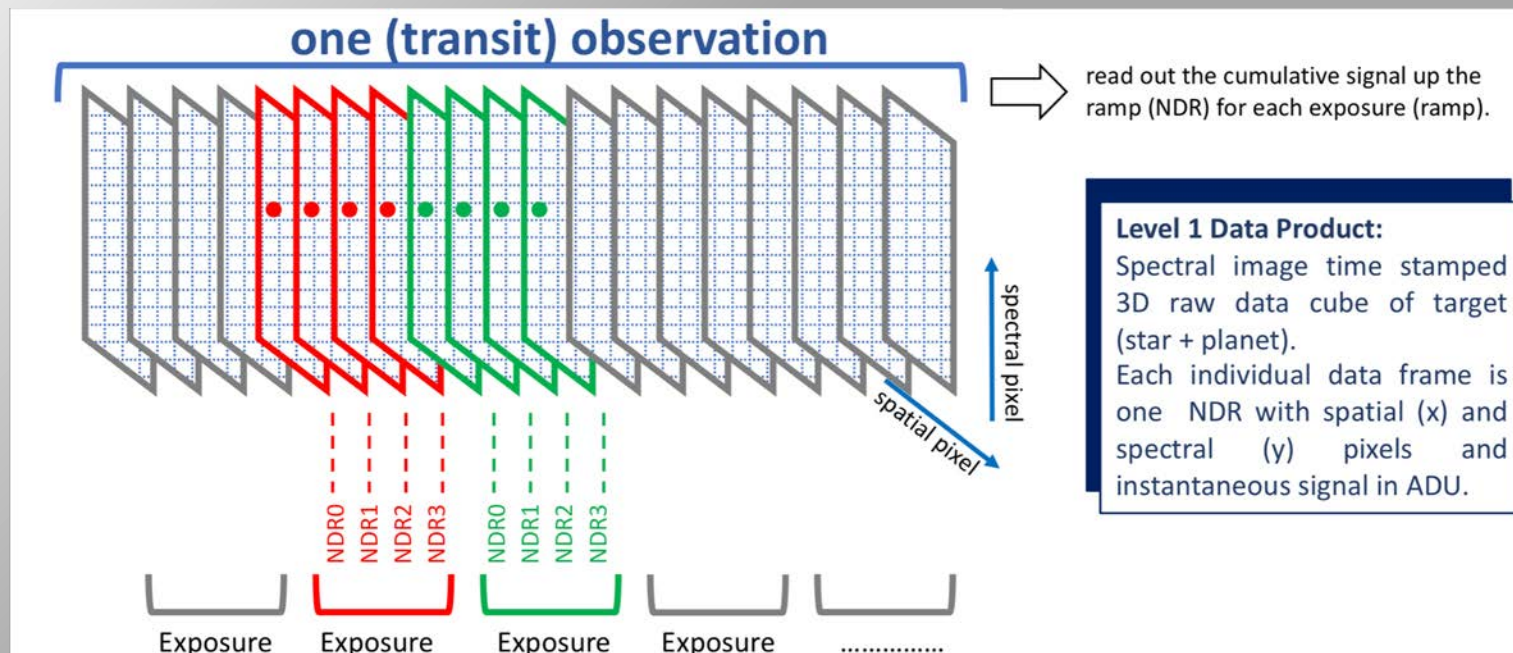
# Data analysis challenge

- Produce end-to-end simulations for a couple of ARIEL observations
- Optimise the data-analysis process to deal with current systematics
- Identify parts of the data-analysis process that may introduce noise and develop a data-analysis strategy
- Evaluate the uncertainty introduced by fitting strategy
- Evaluate the level of noise correlation across wavelength bins

**TIME!**

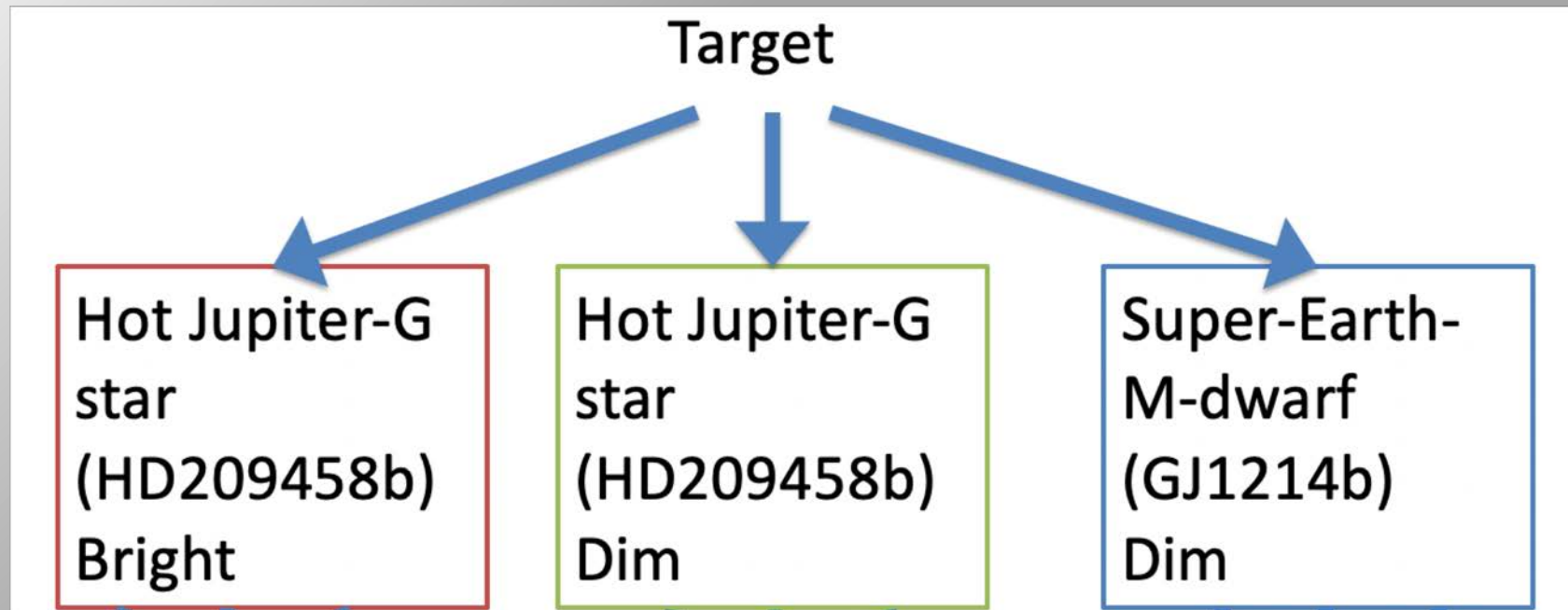
# Data

- Exo-SIM simulations:
  - FITS files, equivalent to level 1 products (no image process)



# Data

- Exo-SIM simulations:
  - FITS files, equivalent to level 1 products (no image process)
  - Planets: GJ-1214 b, HD 209458 b, HD 209458 b faint



# Data


## Exo-SIM simulations:

- FITS files, equivalent to level 1 products (no image process)
- Planets: GJ-1214 b, HD 209458 b, HD 209458 b faint
- Noise:
  - photon noise
  - dark noise
  - read noise
  - emission noise
  - zodi noise
  - jitter noise

<https://ariel-datachallenge.azurewebsites.net/arieldata>

Ariel Data Analysis Challenge

About The Data The Task Upload Current Results Login Sign Up



**ARIEL SPACE MISSION**  
European Space Agency M4 Mission

**Ariel Data Analysis Challenge**

Welcome to the Ariel Data Analysis Challenge. The Ariel Space mission is a European Space Agency mission to be launched in 2028. Ariel will observe the atmospheres of 1000 extrasolar planets - planets around other stars - to determine how they are made, how they evolve and how to put our own Solar System in the galactic context. You can find our press release [here](#).

### Data Analysis

In preparation to the Ariel Red Book, we run this data analysis challenge to compare the efficiency of different data analysis practices on simulated Ariel data. The aim is to optimise the steps of the Ariel data analysis pipeline towards minimising the uncertainty of the final data products, as well as the total execution time. [Find out more!](#)

# Challenge

- Reduction:
  - Background subtraction
  - Jitter correction
  - Photometry
  - Uncertainties propagation
- Fitting:
  - Exposure-integrated model
  - Limb-darkening handling
  - Orbital parameters
  - Planetary spectrum extraction

# Challenge

- Reduction:

- Background subtraction
- Jitter correction
- Photometry
- Uncertainties propagation

- Fitting:

- Exposure-integrated model
- Limb-darkening handling
- Orbital parameters
- Planetary spectrum extraction

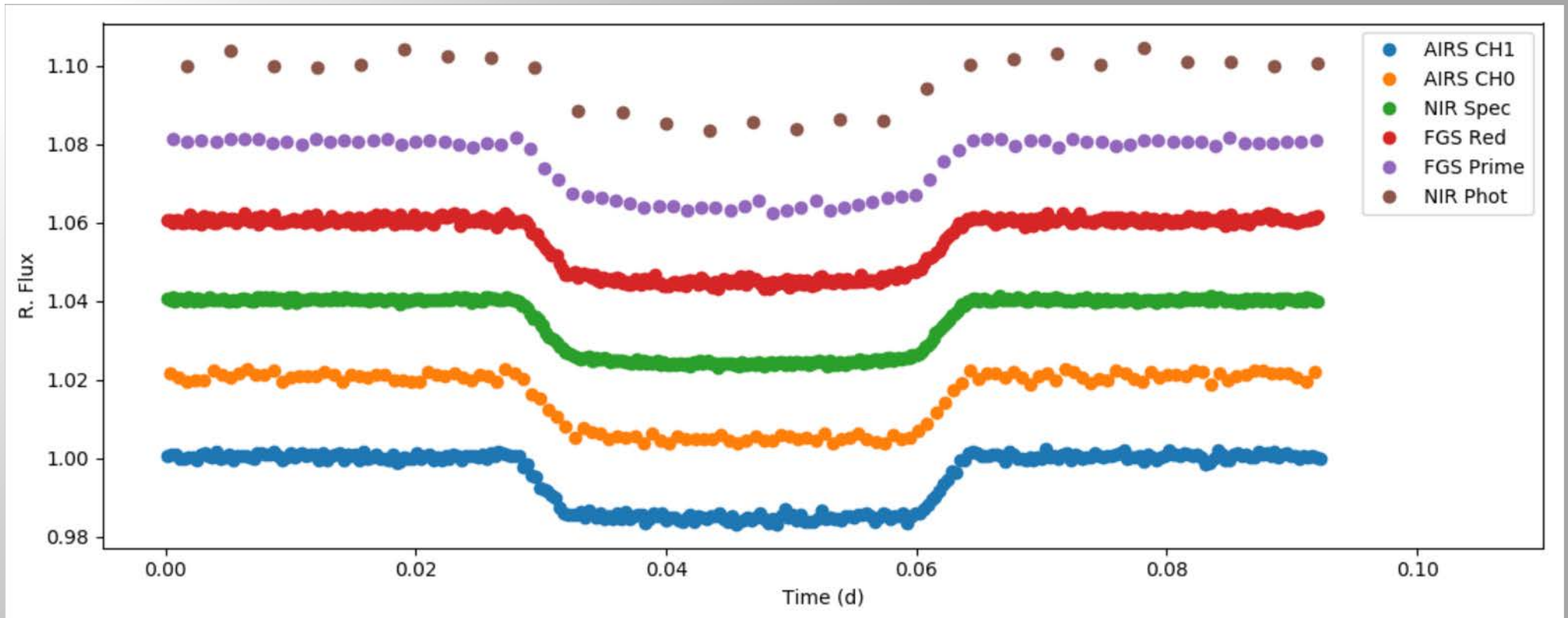
A black oval containing the text "ADaRP" in red, bold, serif font.

**ADaRP**

A black oval containing the text "Iraclis" in red, bold, serif font.

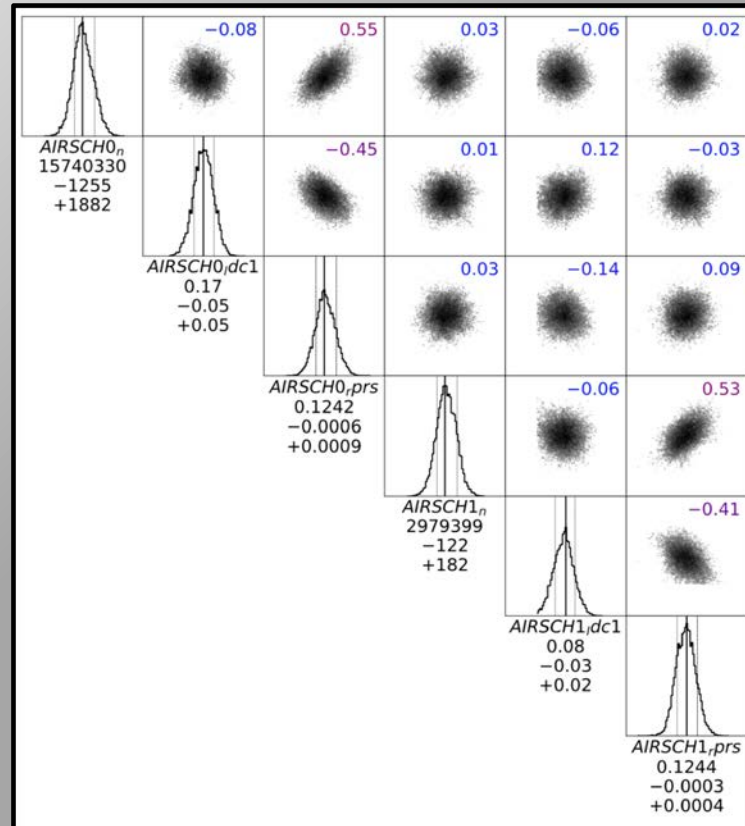
**Iraclis**

# Analysis

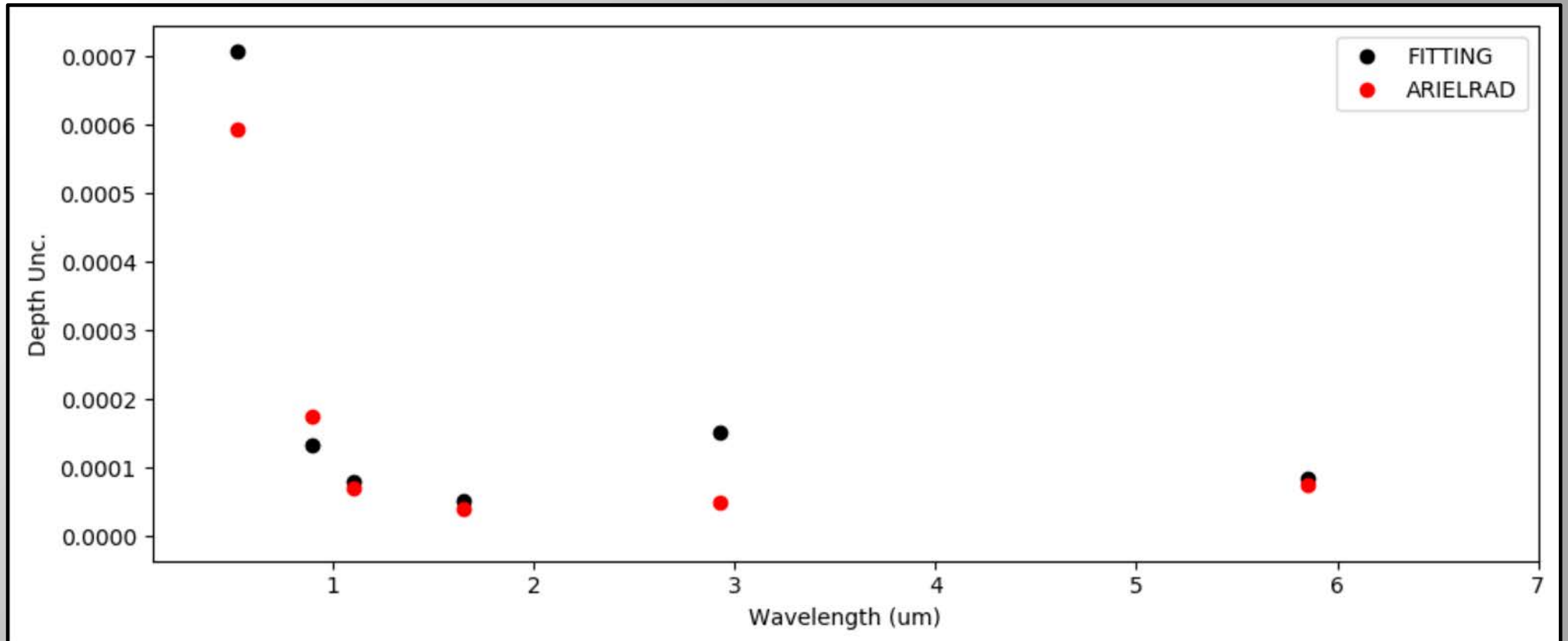




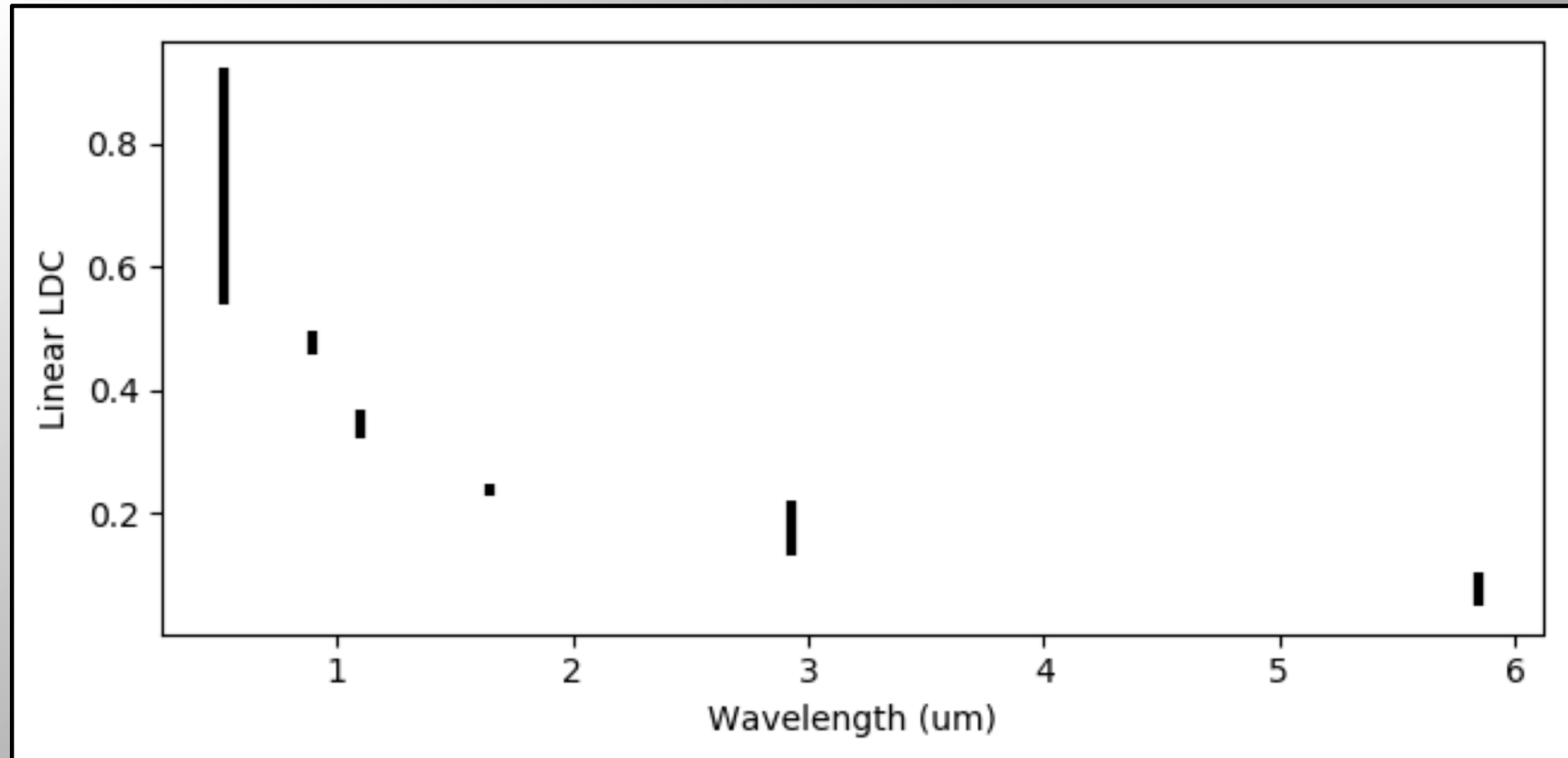
# Analysis



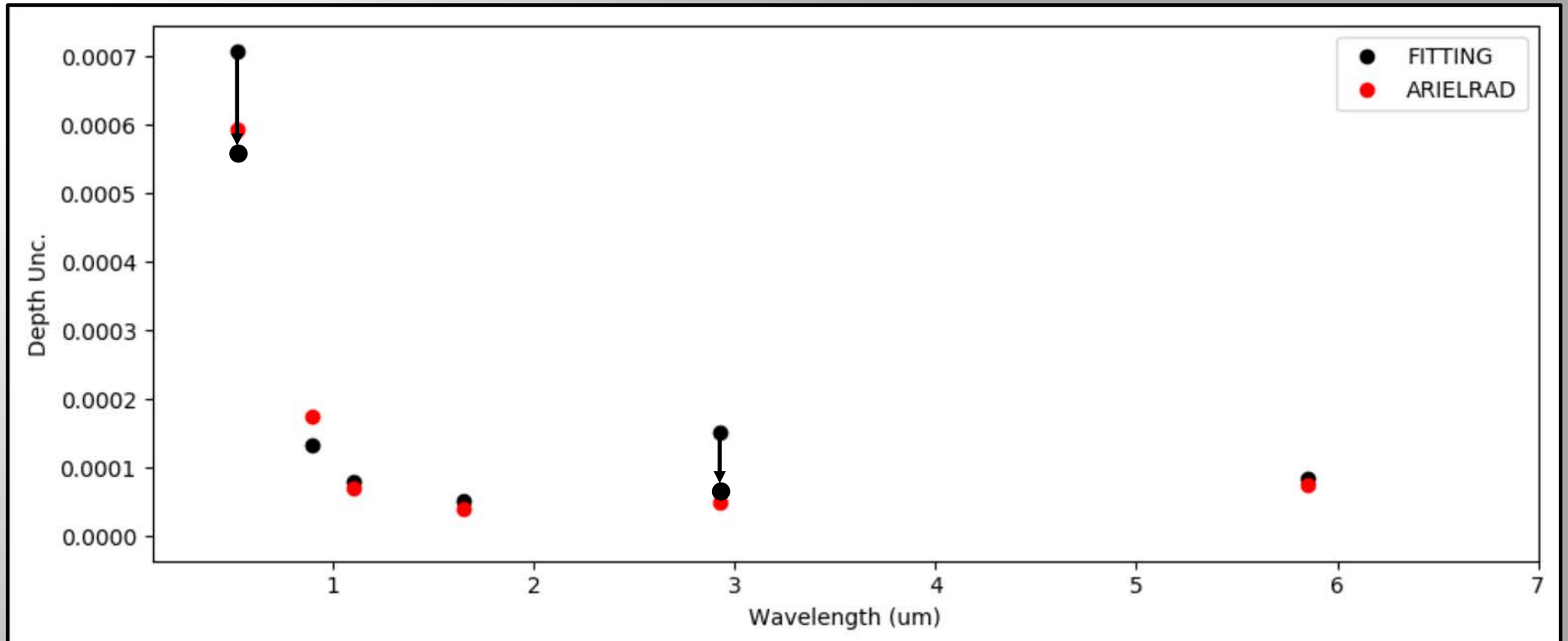
# Analysis



# Analysis




# Analysis



<https://ariel-datachallenge.azurewebsites.net/arieldata>

Ariel Data Analysis Challenge

About The Data The Task Upload Current Results Login Sign Up



**ARIEL SPACE MISSION**  
European Space Agency M4 Mission

**Ariel Data Analysis Challenge**

Welcome to the Ariel Data Analysis Challenge. The Ariel Space mission is a European Space Agency mission to be launched in 2028. Ariel will observe the atmospheres of 1000 extrasolar planets - planets around other stars - to determine how they are made, how they evolve and how to put our own Solar System in the galactic context. You can find our press release [here](#).

### Data Analysis

In preparation to the Ariel Red Book, we run this data analysis challenge to compare the efficiency of different data analysis practices on simulated Ariel data. The aim is to optimise the steps of the Ariel data analysis pipeline towards minimising the uncertainty of the final data products, as well as the total execution time. [Find out more!](#)