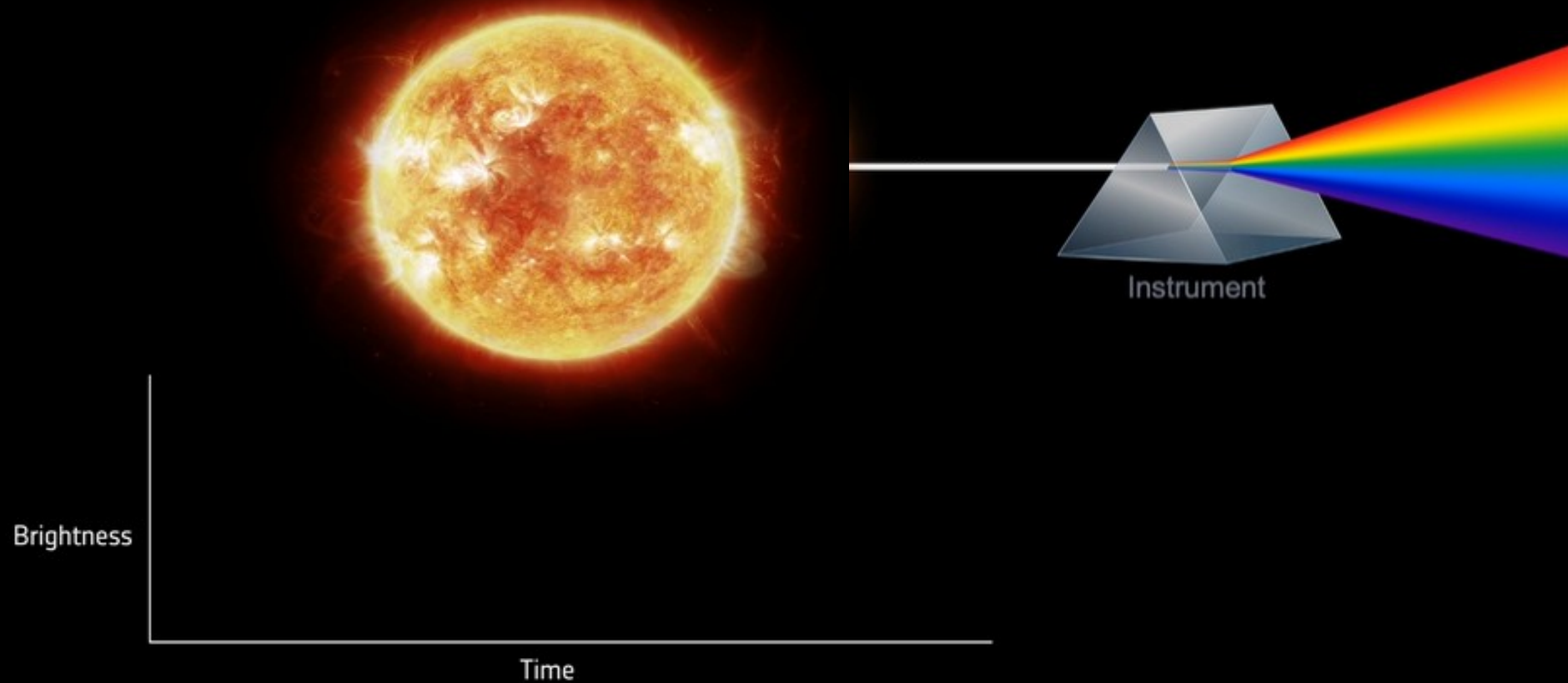




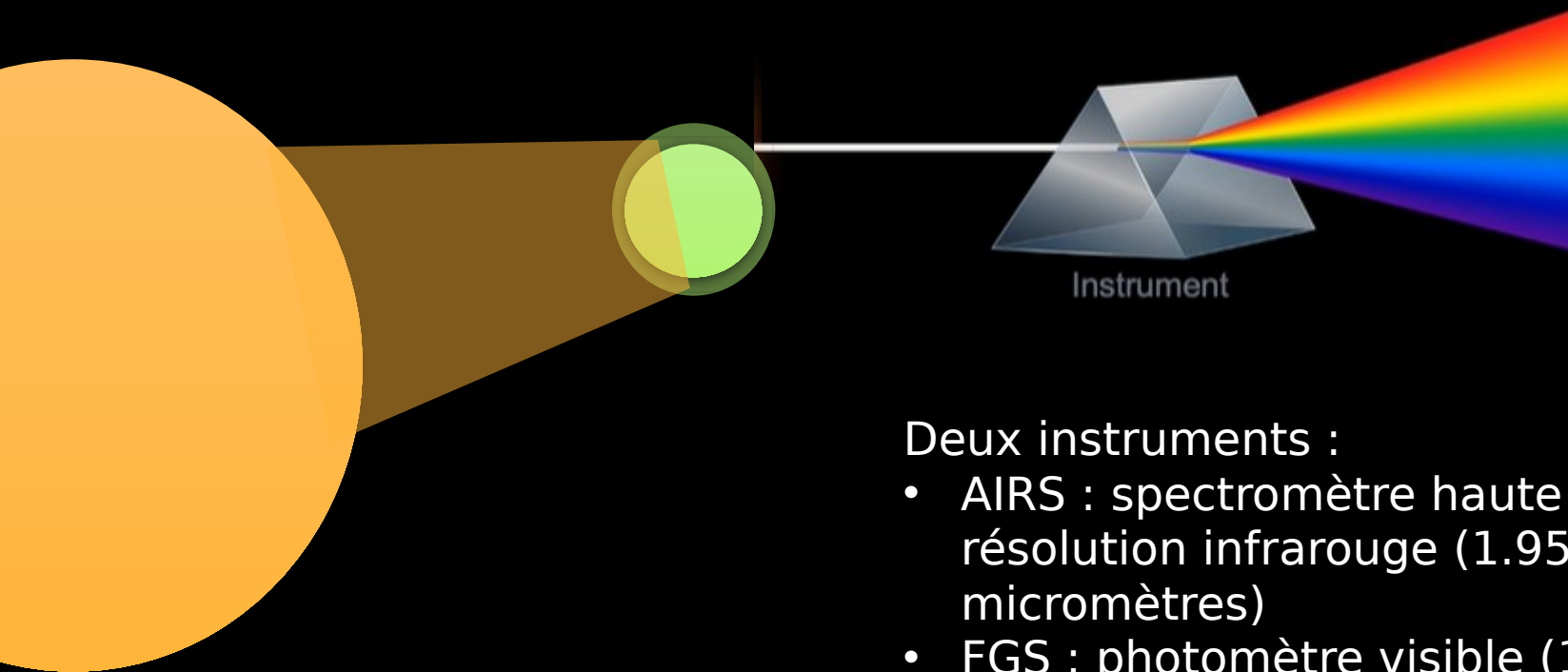
ORPHÉE FAUCOZ (CNES)

15/10/2025

# | CONTEXTE

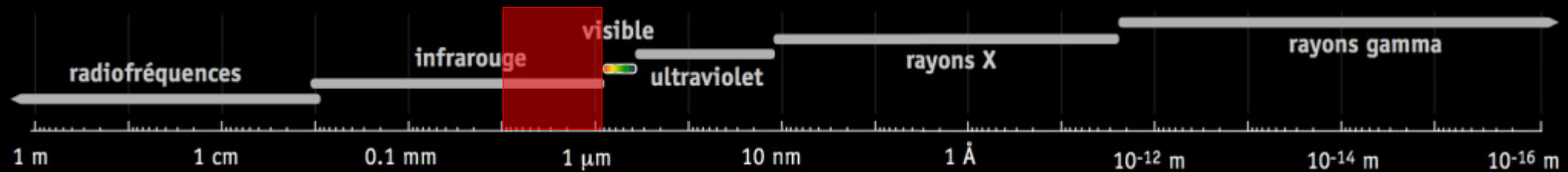


# | CONTEXTE

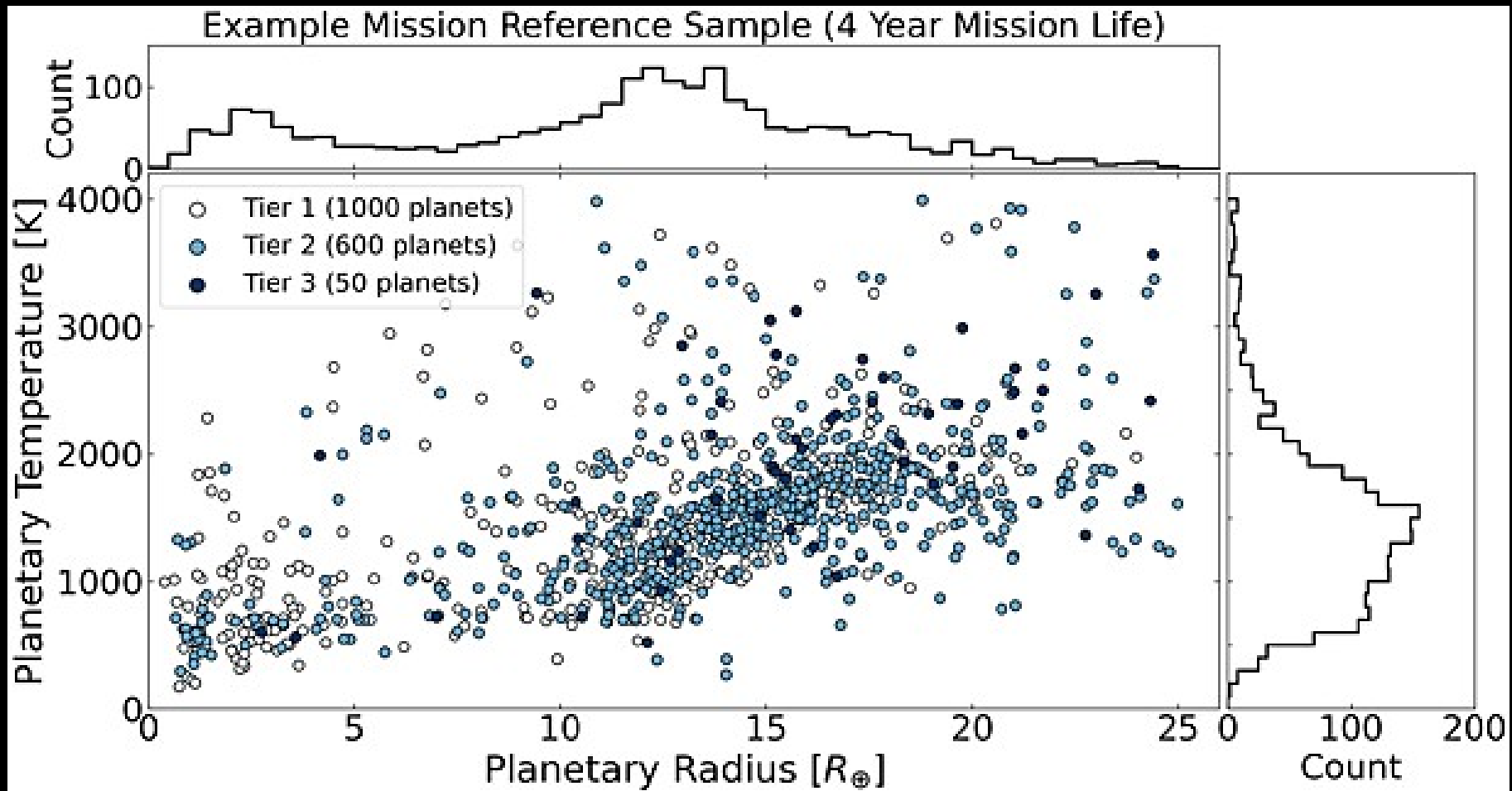


Deux instruments :

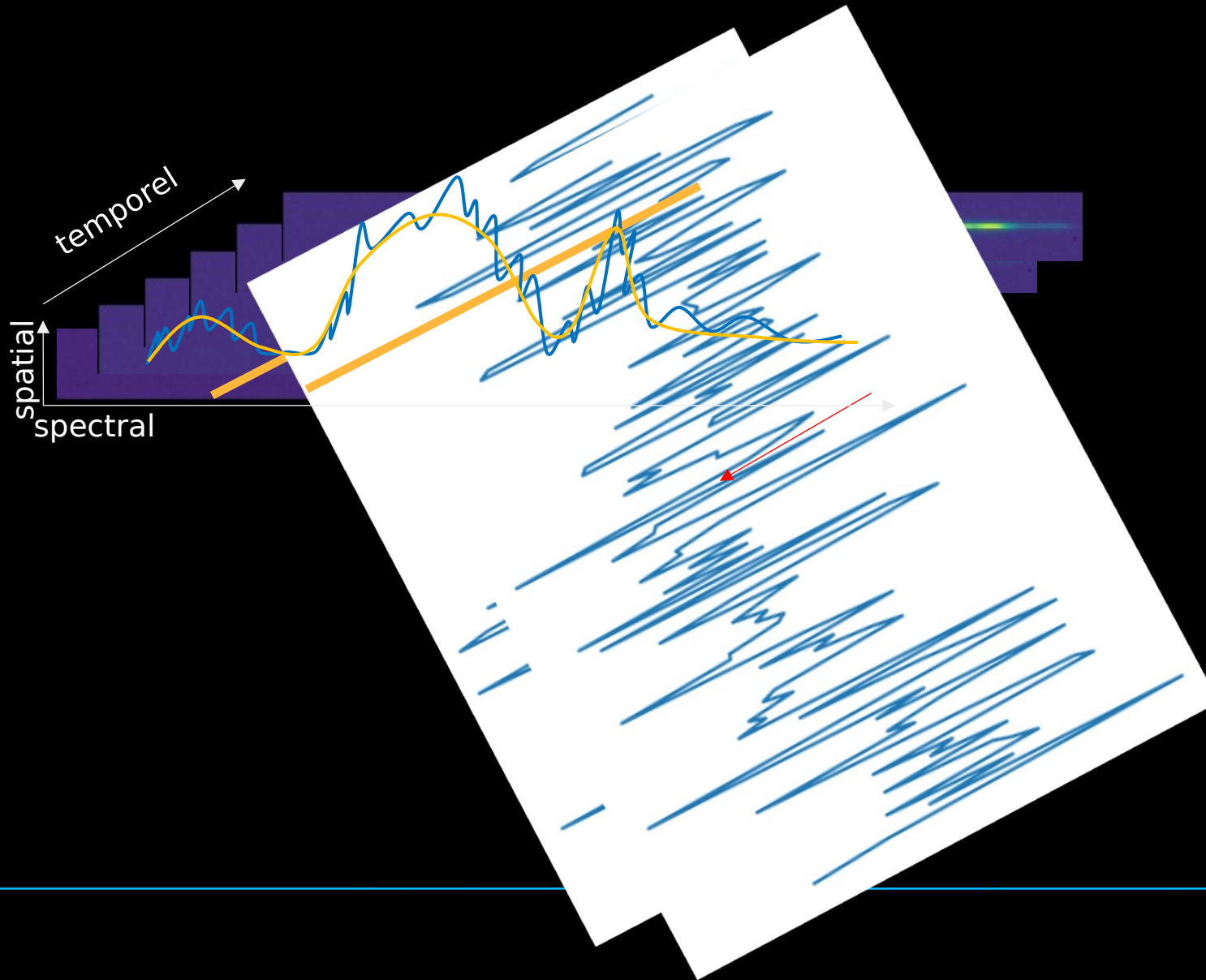
- AIRS : spectromètre haute résolution infrarouge (1.95 – 7.8 micromètres)
- FGS : photomètre visible (1.1 à 1.95 micromètres)



# | CONTEXTE



# | OBSERVATION VERS SPECTRE

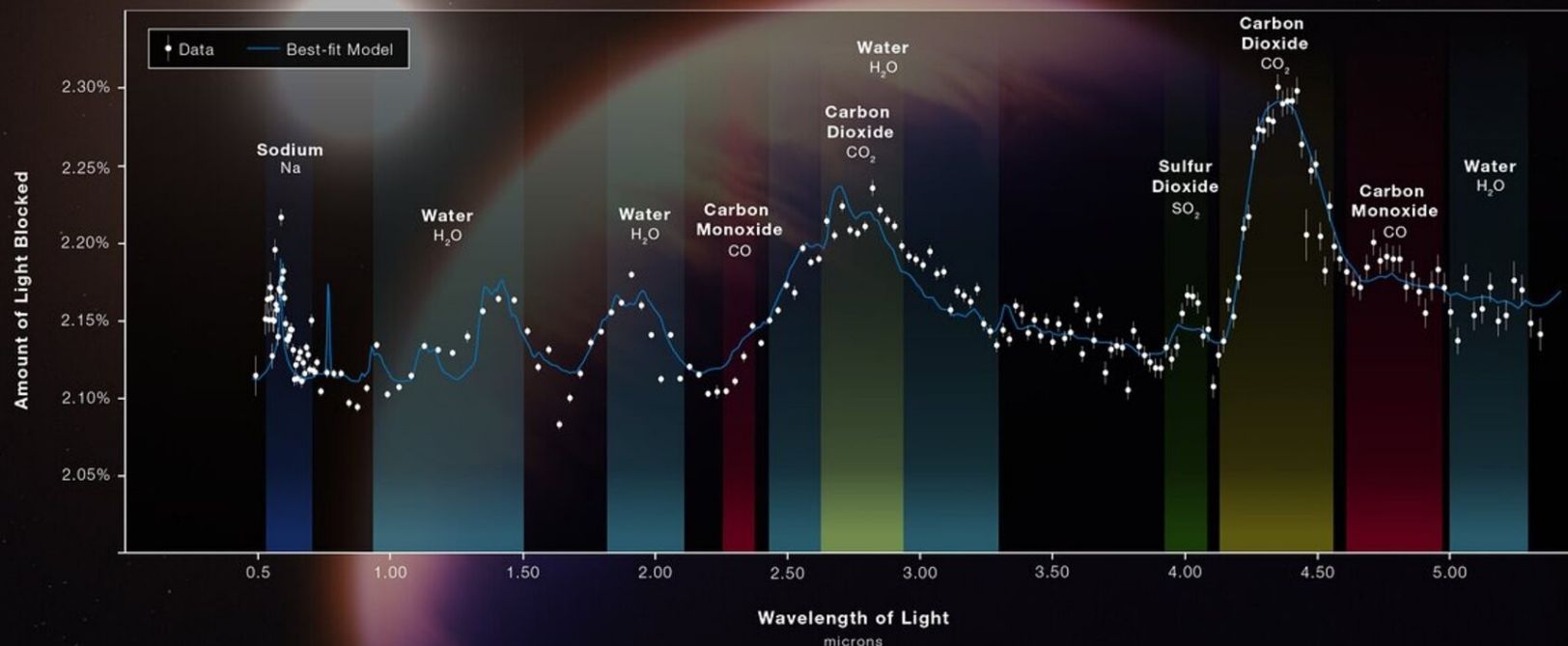


# | CONTEXTE

	2-5 $\mu\text{m}$	5-8 $\mu\text{m}$
H <sub>2</sub> O*	2.69	6.2
CO <sub>2</sub> *	2.03, 4.25	-
C <sub>2</sub> H <sub>2</sub>	3.0	7.53
HCN	3.0	-
O <sub>3</sub>	4.7	-
HDO	2.7, 3.67	7.13
CO*	2.35, 4.7	-
NH <sub>3</sub>	2, 2.25, 2.9,	6.1
PH <sub>3</sub>	4.3	-
CH <sub>4</sub> *	2.2, 2.31, 2.37, 3.3	6.5, 7.7
CH <sub>3</sub> D	3.34, 4.5	6.8, 7.7
C <sub>2</sub> H <sub>4</sub>	3.22, 3.34	6.9
H <sub>2</sub> S	2.5, 3.8, ...	7
SO <sub>2</sub>	4	7.3
N <sub>2</sub> O	2.8, 3.9, 4.5	7.7
NO <sub>2</sub>	3.4	6.2, 7.7
H <sub>3</sub> <sup>+</sup>	2.0, 3-4.5	-
TiO	2-3.5	-
VO	2-2.5	-
FeH	2	-
C <sub>2</sub> H <sub>6</sub>	3.35	6.79

## HOT GAS GIANT EXOPLANET WASP-39 b ATMOSPHERE COMPOSITION

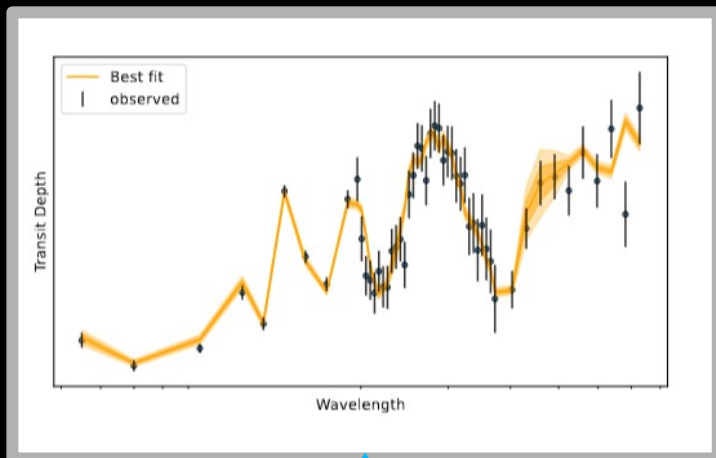
NIRSpec PRISM



**WEBB**  
SPACE TELESCOPE

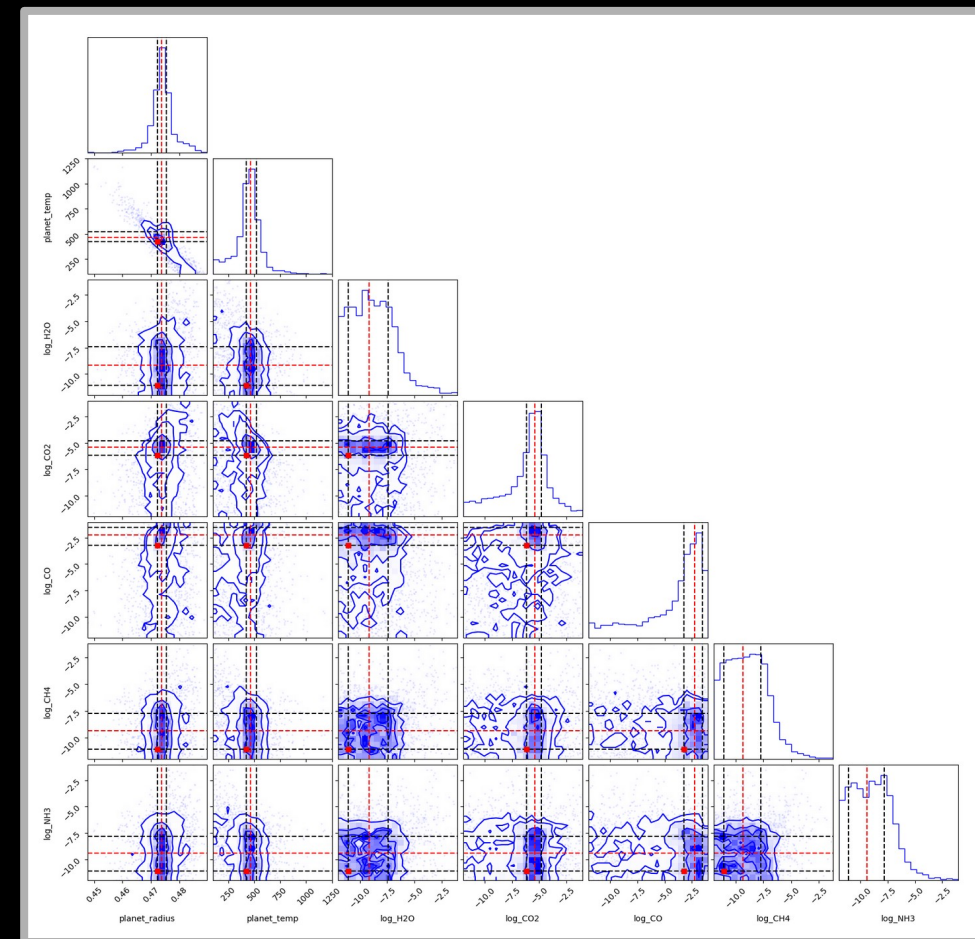
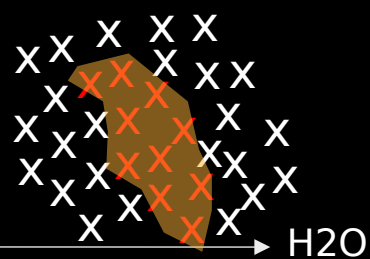


# | LE RETRIEVAL



MCMC / Nested Sampling

Température



# OBJECTIFS

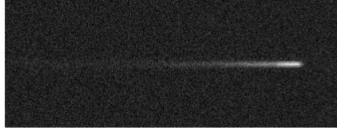
## ADaRP Pipeline Processing Steps

## Science Data Products

### Level 0

- Read Level 0 Product
- Decompression
- Reformat
- Meta data initialisation
- Image array initialisation
- Write Level 1 Product

### LEVEL 0 Product: Raw compressed data files



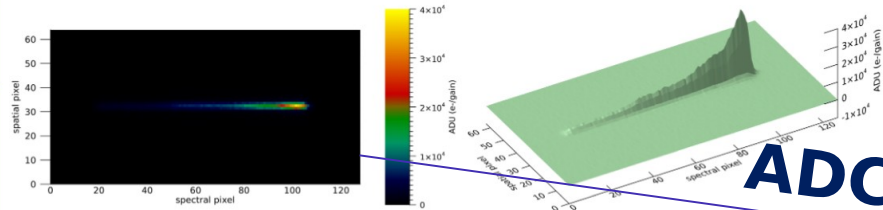
Raw data derived from TLM, delivered from MOC to SOC

### Level 1

- Read Level 1 Product
- DQ Array initialisation
- ADU to e- unit conversion
- Flag saturated pixels
- Subtract zeroth read
- Non-linearity correction
- Pixel cross-talk correction
- Dark current subtraction
- Flat fielding
- Bias drift correction
- Gain drift correction
- Persistence correction
- Background subtraction
- Error array initialisation
- e- to e-/s conversion
- Image statistics (deglitching)
- Bad pixel correction
- Pointing jitter correction
- Write Level 1.5 Product

### LEVEL 1 Product: Uncalibrated science frames of photometric or spectral images

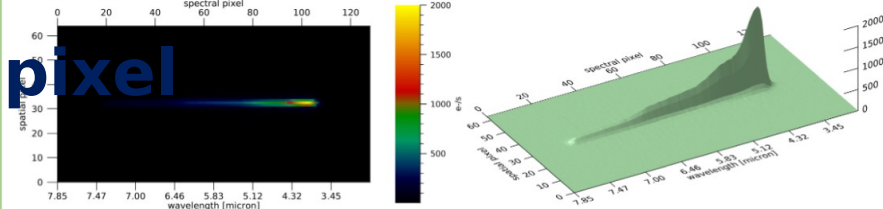
Raw, time stamped (spectral) image cube with each slice corresponding to a science frame with axes; spatial pixel number, spectral pixel number for spectral channels and spatial pixel number for photometric channels and integrated signal in ADU.



Example temporal slice (i.e. one Science Frame) of an ARIEL Data Level 1 data product, for the spectroscopy channel AIRS CH1

### LEVEL 1.5 Product: Calibrated exposures of target observations corrected for instrumental artefacts

Calibrated, time stamped (spectral) image cube with each slice, an array of fitted ramps (slopes) with axes; spatial pixel number, spectral pixel in microns for the spectral channels spatial pixel number, spatial pixel number for photometric channels (or spectral pixel in microns) and signal (slope) in e-/s.



Example of temporal slice (corresponding to a single exposure) of an ARIEL Level 1.5 data Product, for the spectroscopy channel AIRS CH1  
Note that wavelength in Level 1.5 is not corrected for optical distortion.

### Level 1.5

- Read Level 1.5 Product
- Apply spectroscopic mask s
- Extract 1-D spectrum s
- Spectral binning s
- Apply photometric mask p
- e-/s to flux conversion
- Write Level 2 Product

### Level 2

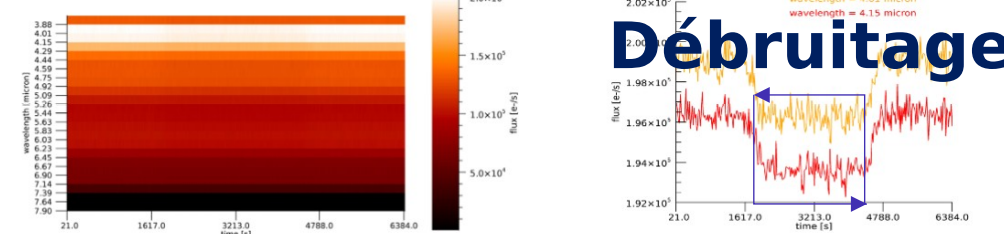
- Read Level 2 Product
- Residual detrending
- Stellar activity correction 1
- Fit model light curves
- Stellar activity correction 2
- Write Level 3 Product

### Level 3

Example of temporal slice (corresponding to a single exposure) of an ARIEL Level 1.5 data Product, for the spectroscopy channel AIRS CH1  
Note that wavelength in Level 1.5 is not corrected for optical distortion.

### LEVEL 2 Product: Spectrally resolved light-curves of the target (star + planet(s))

Wavelength binned 2D set of light curves with axes; time, wavelength (microns), signal (e-/s) for the spectral channels and time, photometric band, signal (e-/s) for photometric channels.



Example ARIEL Data Level 2 for spectroscopy channel AIRS CH1 showing overall Level 2 data array and slices through wavelength for light curve.  
Note that time axes may be in HJD

### LEVEL 3 Product: Exoplanet broad-band spectra

Individual planet(s) spectrum (e.g.: ppm vs wavelength) for each observation with time (s)

Legacy Co-added planet(s) spectrum (e.g.: ppm vs wavelength) for all observations

Stellar properties

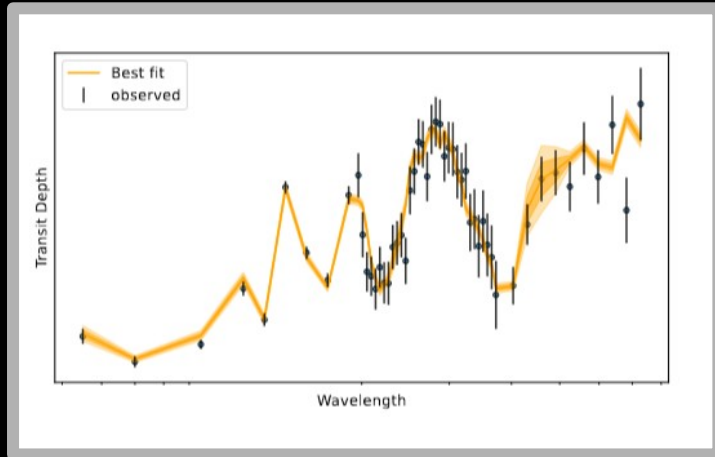


Chimie  
Abondances chimiques, profil thermique, nuage

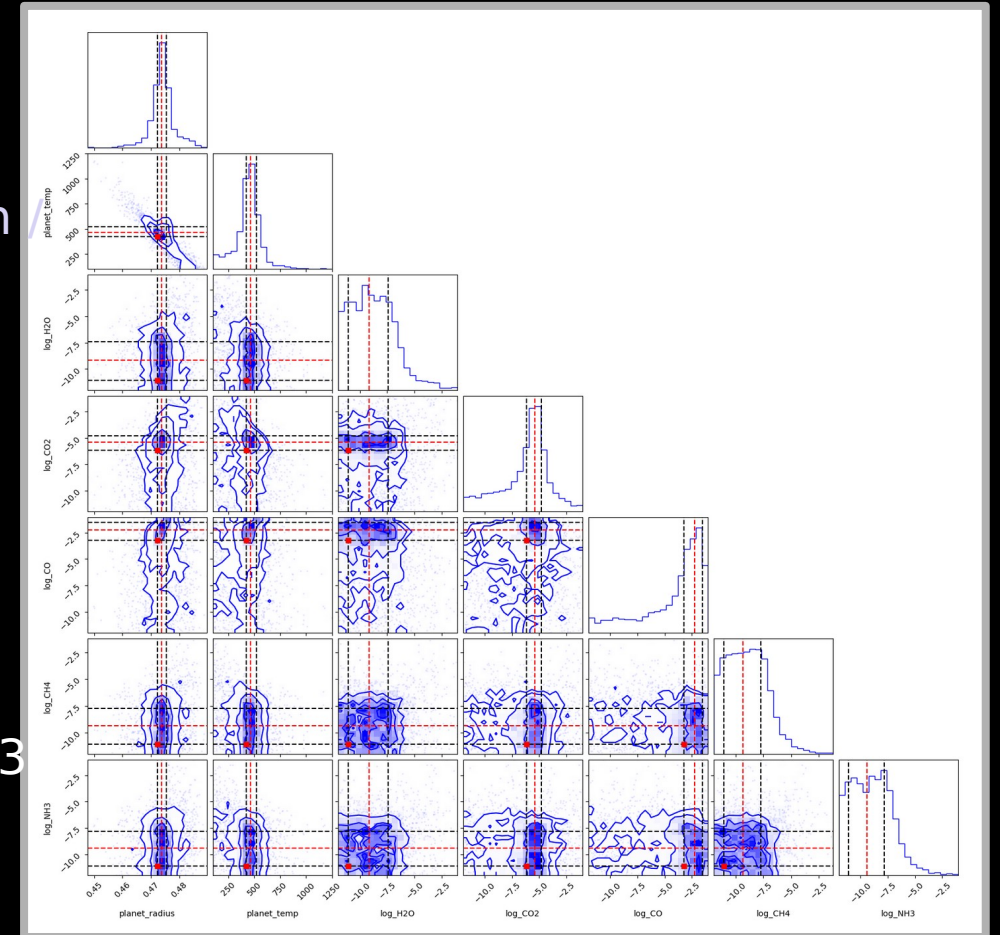
ADC 2022/2023



# | ADC 2022/2023



Réseau de convolution  
NF

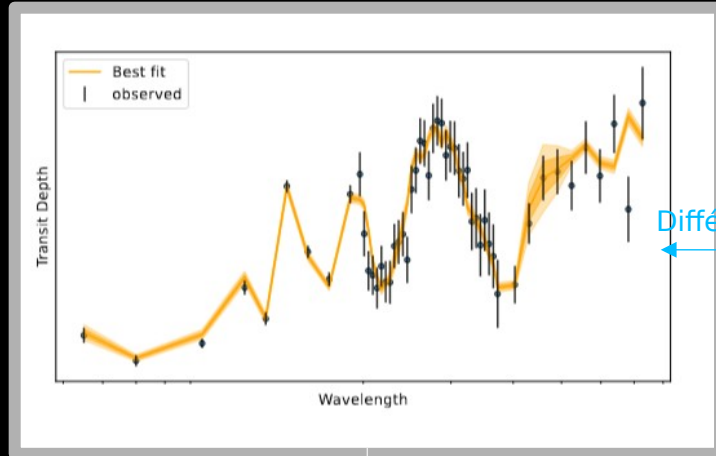


~26 000 spectres simulés et leurs abondances (Taurex 3)  
**+80 000 spectres sans abondance**

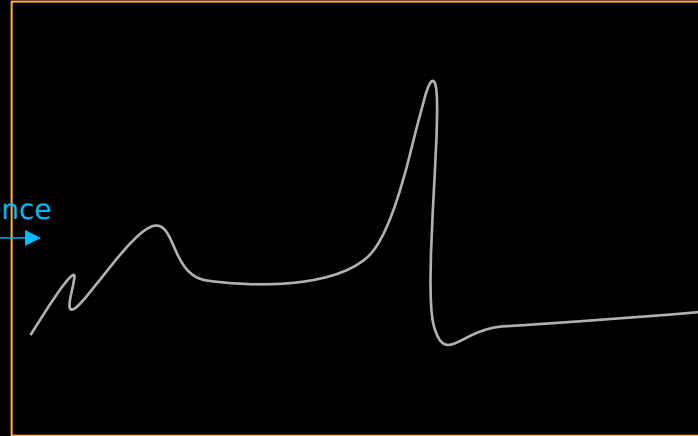
448 participants

Métrique : distance de Wasserstein-2 (2022), Kolmogorov-Smirnov-2 (2023)

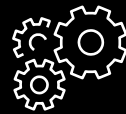
# | POUR ALLER AU-DELÀ ?



Différence

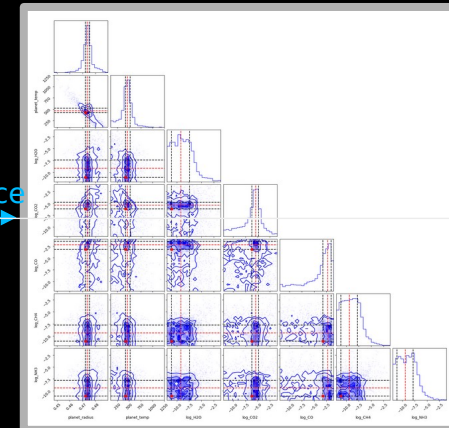


Taurex 3



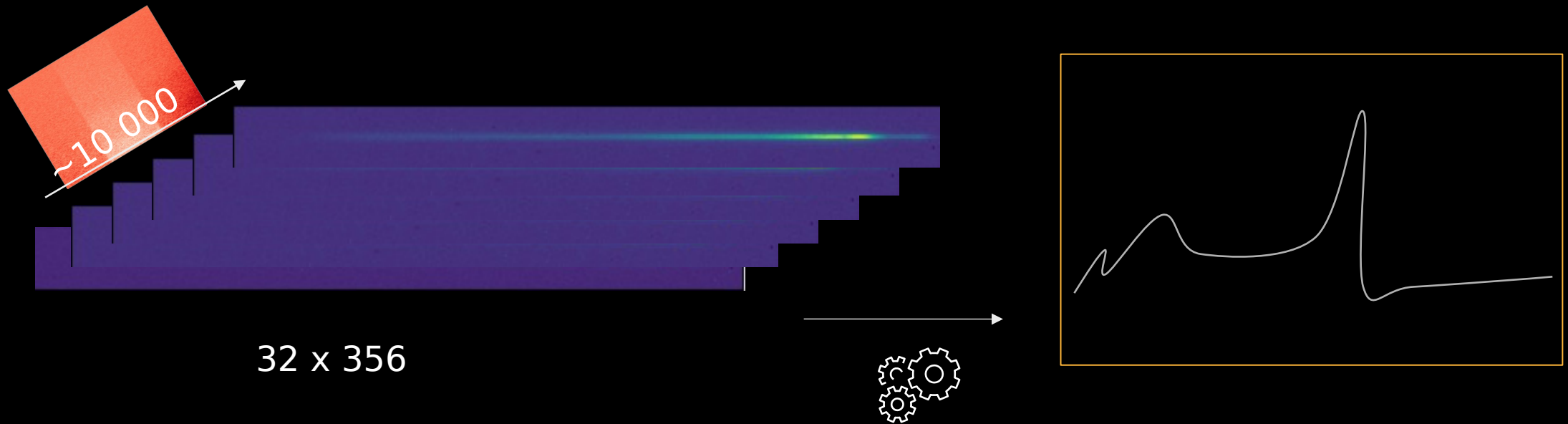
Sorties

Différence



Paramètres connus

# | ADC 2024



644 observations de transit  
simulées par ExoSim 2 sur FGS1  
et AIRS-CH0

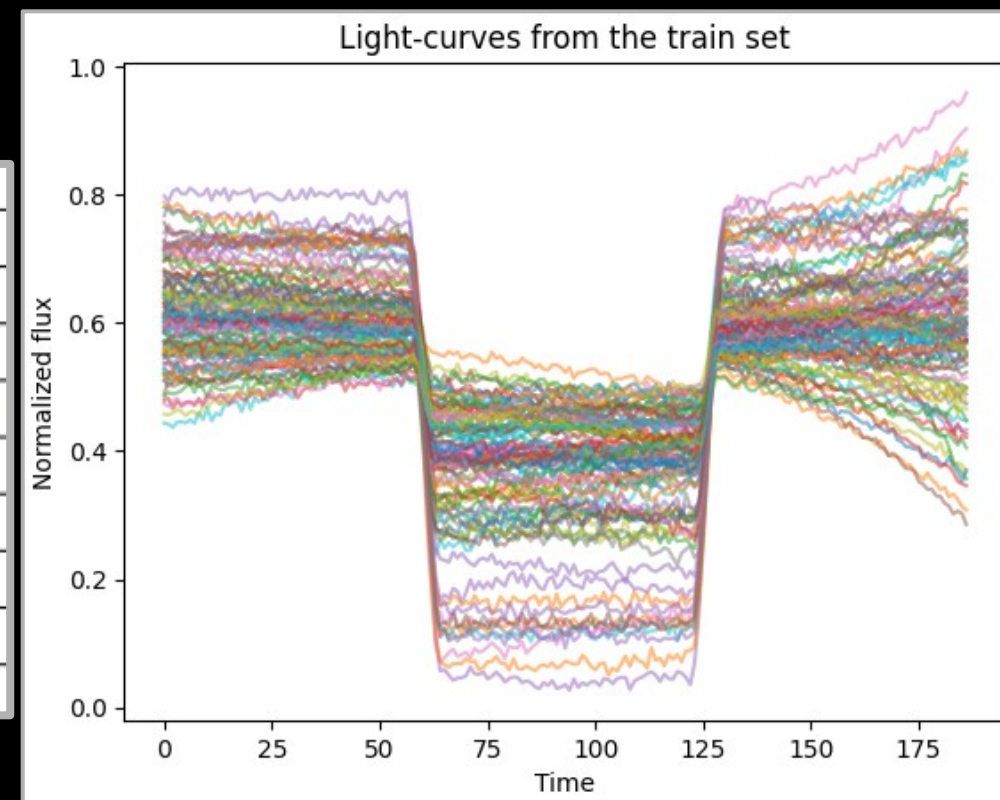
1150 participants (1080  
soumissions) !

644 spectres

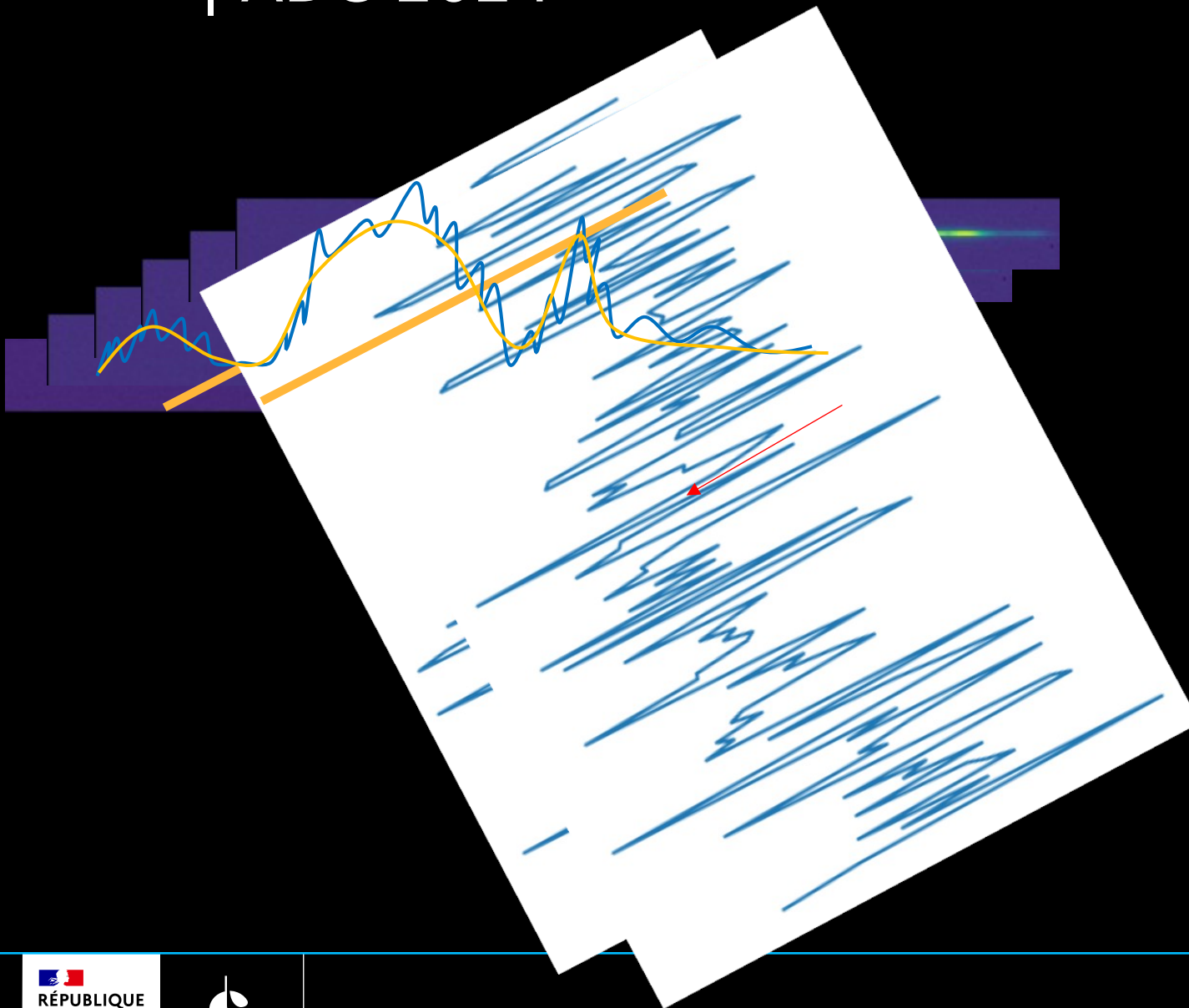
# | ADC 2024

## Diversité du jeu de données

	Training	Training-Ext	Test case 0	Test case 1
$R_p$ range ( $R_{Jup}$ )	0.5 - 1.2	0.5 - 1.2	0.5 - 1.2	0.5 - 1.2
Host Stars	KELT-11, HD 17194	KELT-11, HD 17194	KELT 11, HD 17194	KELT-11, HD 17194
Trace gases	H <sub>2</sub> O, CH <sub>4</sub> , CO <sub>2</sub> , Cloud	H <sub>2</sub> O, CH <sub>4</sub> , CO <sub>2</sub> , Cloud	NH <sub>3</sub> , C <sub>2</sub> H <sub>2</sub> , SO <sub>2</sub> , SiO	NH <sub>3</sub> , C <sub>2</sub> H <sub>2</sub> , SO <sub>2</sub> , SiO
Detector type	Model A	Model A	Model A	Model B
Semi-Major Axis	Fixed	$\mathcal{U}(8.81, 11)$	Fixed	Fixed
Inclination	Fixed	$\mathcal{U}(86.71, 88)$	Fixed	Fixed
Mid Transit Time	Fixed	$\mathcal{U}(t_0-0.1, t_0+0.1)$	Fixed	Fixed



# | ADC 2024



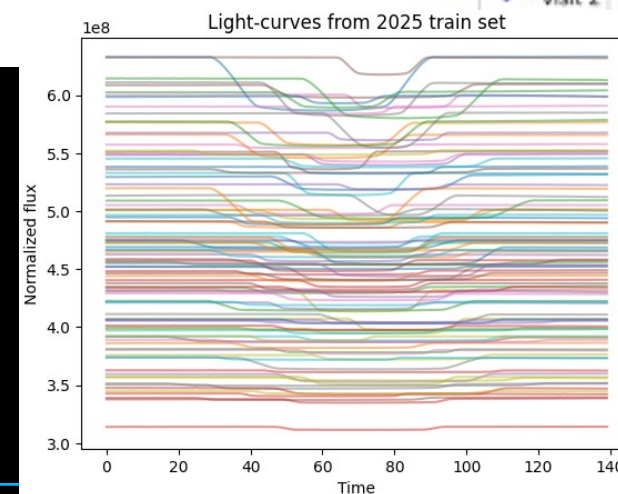
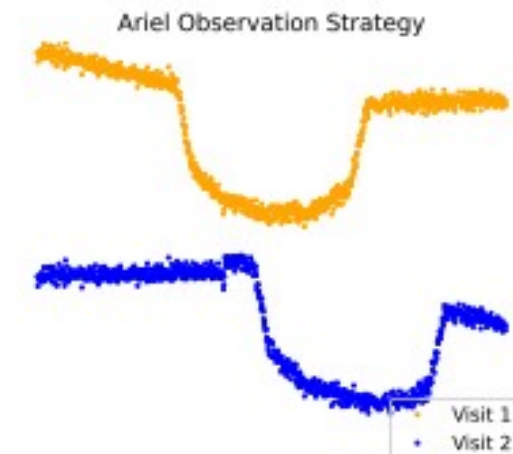
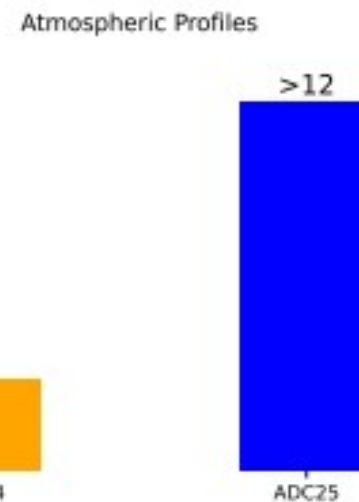
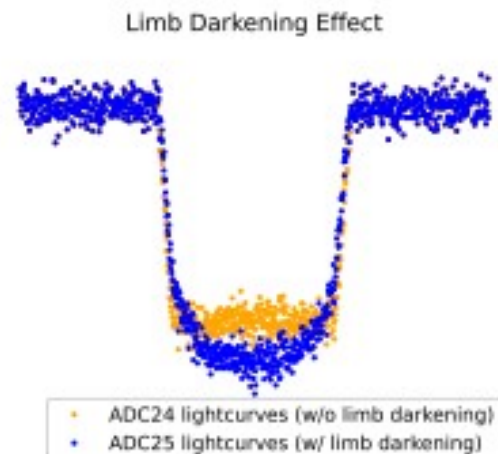
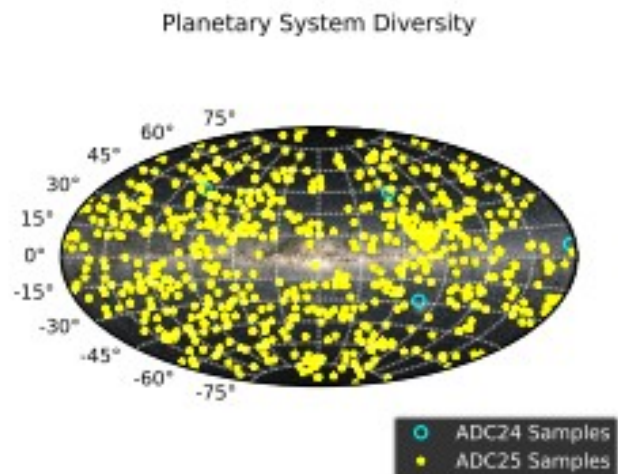
Estimation de la profondeur de transit par wl

**Comment ?**

- Estimation de début et de fin de transit par **interpolation**
- Estimation de profondeur par **interpolation**
- Amélioration par méthode de **machine learning** (Autoencodeur, GPR, NMF)
- Estimation du rayon par transit moyen sur la courbe de lumière



# | ADC2025



1000 observations ?

IAP/CNES dans la Baseline  
Team

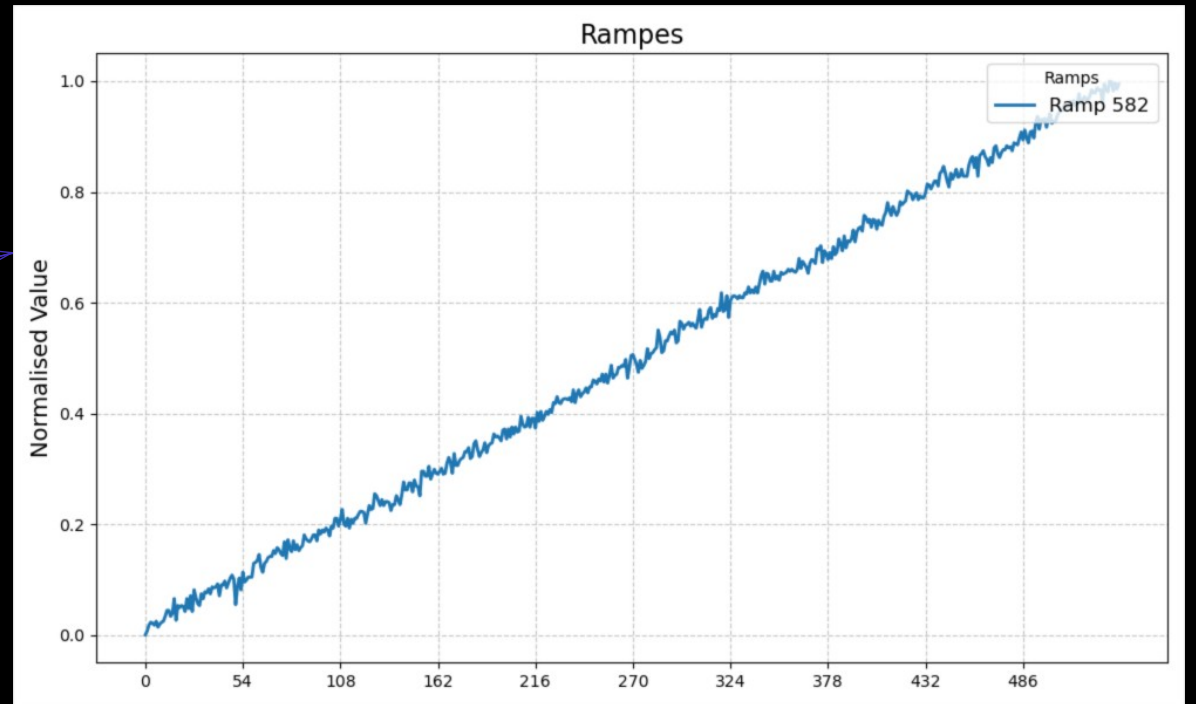
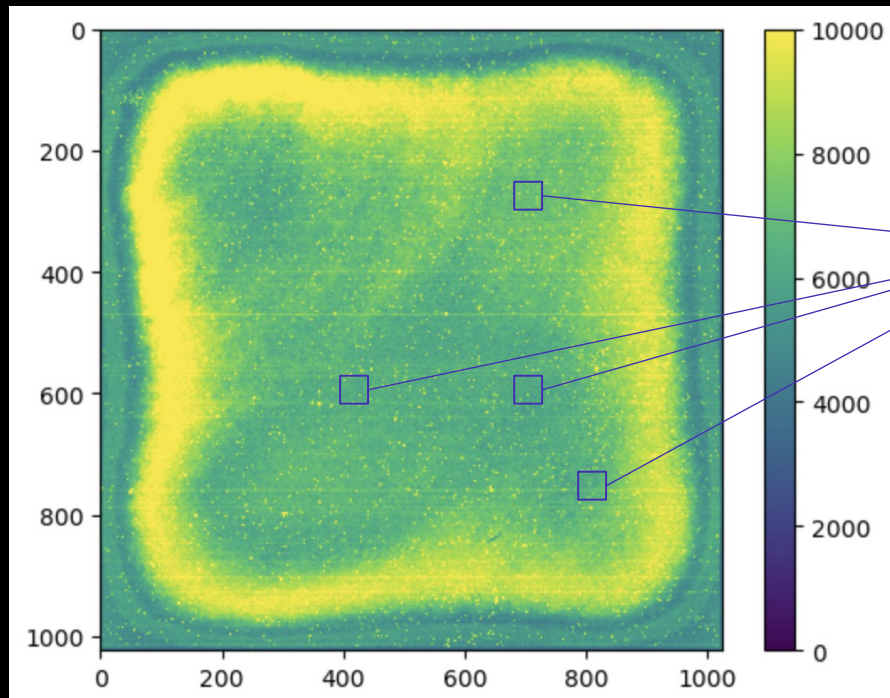
# | ADC 2025

Estimation de la profondeur de transit par wl

**Comment ?**

- Estimation de début et de fin de transit par **interpolation**
- Estimation de profondeur par **interpolation**
- Amélioration par méthode de **machine learning (GPR)**
- Estimation du rayon par transit moyen sur la courbe de lumière

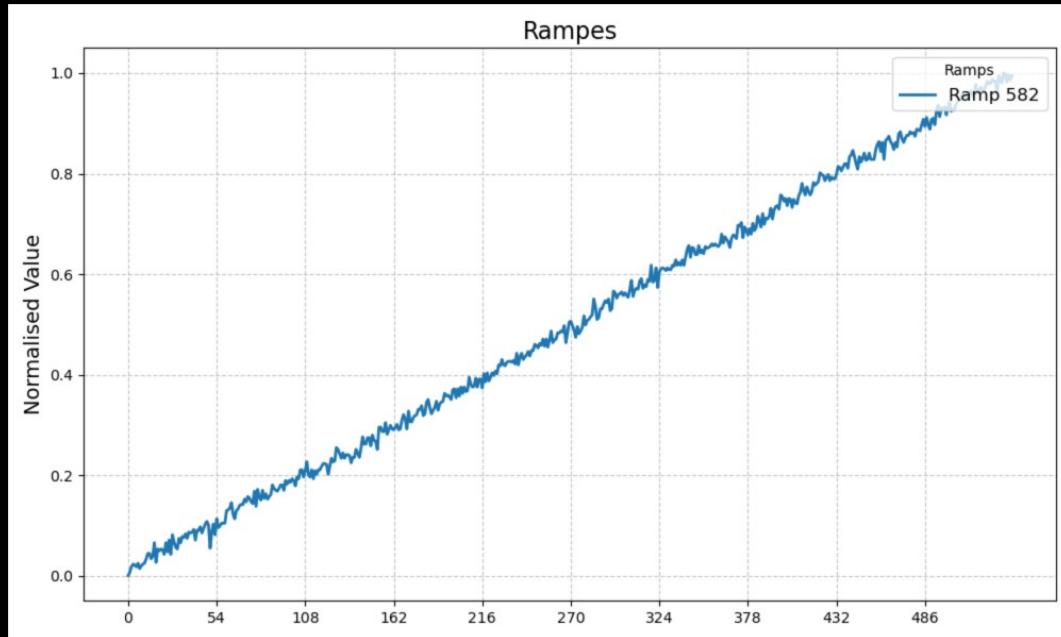
# | BAD PIXELS



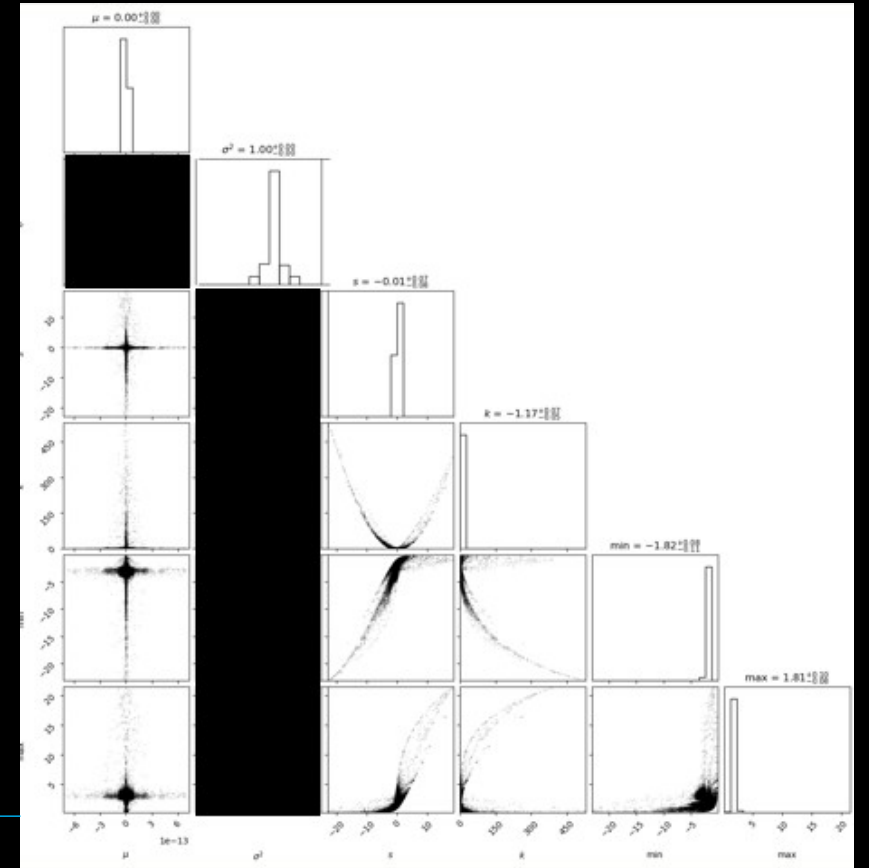
40 fichiers de dark mesurés au CEA

# | BAD PIXELS

Non supervisé



N profiles ←

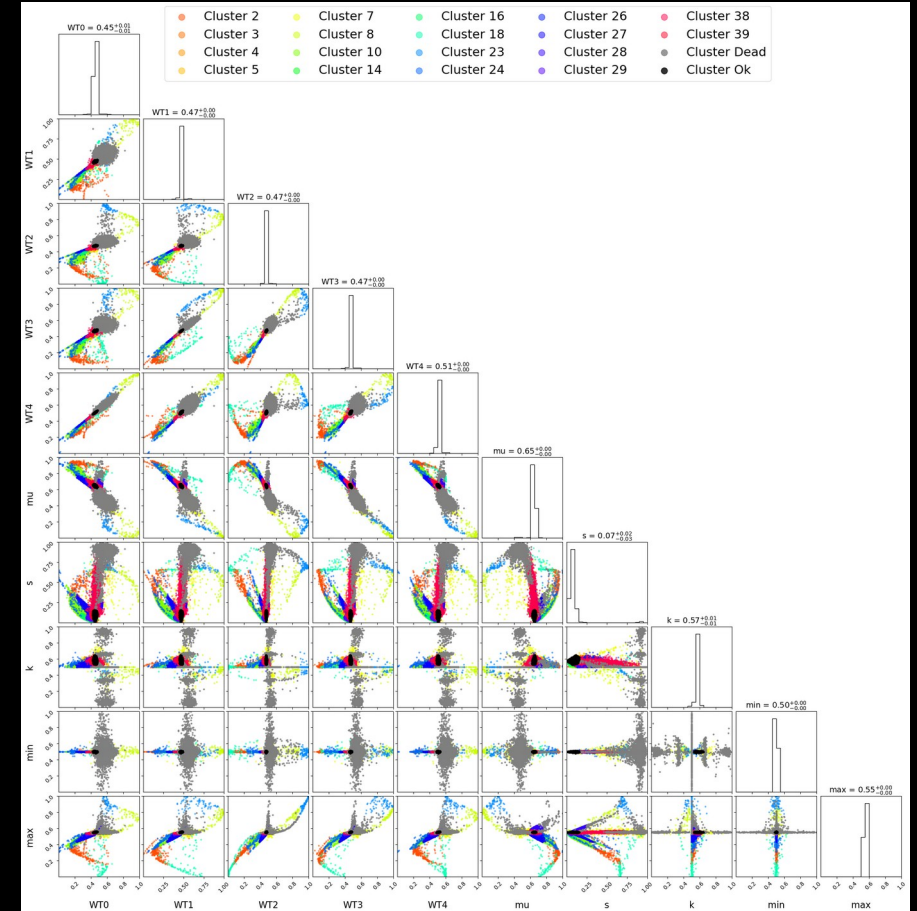
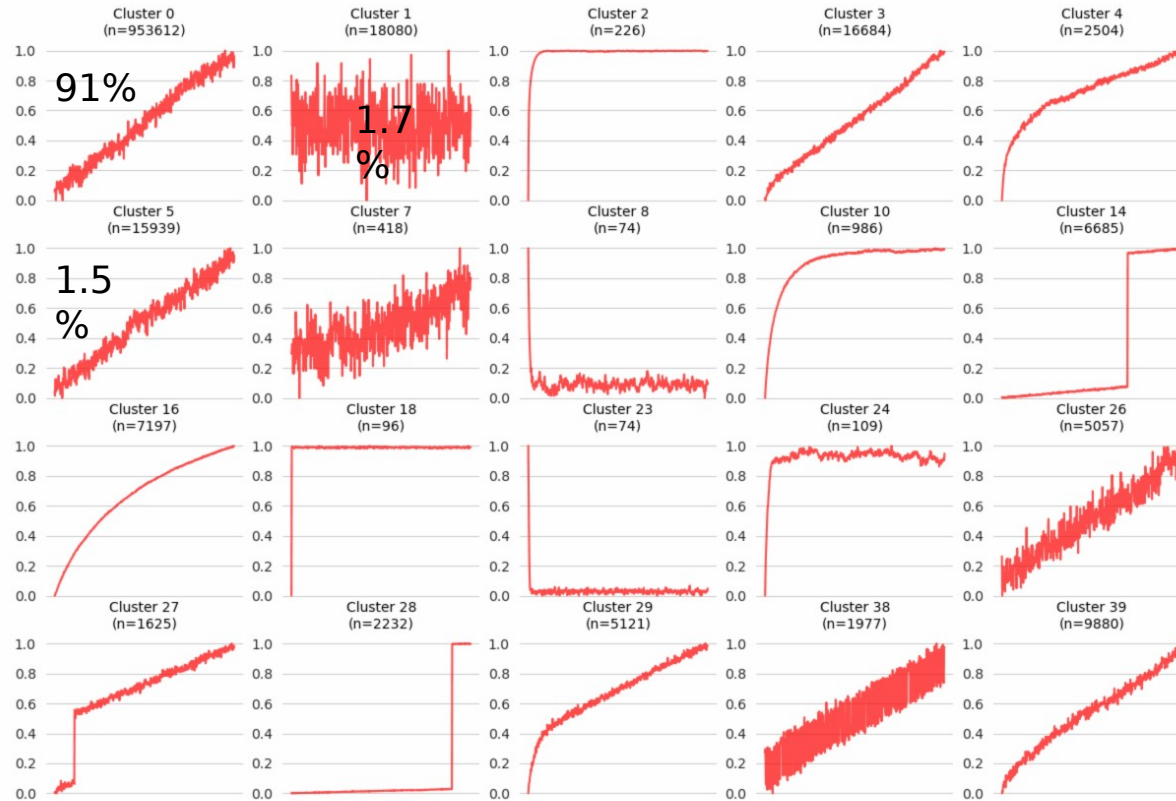




# | BAD PIXELS

Non supervisé

Clusters Visualization

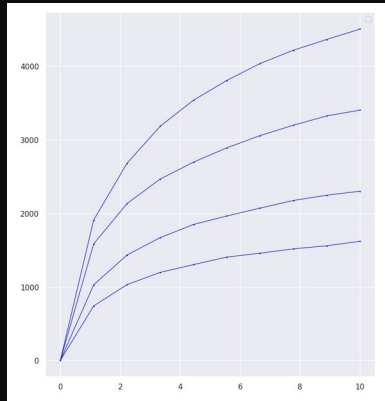




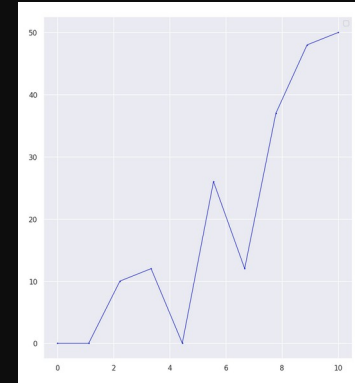
# | BAD PIXELS SUR JWST

PIXELS

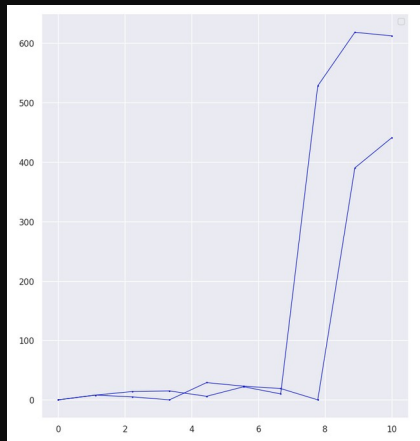
Hot /dead



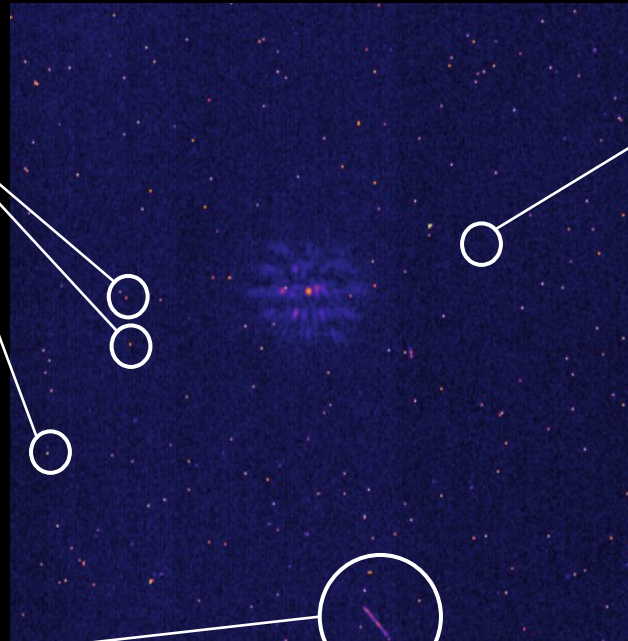
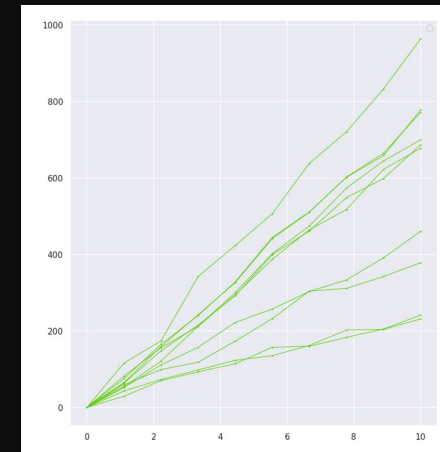
Low



Jum



Good

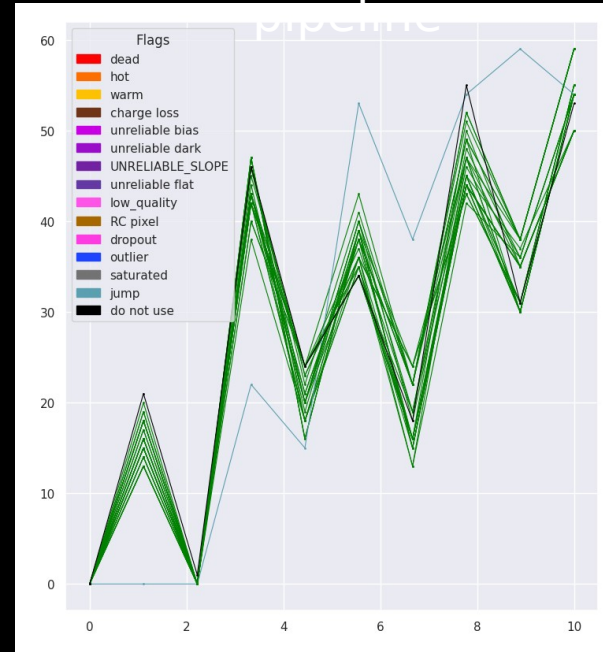
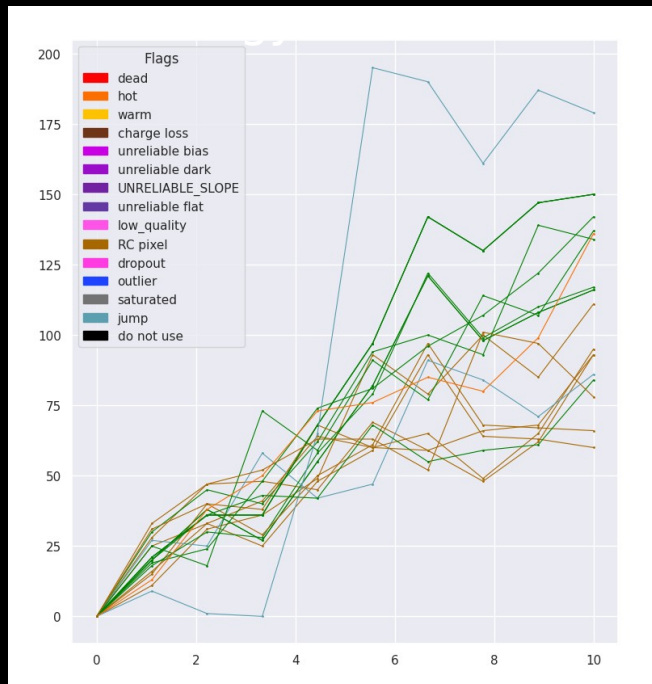


# | BAD PIXELS SUR JWST

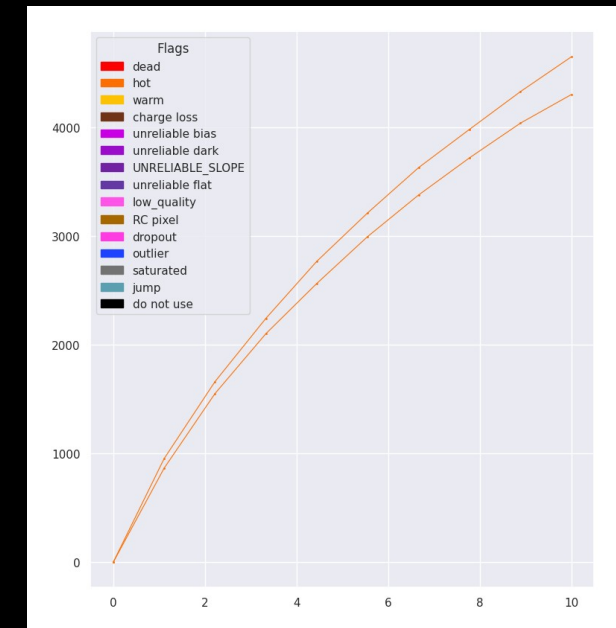
Défauts de la pipeline

« good » pixel d'après la pipeline

Le chaos des low-

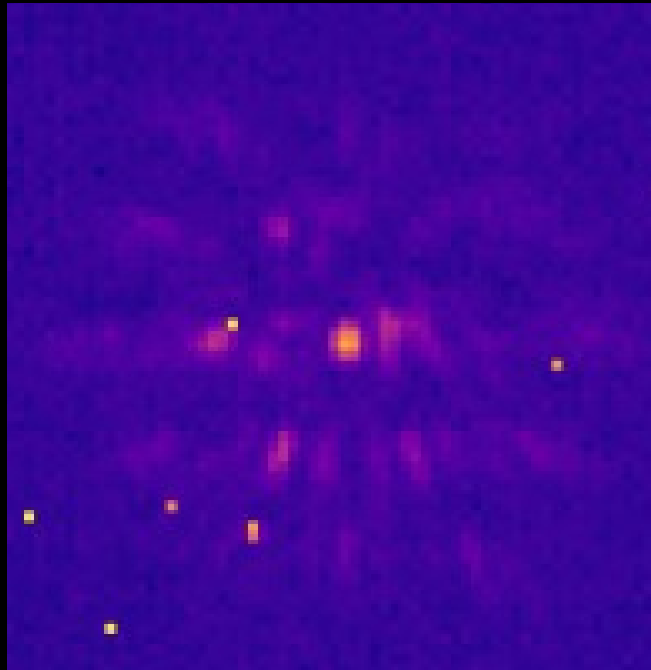


Les presque linéaires

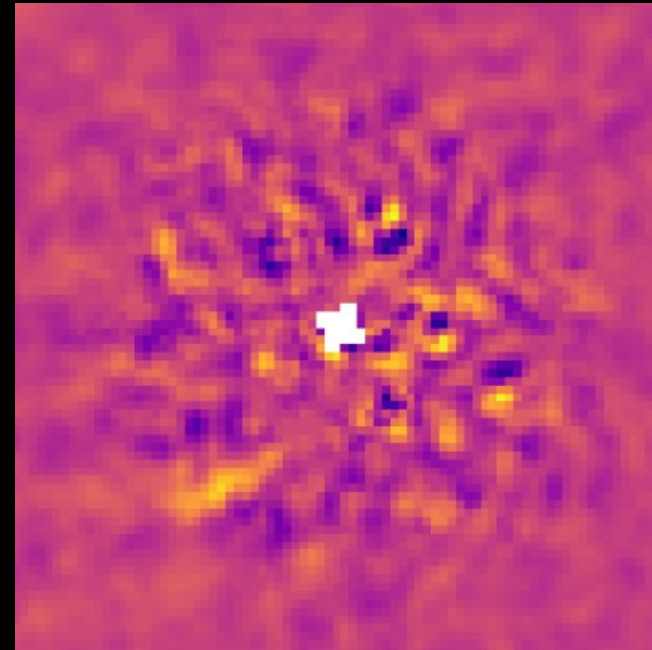


	good	bad
Précision	0.29	0.78
Rappel	0.57	0.52

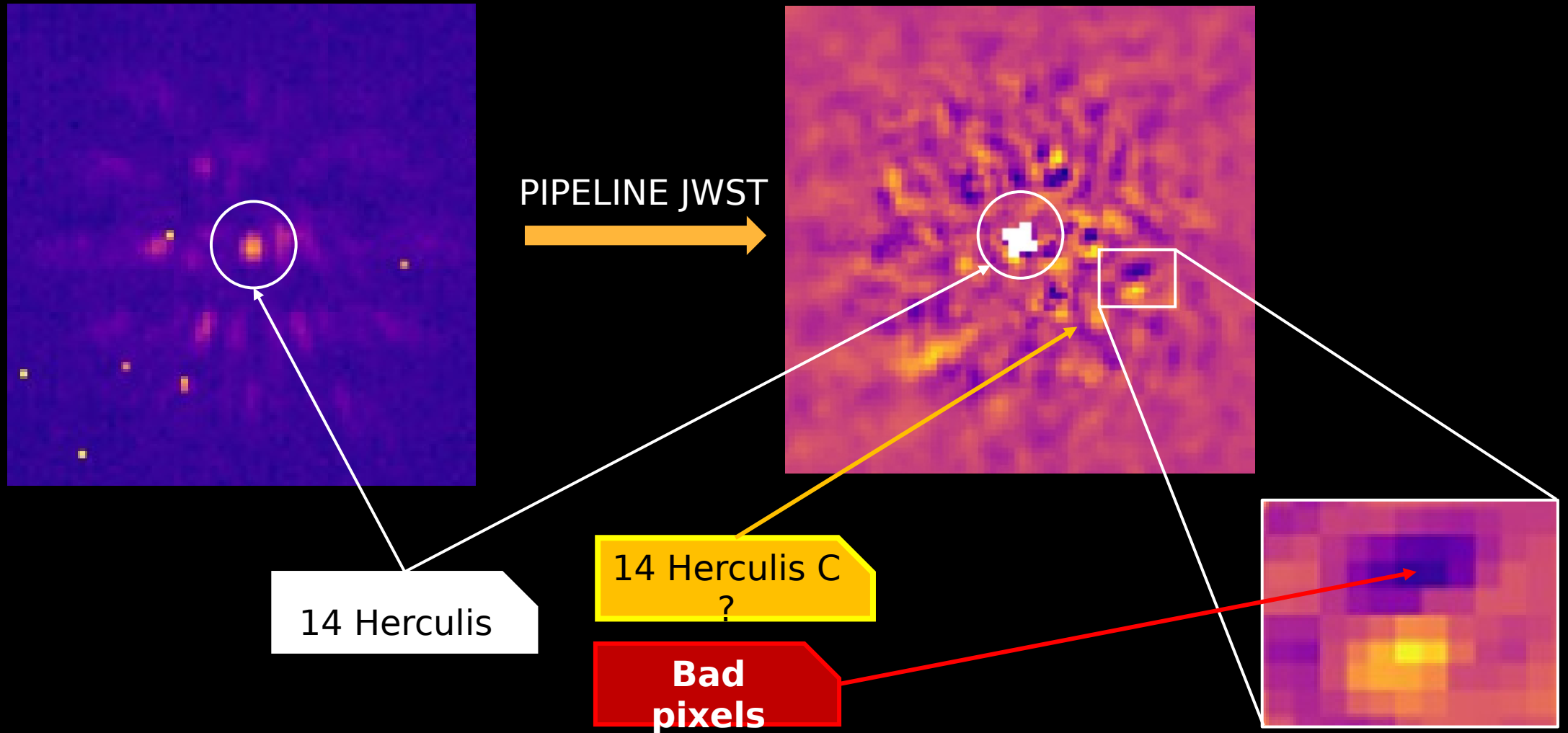
# | 14 HERCULIS C



PIPELINE JWST



# | 14 HERCULIS C



# | BAD PIXELS JWST - NON SUPERVISÉ



Bayesian  
Gaussian mixture

	<i>good</i>	<i>bad</i>
Précision	0.79	0.94
Rappel	0.84	0.92



DBScan

	<i>good</i>	<i>bad</i>
Précision	0.82	0.81
Rappel	0.32	0.97



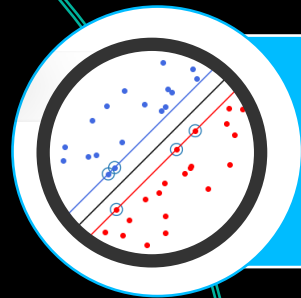
Spectral  
clustering

Pas assez de  
convergence



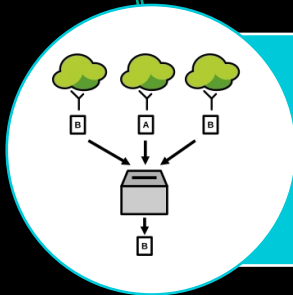
# | BAD PIXELS JWST - SUPERVISÉ

PHASE  
EXPÉRIMENTALE



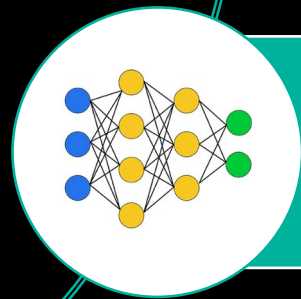
SVM

	<i>good</i>	<i>bad</i>
Précision	0.97	0.97
Rappel	0.91	0.99



Random  
Forest

	<i>good</i>	<i>bad</i>
Précision	0.90	0.96
Rappel	0.87	0.97

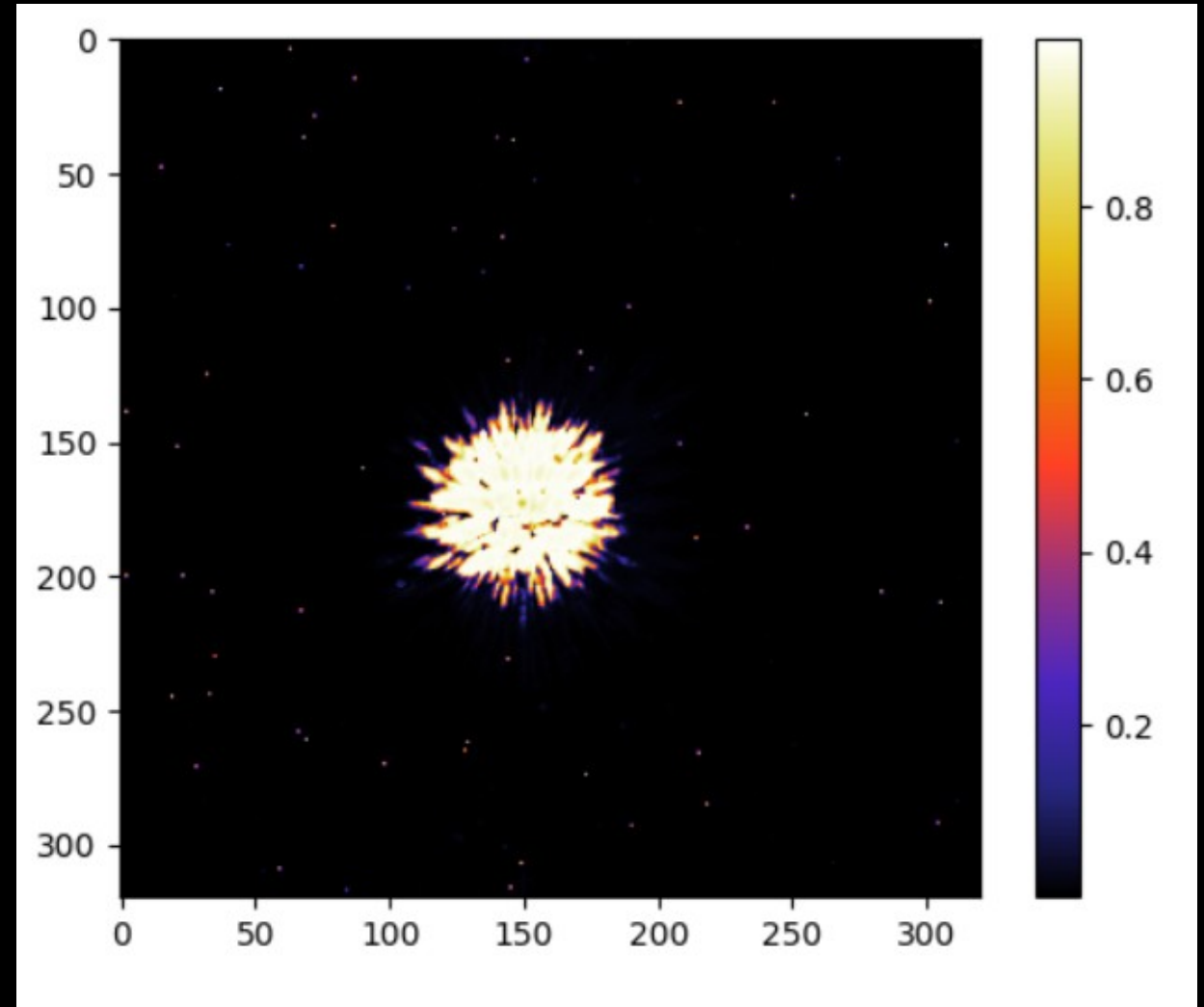
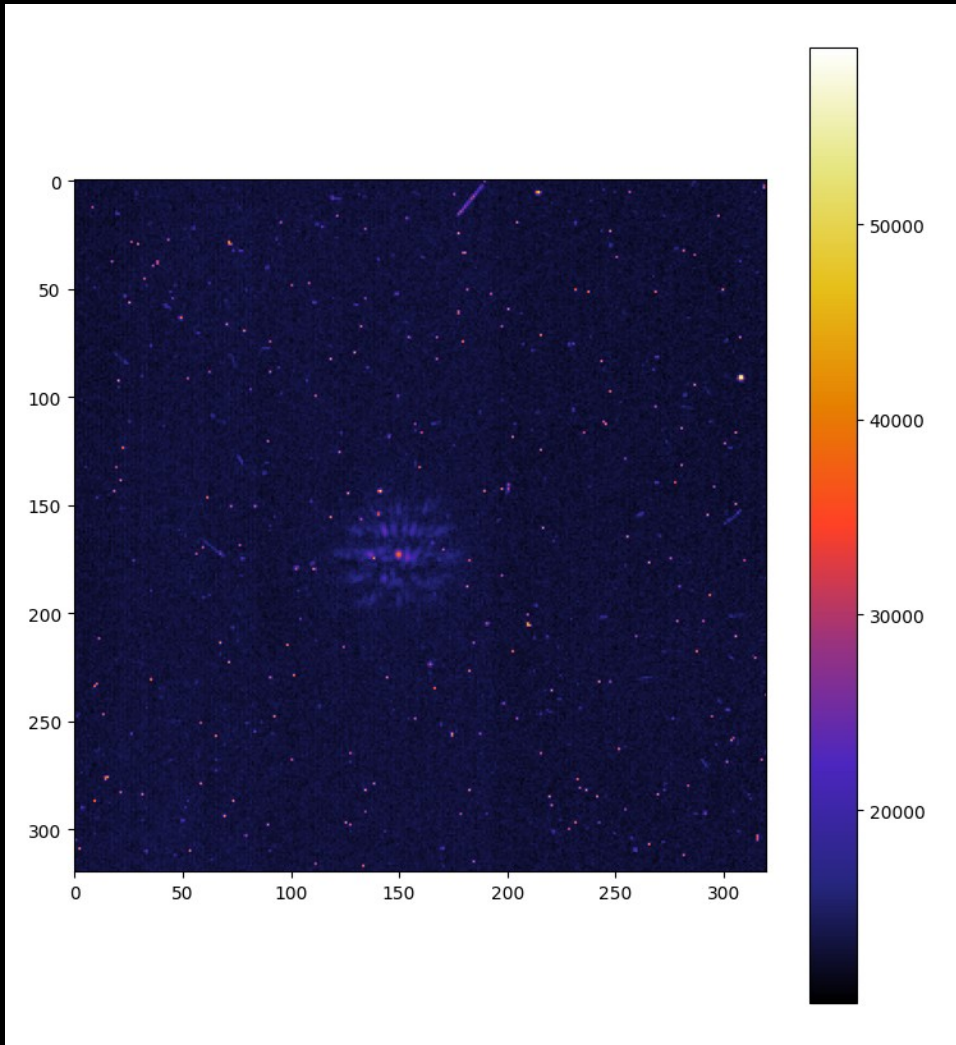


DNN

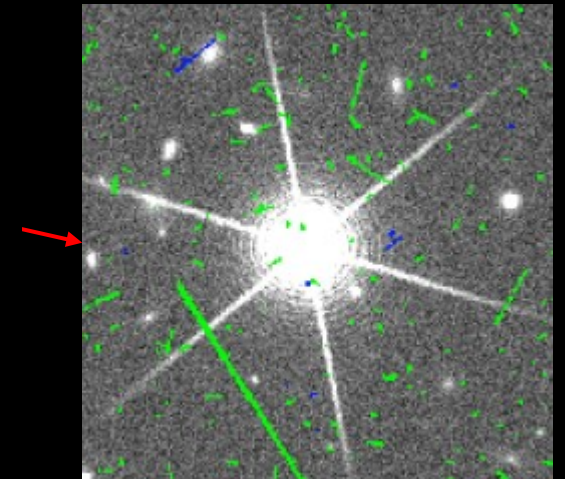
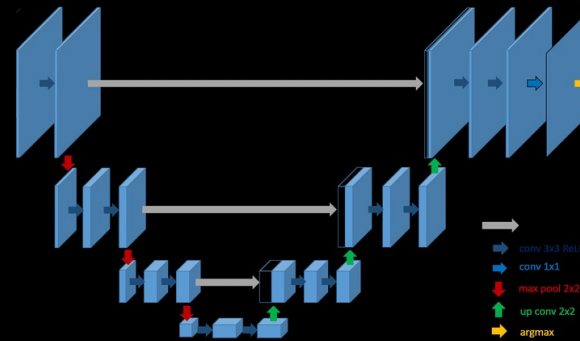
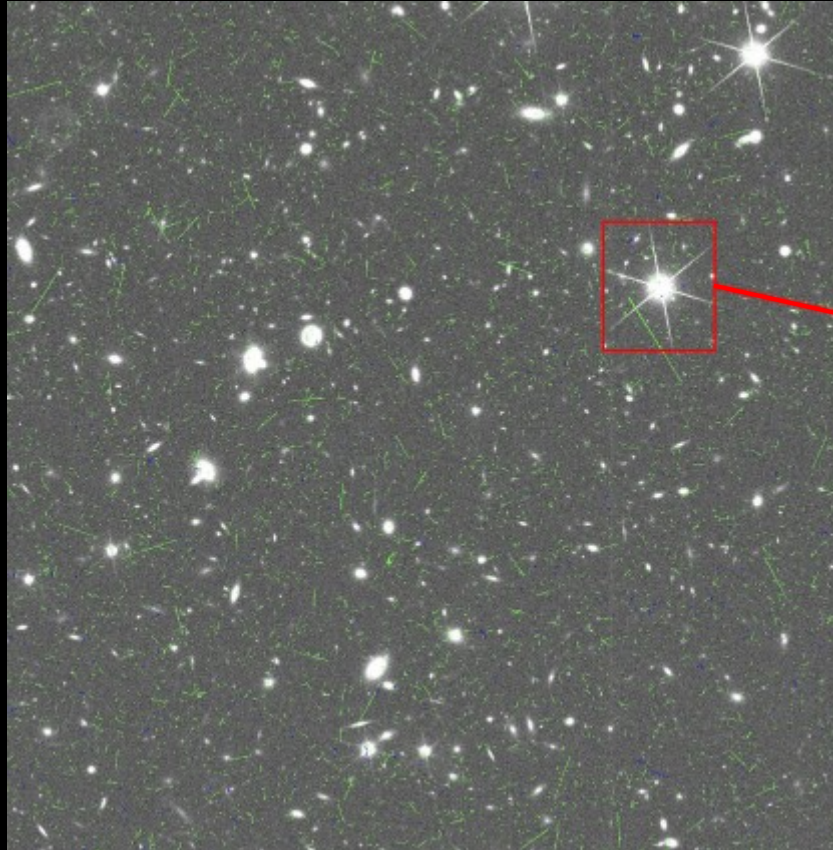
	<i>good</i>	<i>bad</i>
Précision	0.90	0.95
Rappel	0.83	0.97

# | CARTE DE CONFIANCE REVUE

PHASE  
EXPÉRIMENTALE



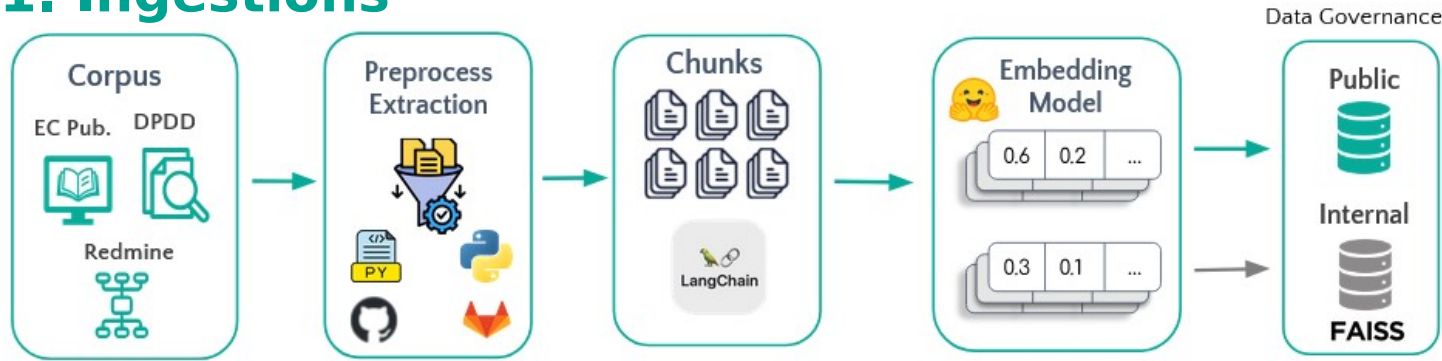
# | RAYON COSMIQUE SUR IMAGE VIS



IoU 93%

# #EUCLID | LLM RAG

## 1. Ingestions



Documentation du modèle de données + articles publiés

Redmine

## 2. Retrieval

