Using ESA Datalabs to create data analysis tools for the James Webb Space Telescope

Alice Young- ESAC YGT

Datalabs Workshop-November 24th 2022

Introduction

- One of the first projects at ESA which aims to utilise the new ESA Datalabs platform
- *Why Datalabs:*
    - All JWST data can be accessed as **mountable volumes** in datalabs
        - Data accessible to whole teams without downloading files onto each personal computer
    - NIRSpec data processing **pipeline** can easily be run in datalabs
    - Facilitates **collaboration** and sharing of data products within teams (i.e., NIRSpec Science Readiness Team during commissioning)

- *My project:* developing **alternate data analysis** tools than NIRSpec default to explore methods which could better utilise data for certain science goals
    - Build analysis tools using functions designed by instrument scientists for specific data products
    - Scientific community will have access to these notebooks when using NIRSpec data on ESA datalabs
    - **Specific task:** use background flux from observations taken using Multi-Object Spectroscopy (MOS) mode to create a master 2D spectrum to be used for background subtraction from science data

NIRSpec MOS Overview

- Multi-Object Spectroscopy utilises the Micro-Shutter Array (MSA)
- Allows spectra to be obtained for **several objects** within a single exposure
- Typical background subtraction method is **nodded background subtraction**
    - Nodd 1x3 slitlet containing three open shutters such that the target object appears **once in each shutter**
    - Subtract background captured by nodded sub-exposure with **pixel-to-pixel** subtraction
- **Caveats for Nodded Background Subtraction:**
    - Demands use of **1x3 slitlet** confi
    - **limits numb** observe in field of view in order to ing
    - also susceptible to **observational artefacts**
- **Pros for alternate method: Master Background Subtraction**
    - High signal to noise spectra → **extended object** use case
    - Allows for deviation from 1x3 slitlet configuration → **gain in multiplexing**

Possible alternative method: Master Background Subtraction!

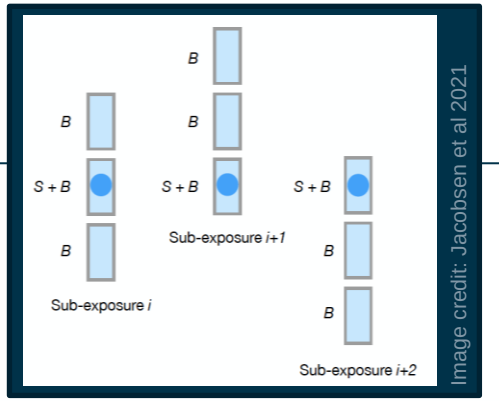


Image credit: Jacobsen et al 2021

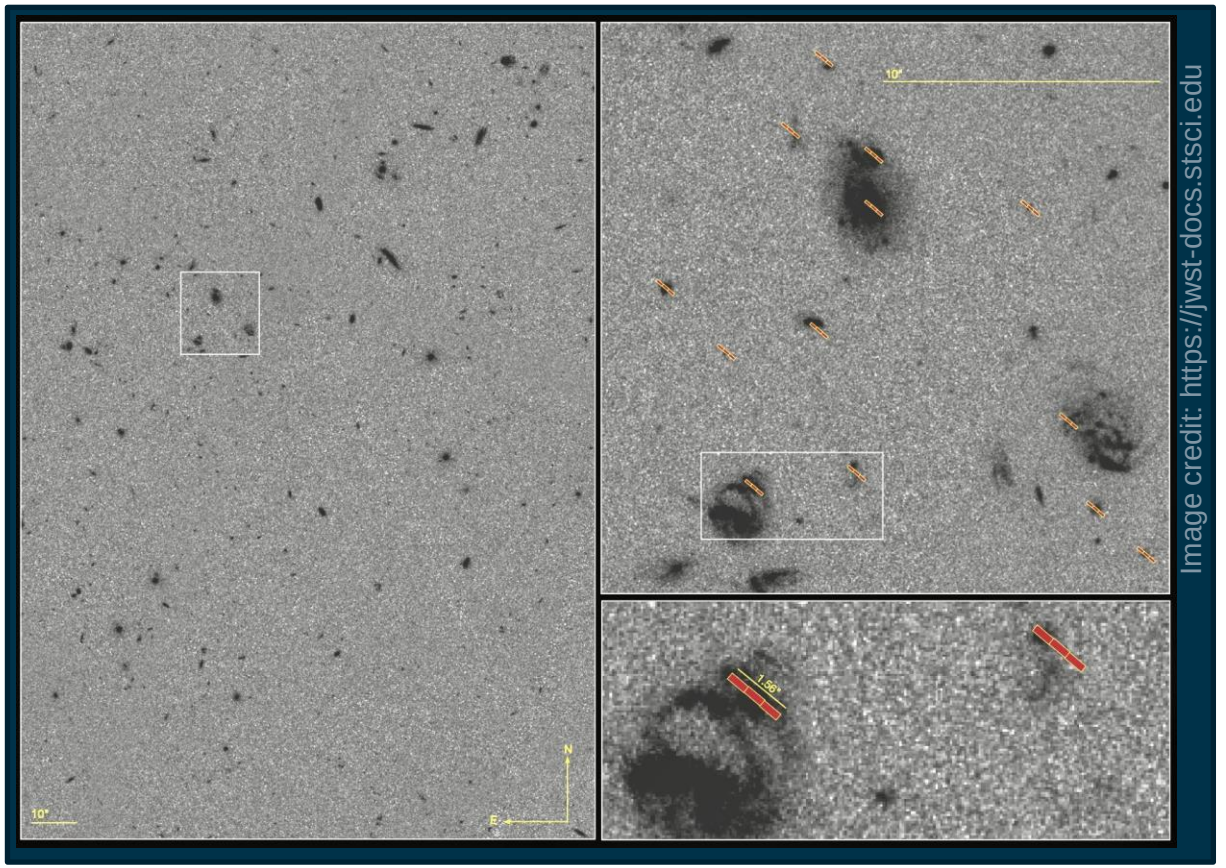
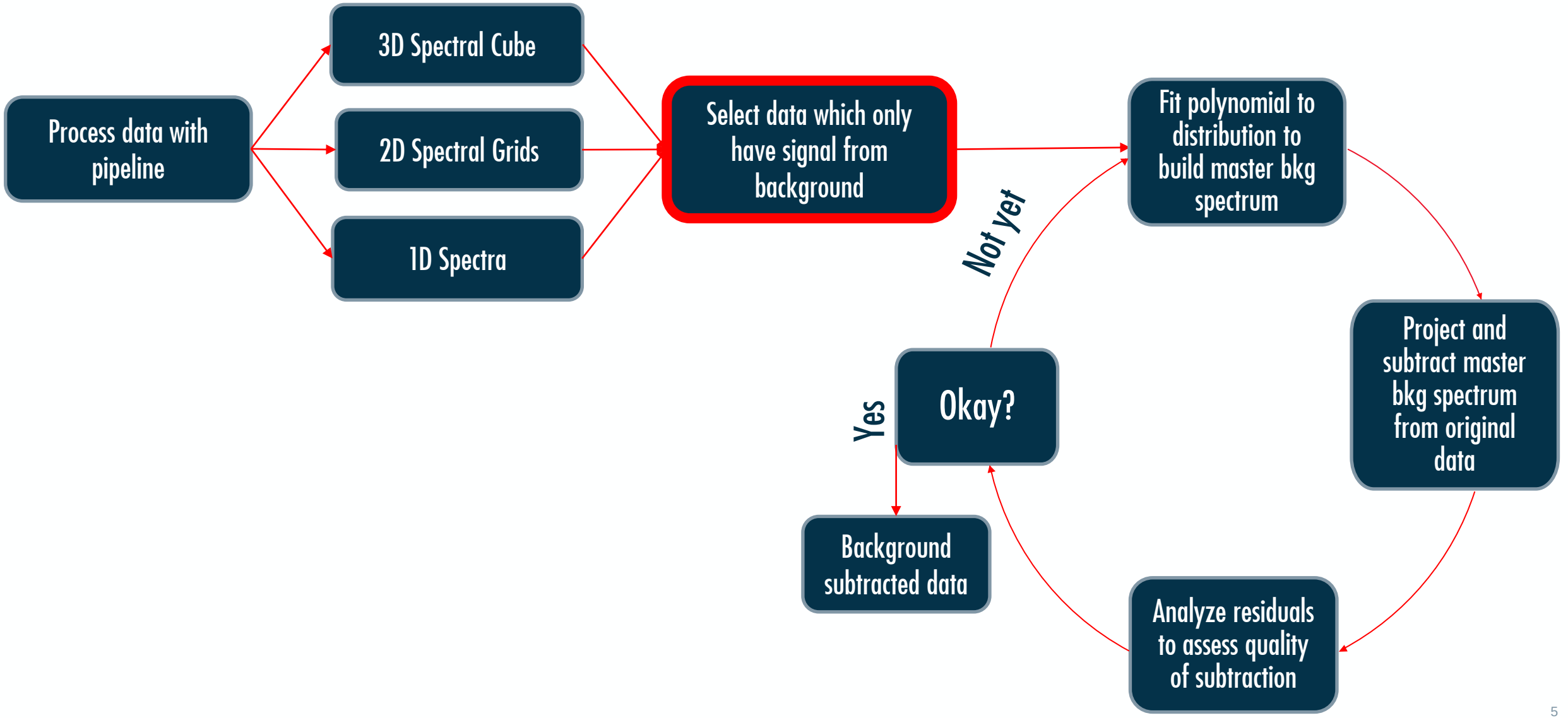


Image credit: https://jwst-docs.stsci.edu

Master Background Subtraction

- Break method up into three parts:
  - Define master background 2D spectrum
    - Use **Irregular** data products to avoid introducing correlations to data that enter after recalibration
    - Extract flux from slitlets containing only background in FOV
    - Fit polynomial to cloud of points in flux vs. wavelength profile from all relevant slitlets
  - Applying the master background 2D spectrum to science data
    - Define master background I2D using model profile and sampling into grid
    - Assume **constant** in spatial direction to build 2D grid
    - Project master background I2D to dimensions of science I2D for subtraction → f_rebin2d()
  - Testing the master background subtraction by comparison to nodded subtraction
    - Examine **normalised residuals** for both methods

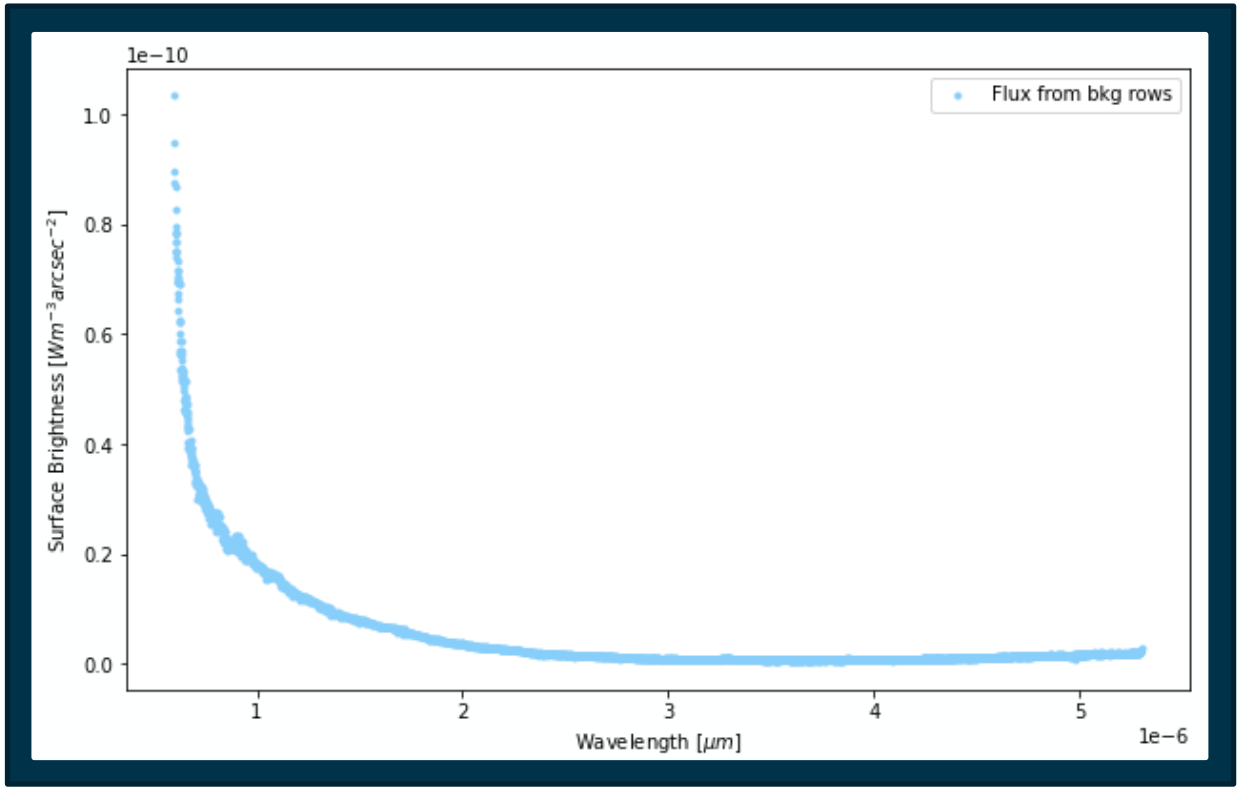
Example Workflow in ESA datalabs:



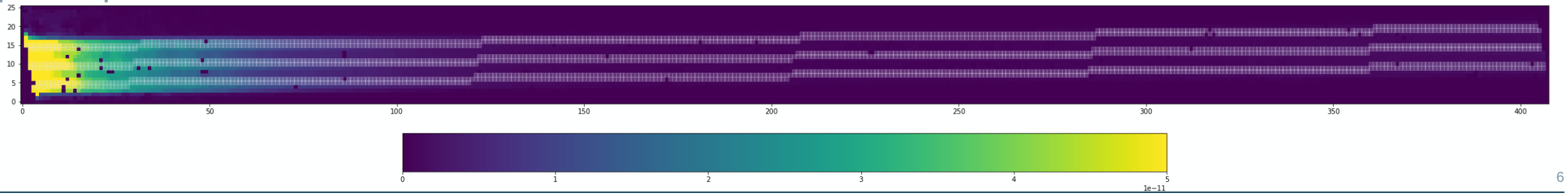
Select data which only have signal from background

- Load I2Ds corresponding to slitlets identified as containing only bkg flux in FOV
- For each I2D:
- Apply **quality flags** to get masked arrays of flux and error in each pixel
    - Also apply **median filter** to data
    - Use wavelength array from **AWCS extension**
- Don't use pixels affected by **bar shadow**
    - Extract flux from pixels with **yrel** values:

  yrel = [-1.2, -0.8], yrel = [-0.2, 0.2], yrel = [0.8, 1.2]

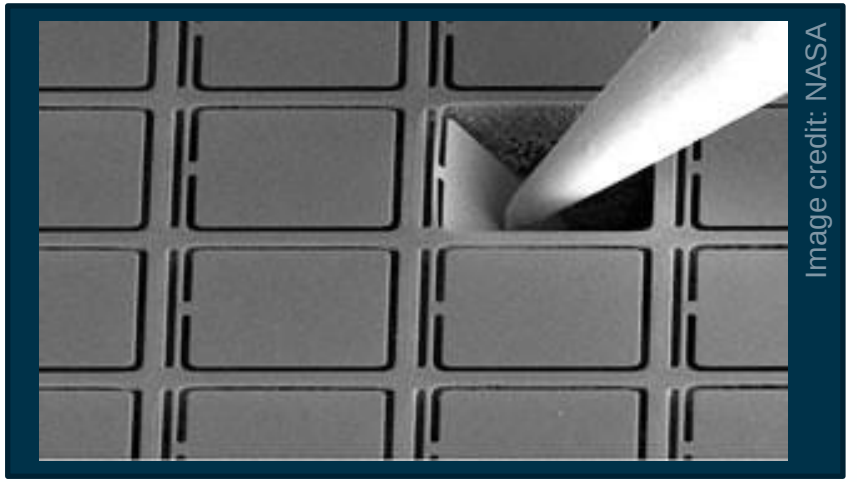


- Typical Example:



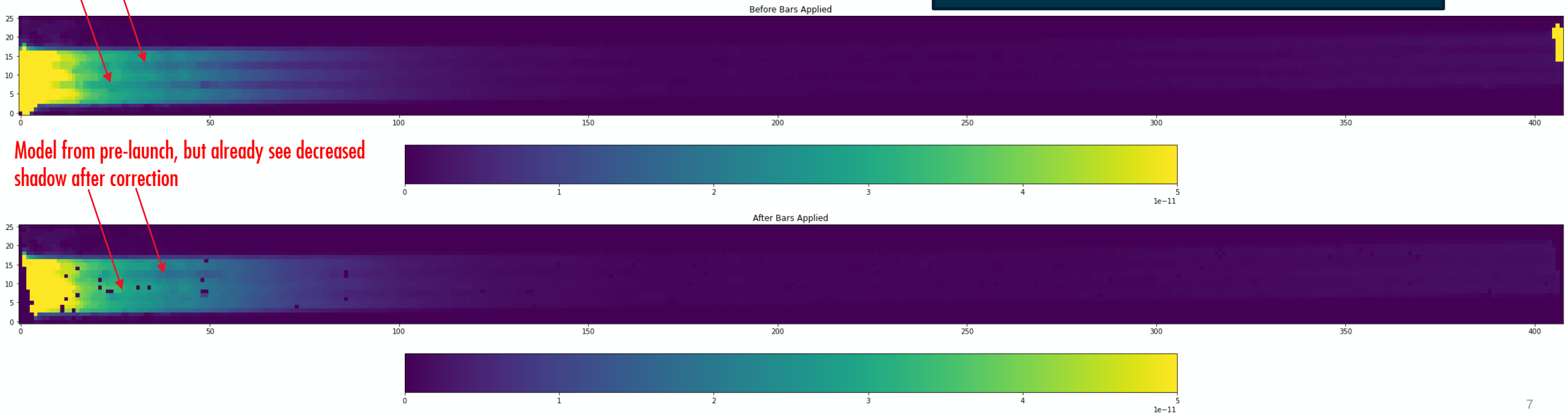
A note about the bar shadow correction

- Unique to NIRSpec observations
- Corrects for shadow of MSA on detector
  - accounts for fact the PSF of instrument is convolved with top-hat function
  - Would expect:
  - But once convolved with PSF get smoothing, e.g.:



The bar shadows

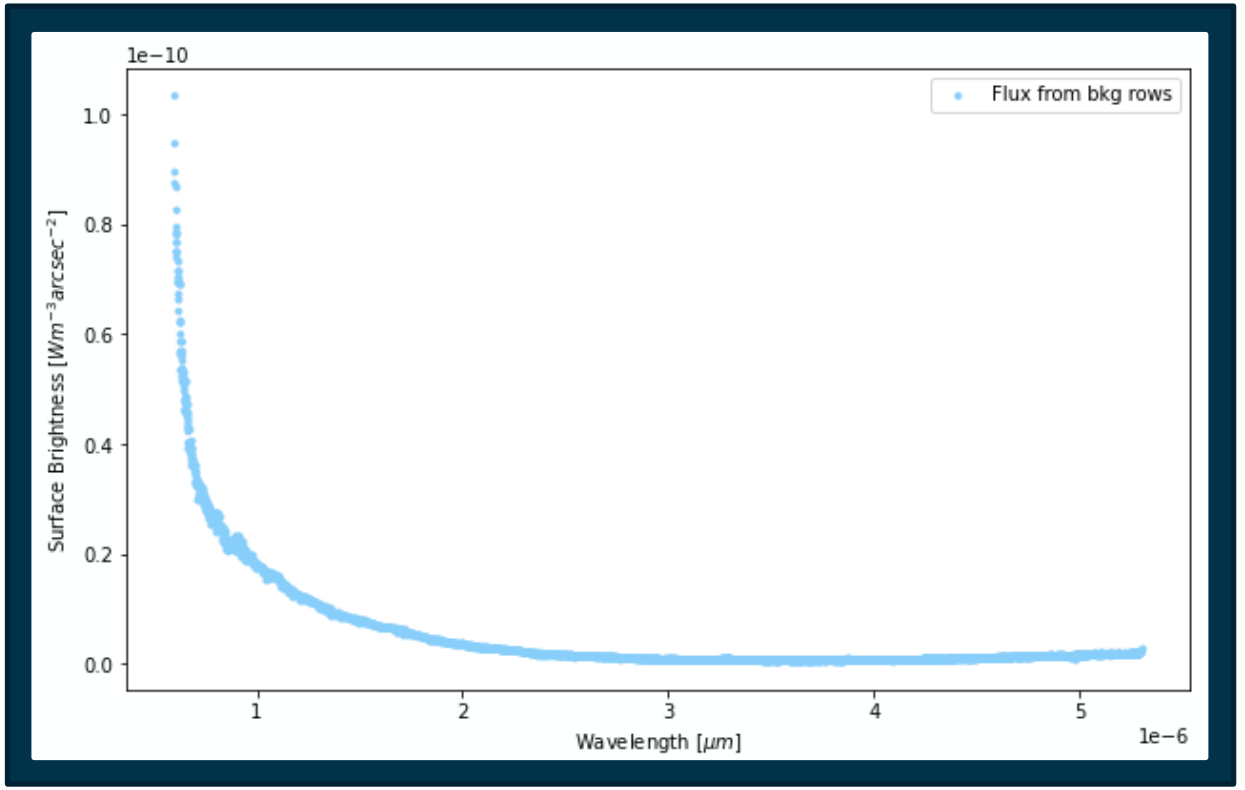
Model from pre-launch, but already see decreased shadow after correction



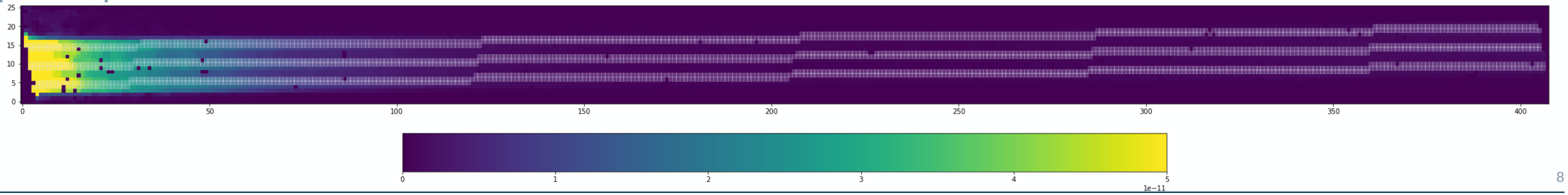
Select data which only have signal from background

- Load I2Ds corresponding to slitlets identified as containing only bkg flux in FOV
- For each I2D:
- Apply **quality flags** to get masked arrays of flux and error in each pixel
    - Also apply **median filter** to data
    - Use wavelength array from **AWCS extension**
- Don't use pixels affected by **bar shadow**
    - Extract flux from pixels with **yrel** values:

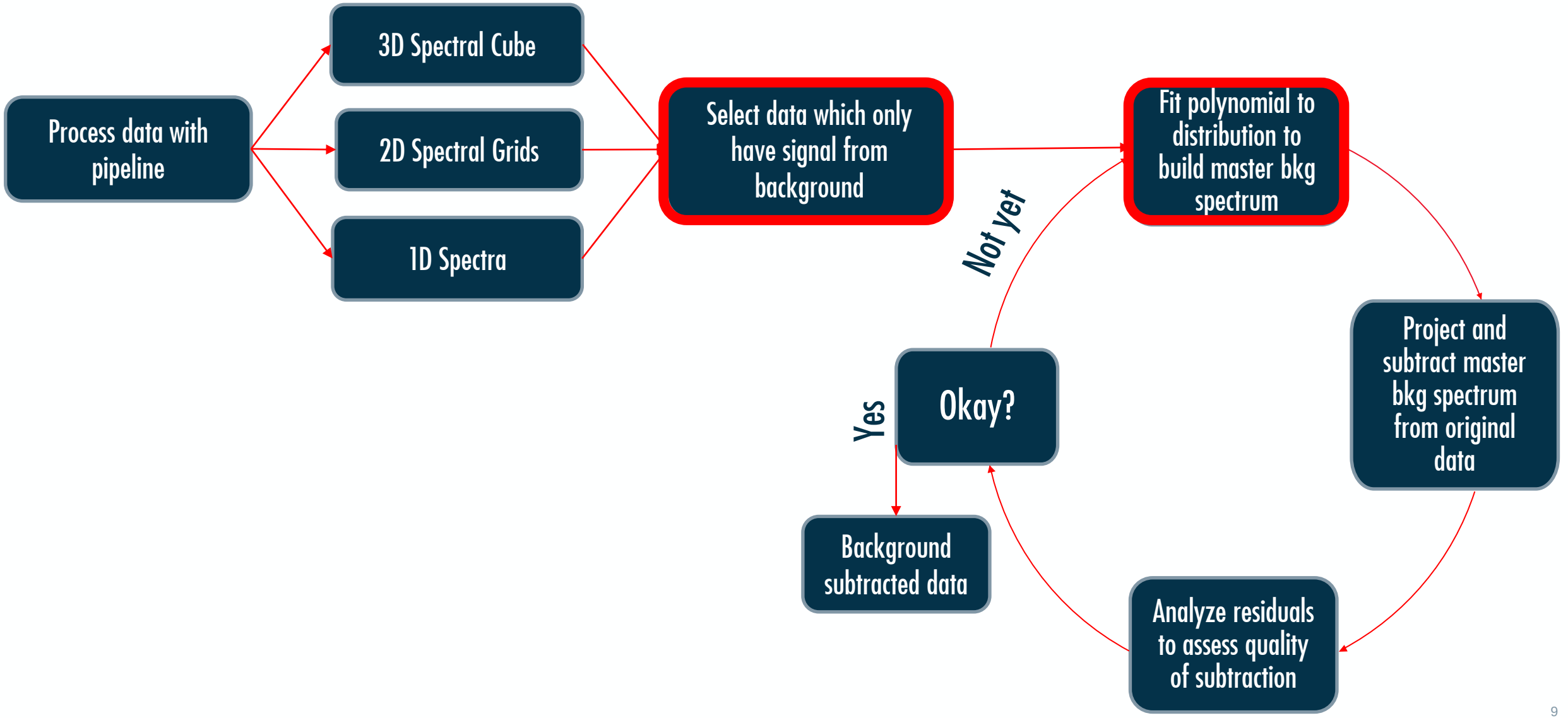
  yrel = [-1.2, -0.8], yrel = [-0.2, 0.2], yrel = [0.8, 1.2]



- Typical Example:

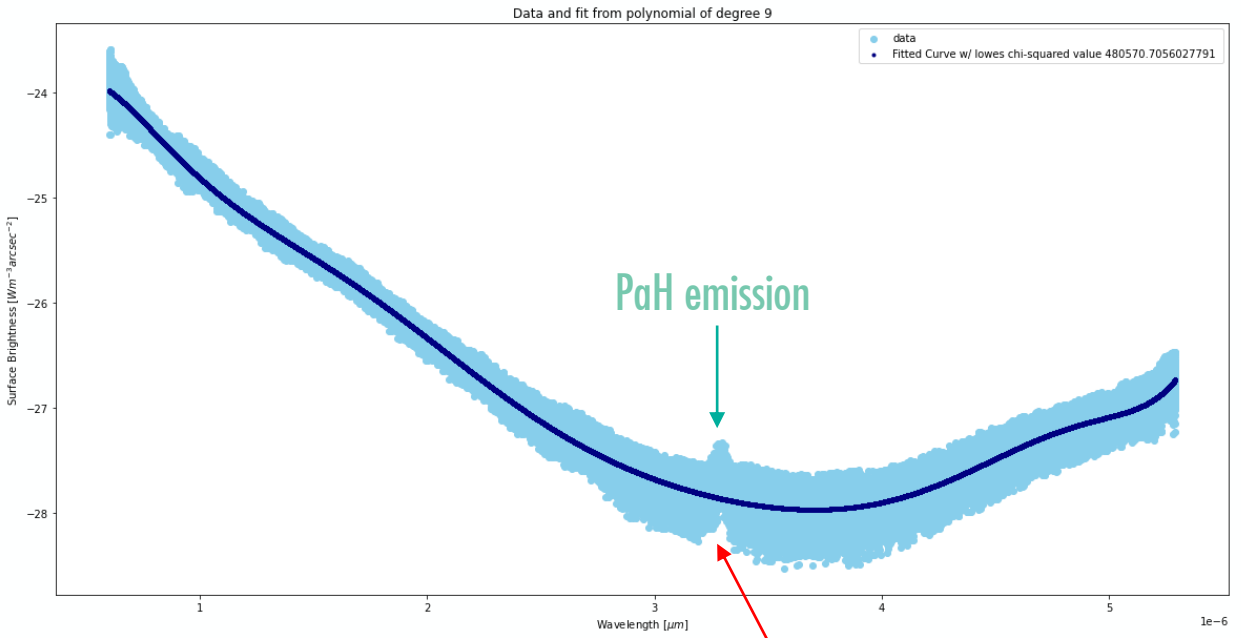


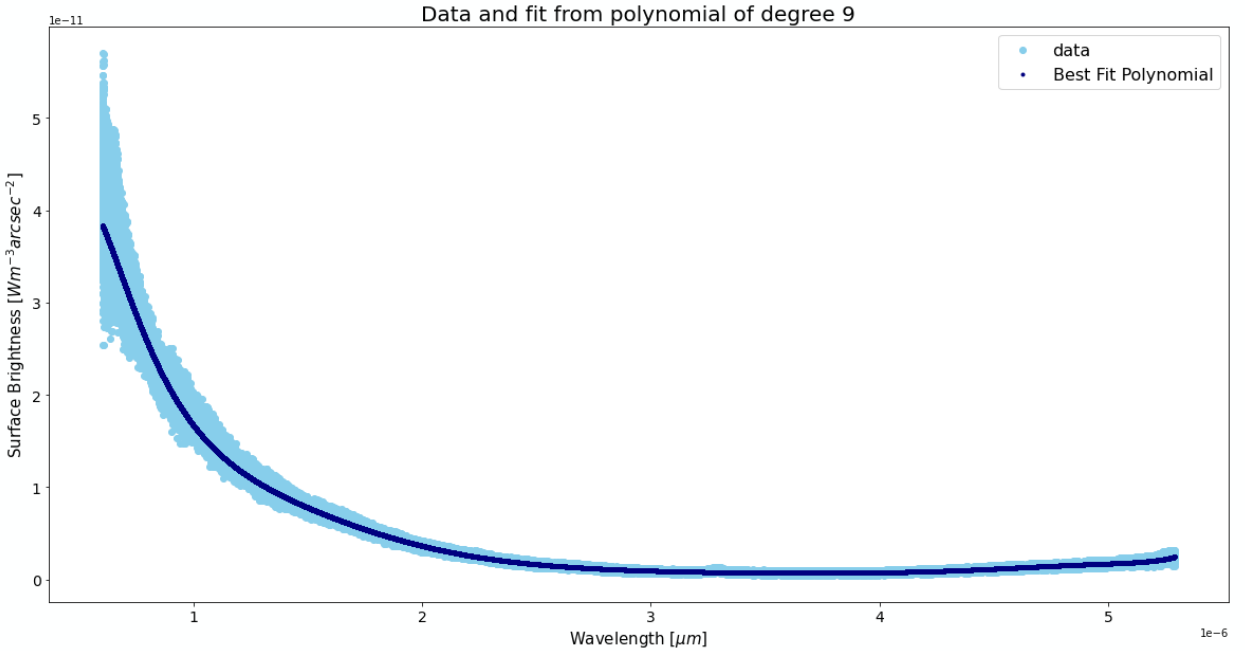
Example Workflow in ESA datalabs:



Fit polynomial to distribution to build master bkg spectrum

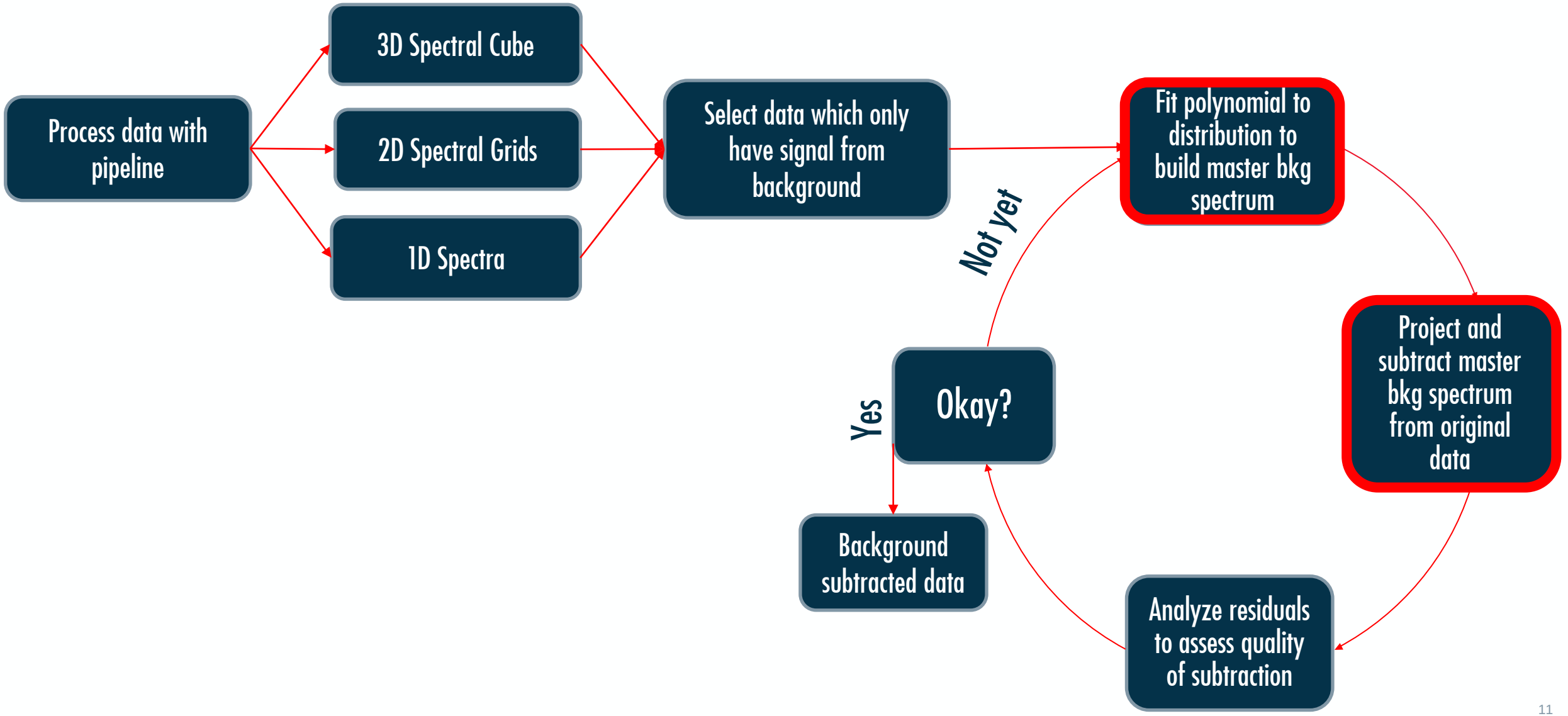
- Use extracted **sigma-clipped** flux from relevant pixels in **all I2Ds** and plot vs. wavelength
- Take the log and interatively fit polynomials of increasing degrees
  - Calculate χ^2 for each fit
- Return best fit polynomial coefficients corresponding to lowest χ^2 fit





Not subtracted with current fitting, but can be added

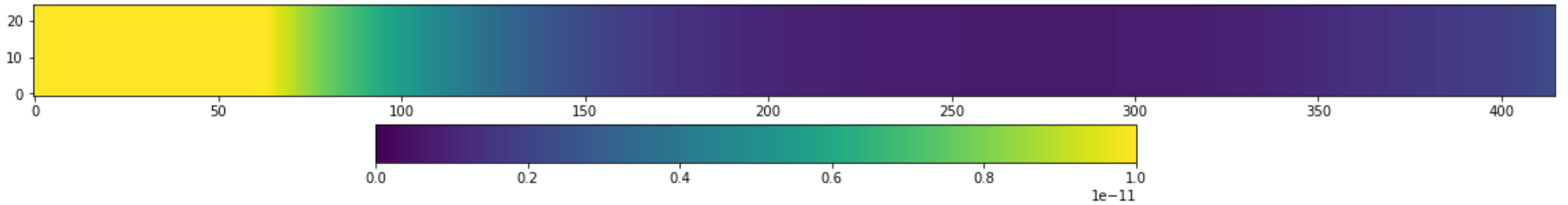
Example Workflow in ESA datalabs:



Project and subtract master bkg spectrum from original data

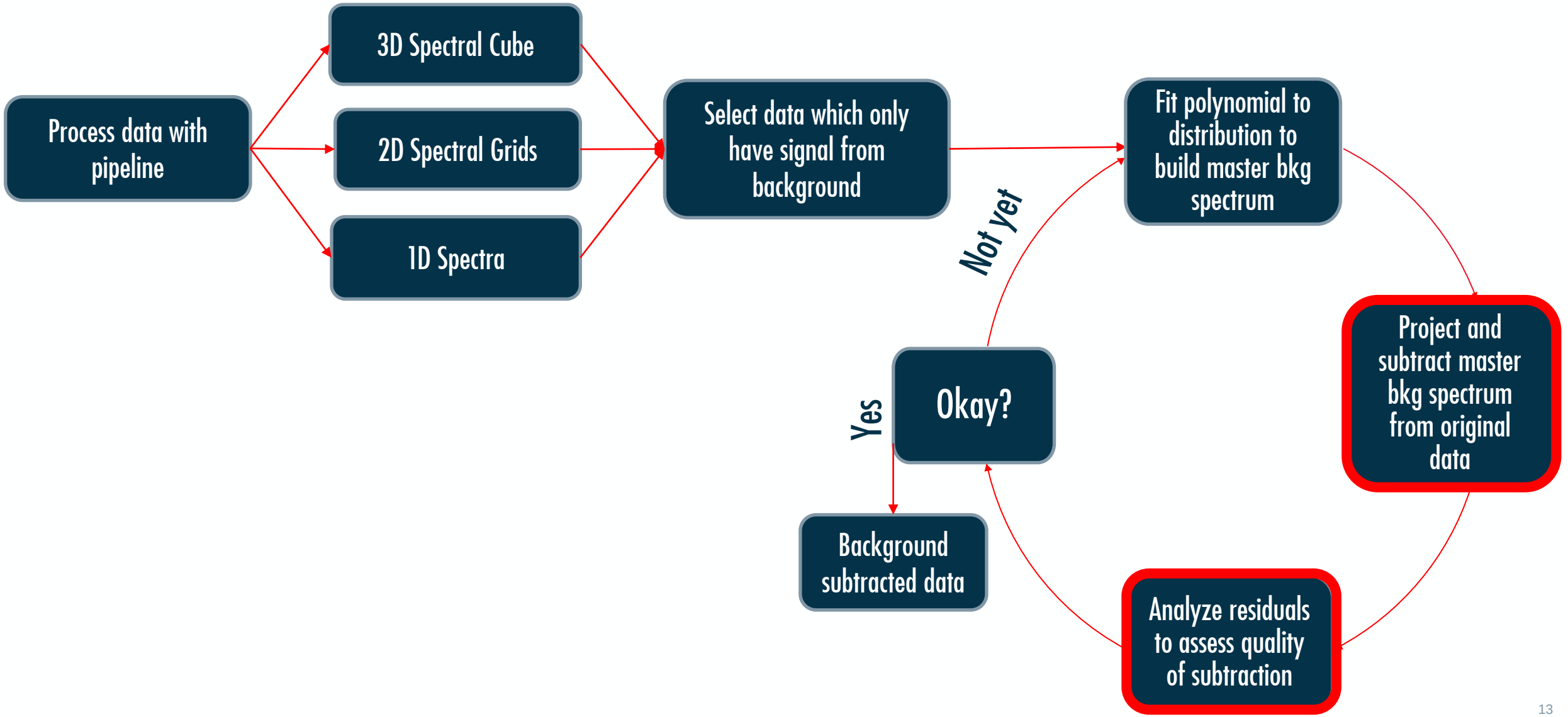
- Now have model for flux vs. wavelength
  - Need to **sample flux** data in wavelength to create pixel grid in spectral direction
  - Must make **sufficiently small** wavelength intervals to avoid losing information (doesn't matter too much in this case, but for other use cases with spectral lines it will be important)
- Copy flux from spectral direction in spatial direction to make grid

→ **Now have master background I2D spectrum**



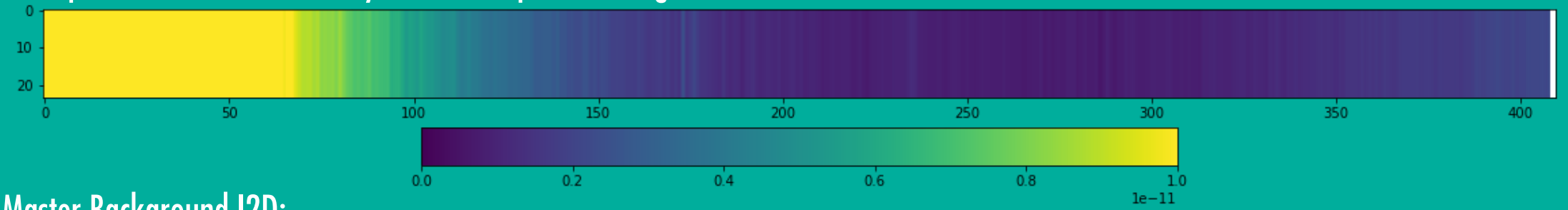
- To subtract model from each I2D, need same grid dimensions
- Iteratively **project model grid** to the dimensions of I2D which needs bkg subtraction
- Use **f_rebin2d()** from rectification pipeline step
- Subtract projected model grid from I2D data

Example Workflow in ESA datalabs:

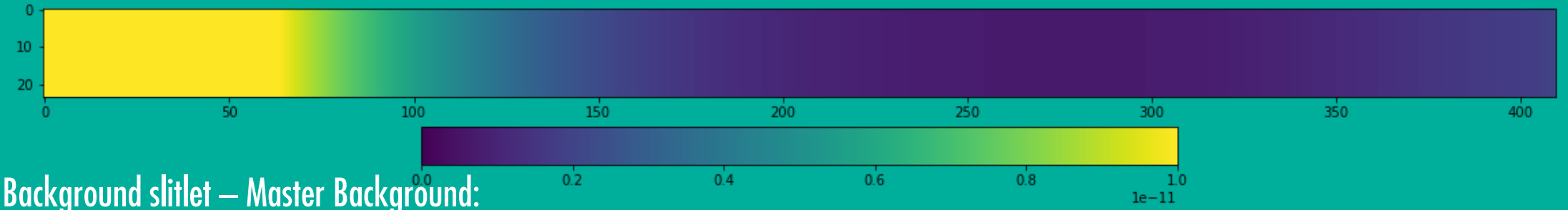


Analyze residuals to assess quality of subtraction

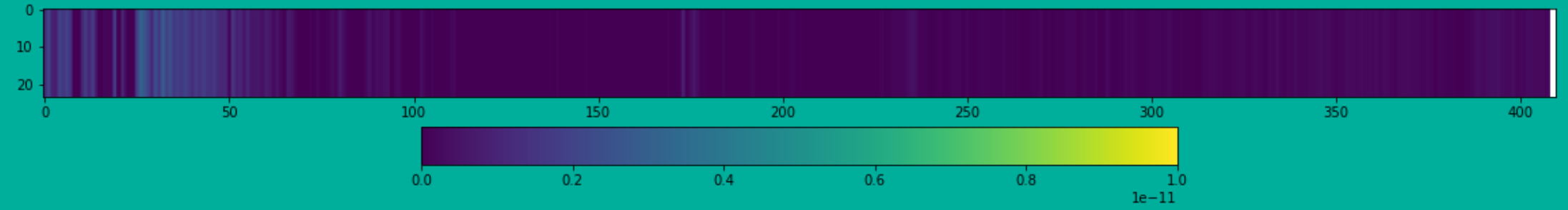
Example I2D constructed from yrel selected pixels of bkg slitlet:



Master Background I2D:



Background slitlet – Master Background:



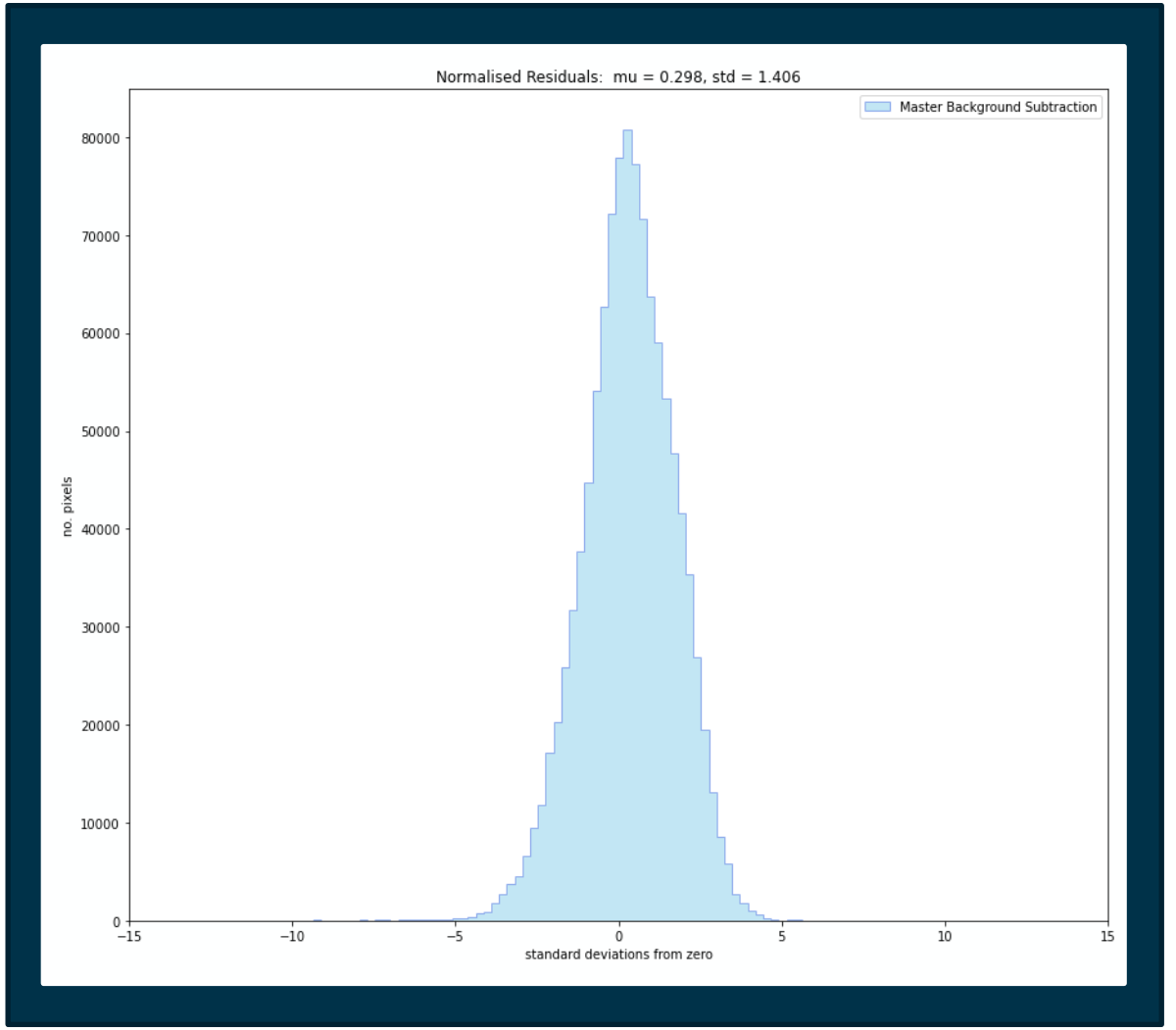
Analyze residuals to assess quality of subtraction

- Calculate **normalised** residuals for each pixel of each I2D after subtraction:

$$\sigma_{from\ zero} = \frac{(data\ flux_{pix} - model\ flux_{pix})}{data\ err_{pix}}$$

- Approximately Gaussian residuals
  - Widening due to **imperfections in model**

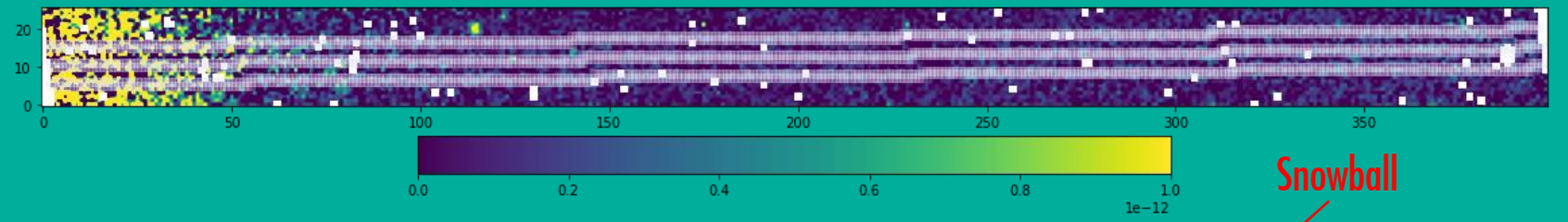
- More under-subtracted pixels than over-subtracted



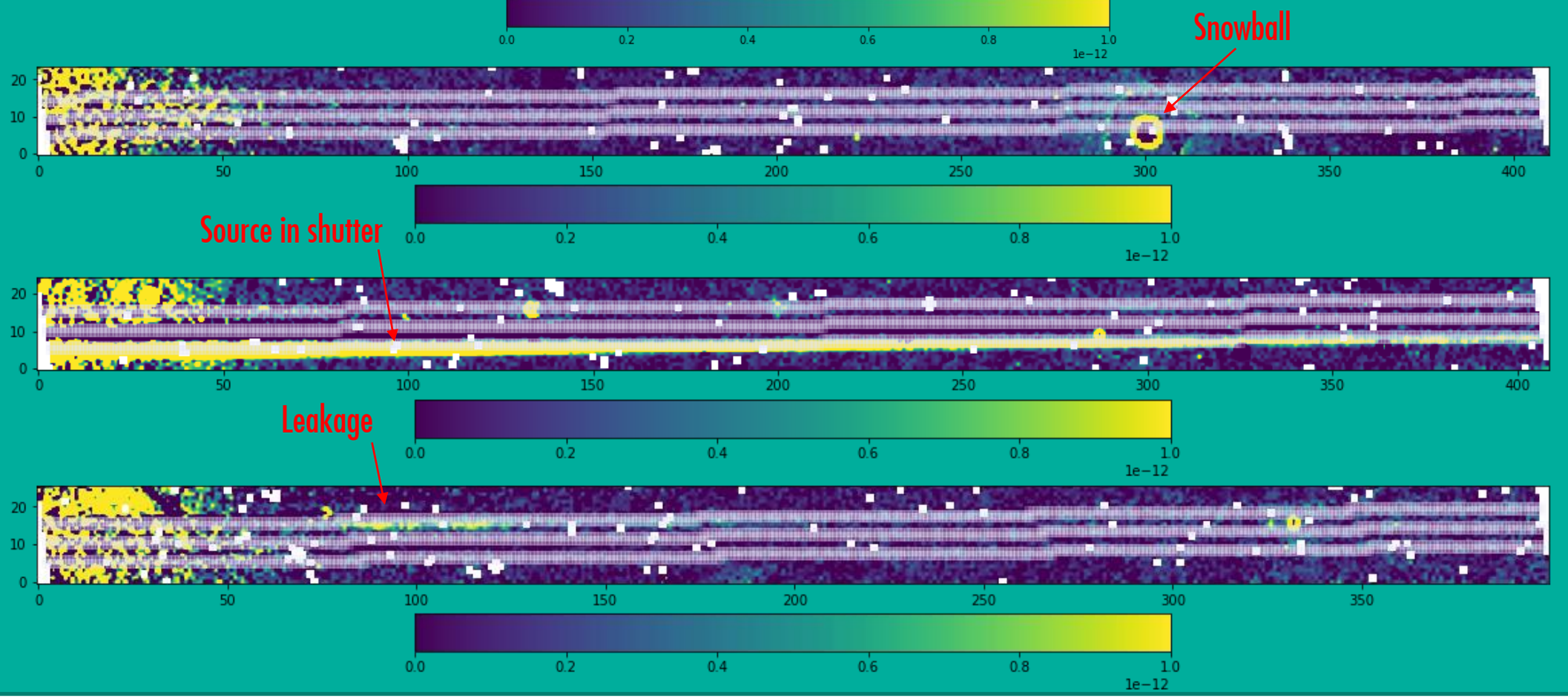
Nodded Subtraction

Select flux from pixels with required yrel values from I2Ds which have **nodded bkg subtraction**

Typical I2D:

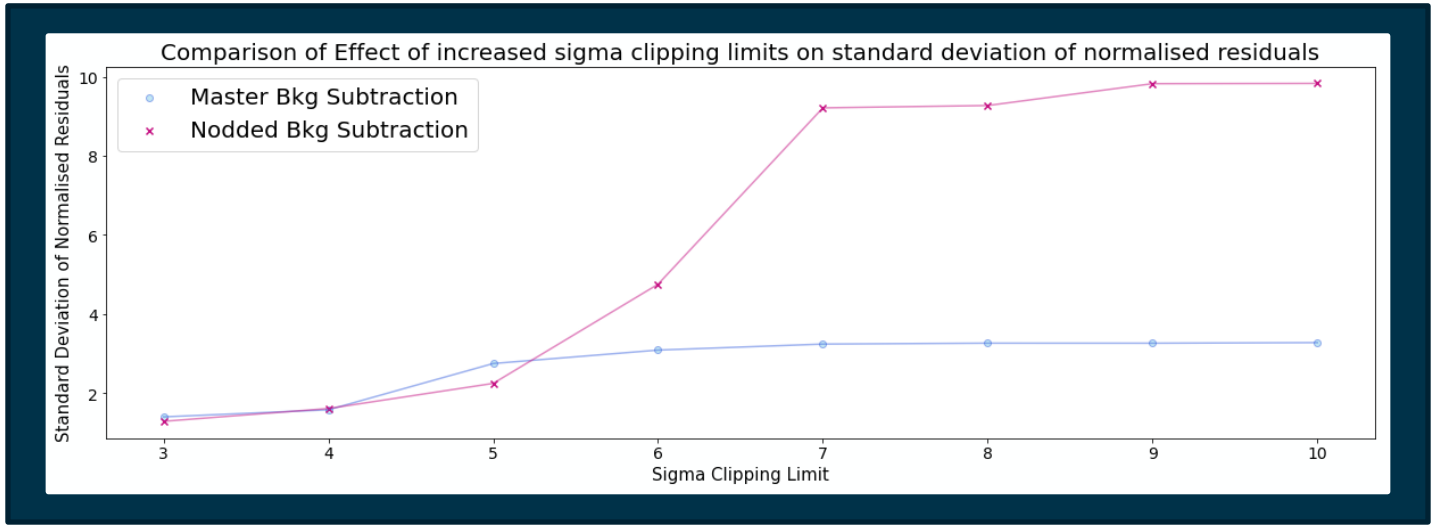


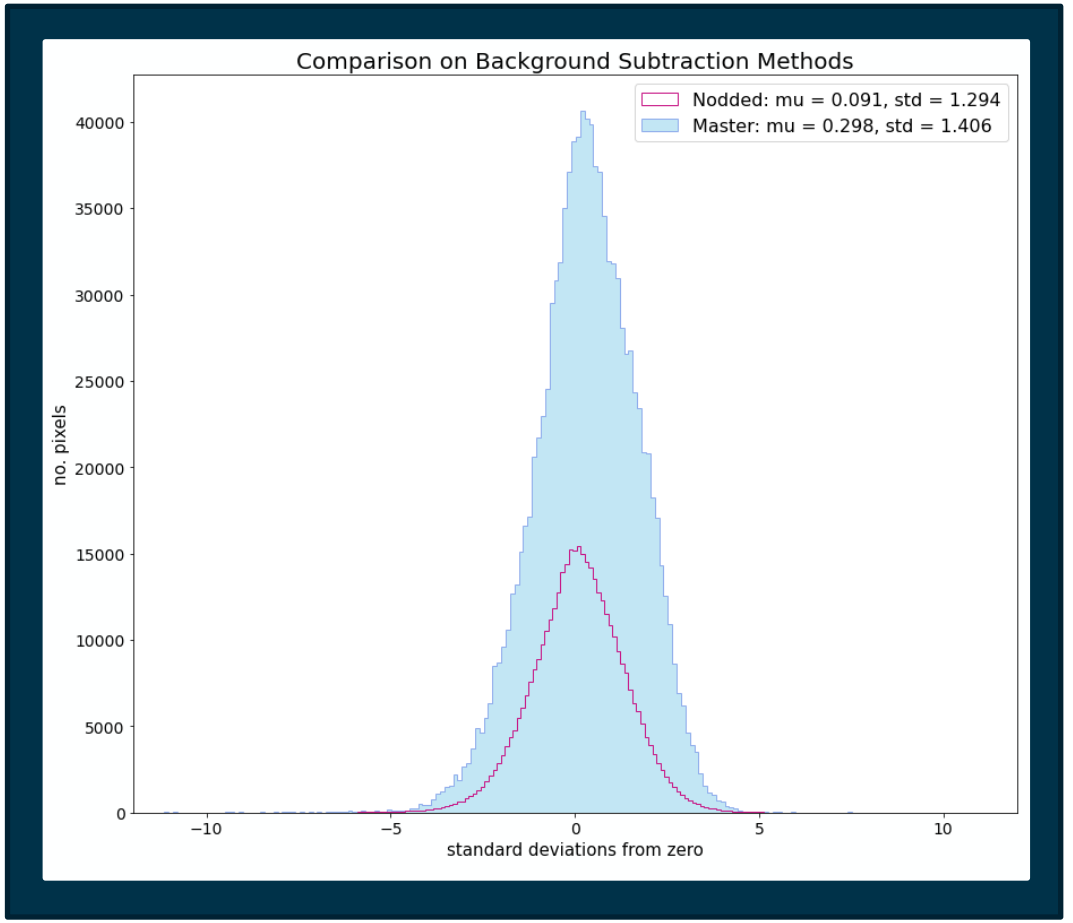
Outliers:



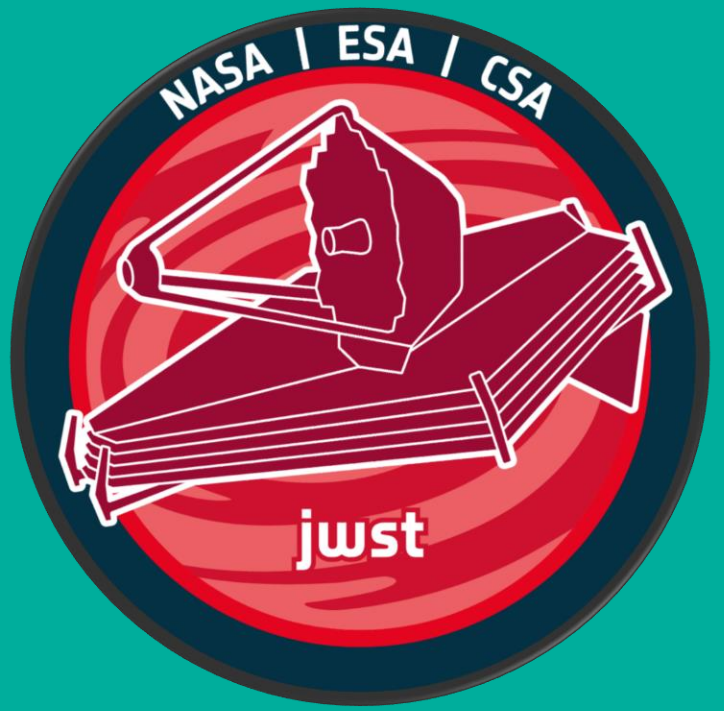
Comparison of two methods

- **Results consistent** with work done at R1D level by NSRT team during comissioning
- **Slightly smaller** standard deviation for nodded bkg subtraction
- Could be useful for observations which want to **maximise shutter use** for science targets
  - Gain in multiplex
- Confirms master bkg subtraction scheme is very promising





What does this look like in Datalabs?



Thank you!
Any questions?