

Data reduction and analysis of

FIFI-LS data

Dario Fadda

SOFIA Science Center



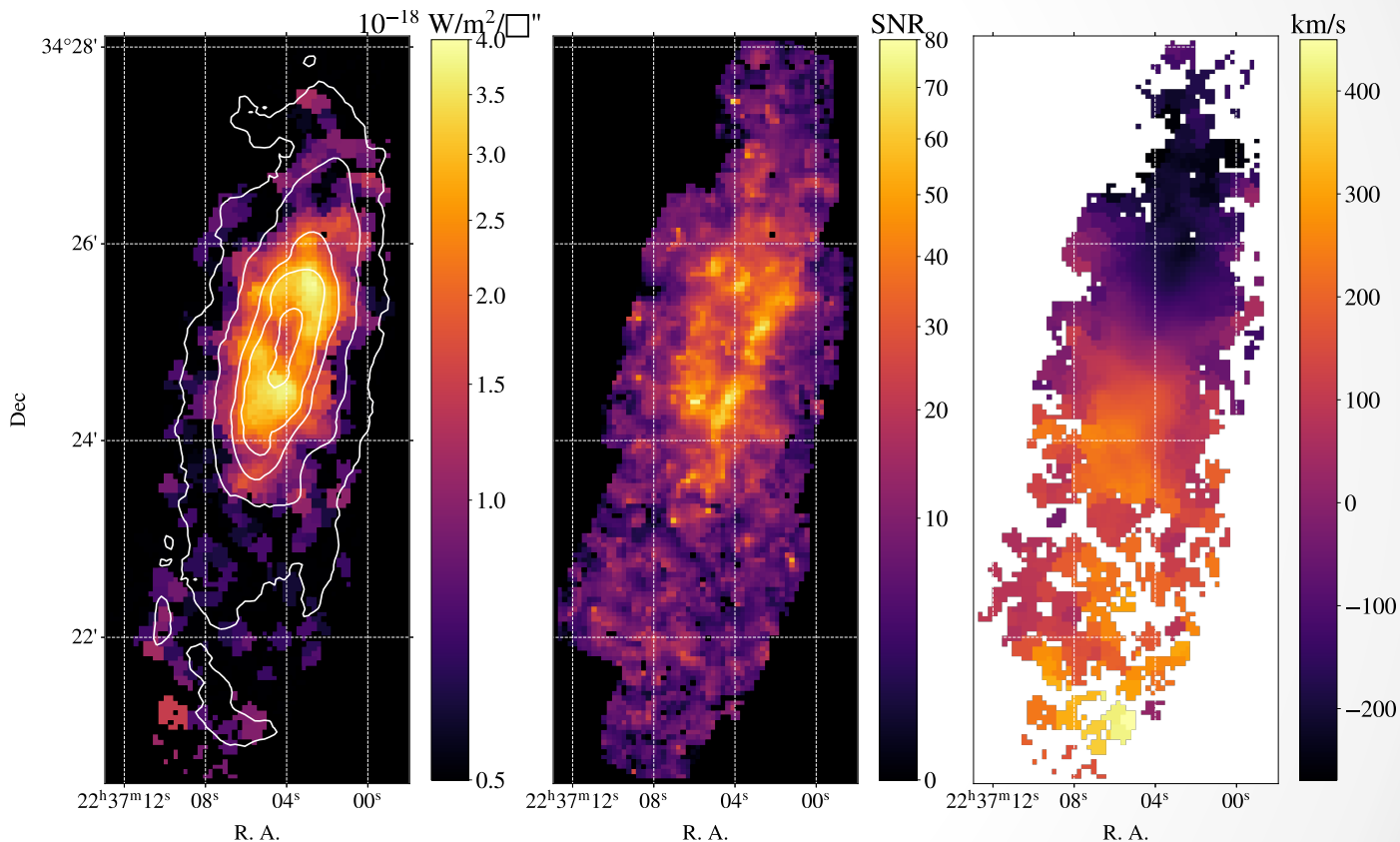
Dario Fadda
SOFIA data school - 16 Apr 2023



Outline

We will go through three essential steps with FIFI-LS data:

1. Download
2. Reduce
3. Analyze



Getting the data: selection

The screenshot shows the SOFIA Search web interface. At the top, there is a navigation bar with the SOFIA logo, IRSA logo, and menu items: IRSA | DATA SETS | SEARCH | TOOLS | HELP. A 'Login' link is on the right. Below the navigation bar are three buttons: Search, Catalogs, and Help. The main content area is titled 'SOFIA Search' and contains several sections for filtering search results:

- Spatial Constraints:** Includes radio buttons for 'Object/Position', 'Multiple Positions', 'Solar System Target', 'Precovery', and 'All-Sky'. A text input field contains 'NGC 7331' with a dropdown menu 'Try NED then Simbad'. Below it, coordinates are listed: '339.26672, 34.41552 Equ J2000 or 22h37m04.01s, +34d24m55.9s Equ J2000'. A 'Radius' input field is set to '100' with a unit dropdown set to 'arcseconds'. A note states 'Valid range between: 1" and 18000"'. A help icon (?) is at the bottom right of this section.
- Proposal Constraints:** A section with a help icon (?) at the bottom right.
- Observation Constraints:** A section with a help icon (?) at the bottom right.
- Instrument Constraints:** Includes a dropdown menu set to 'FIFI-LS', a 'Configuration' dropdown set to 'Med Res. Spectroscopy (R ~ 1,000)', a 'Spectral Element' dropdown set to 'Any', and 'Wavelength (um)' fields with 'From' and 'To' inputs. A help icon (?) is at the bottom right.
- Data Product Constraints:** Includes 'Processing Level' checkboxes for Level 0, Level 1, Level 2, Level 3, and Level 4 (which is checked), and an 'Observation Type' dropdown set to 'Any'. A help icon (?) is at the bottom right.

At the bottom of the search form are 'Search' and 'Cancel' buttons. The footer contains contact information, logos for ipac, Caltech, JPL, and NASA, and version information: 'v4.0.2021.3, Built On: 2022-04-18'.

Search tab

Object name

Instrument

Processing level

Search button



Dario Fadda
SOFIA data school - 16 Apr 2023



Getting the data: download

Download

Select AORs

Different observations

Coverage

The screenshot displays the IRSA (Infrared Science Archive) interface. At the top, there is a navigation bar with 'IRSA | DATA SETS | SEARCH | TOOLS | HELP' and a 'Login' button. Below this is a 'Prepare Download' section. The main content area is divided into two panels. The left panel, titled 'AOR', contains a table of observations. The right panel, titled 'Coverage', shows a detailed view of an astronomical image with a grid overlay and various data points.

AOR ID	Target Name	NAIF ID	Instrument	Plan ID	Proposal PI	Abstract
<input type="checkbox"/> 75_0045_1	NGC7331_N0		FIFI-LS	75_0045	Dario Fadda	abstract
<input type="checkbox"/> 75_0045_12	NGC7331_C1-1-3		FIFI-LS	75_0045	Dario Fadda	abstract
<input type="checkbox"/> 75_0045_13	NGC7331_C2-1-2		FIFI-LS	75_0045	Dario Fadda	abstract
<input type="checkbox"/> 75_0045_2	NGC7331_S1		FIFI-LS	75_0045	Dario Fadda	abstract
<input type="checkbox"/> 75_0045_3	NGC7331_N2		FIFI-LS	75_0045	Dario Fadda	abstract
<input type="checkbox"/> 75_0045_4	NGC7331_S2		FIFI-LS	75_0045	Dario Fadda	abstract
<input type="checkbox"/> 75_0045_5	NGC7331_S3		FIFI-LS	75_0045	Dario Fadda	abstract
<input type="checkbox"/> 75_0045_6	NGC7331_S4		FIFI-LS	75_0045	Dario Fadda	abstract
<input type="checkbox"/> 75_0045_8	NGC7331_N1		FIFI-LS	75_0045	Dario Fadda	abstract

The 'Coverage' panel shows a detailed view of an astronomical image with a grid overlay. The image displays a bright, elongated source of light, likely a star or galaxy, with a grid of red and green lines overlaid. The grid is labeled with coordinates in J2000 equinox, such as '22h37m12.12s' and '+34d25m46.0s'. The image also shows various data points and labels, including 'Options: FITS HiPS Auto' and '2MASS K_s FOV: 4.2". The bottom of the panel displays the coordinates 'EQ-J2000: 22h37m13.84s, +34d25m46.0s' and the value 'Value: 486.253998 DN'. There is a 'Lock by click' checkbox at the bottom right.

Raw data only

To download raw data only:

- Select Level 1
- Observation type: OBJECT

Important:

Use a radius larger than that suggested (0.2 degs)

Select:

- Instrument: FIFI-LS
- Spectral element: Red (to download [CII] obs only)

The screenshot shows the SOFIA Search interface. At the top, there are navigation links for IRSA, DATA SETS, SEARCH, TOOLS, HELP, and a Login button. Below this is a header with the SOFIA logo and buttons for Search, Catalogs, Help, and Background Monitor. The main content area is titled "SOFIA Search" and includes a link for "Important notes on archive completeness".

Spatial Constraints Search for observations within a specified radius of a specified position. Enter search criteria below.

Object/Position Coordinates or Object Name:

Multiple Positions **NGC 7331** resolved by NED

Solar System Target 339.26672, 34.41552 Equ J2000 or 22h37m04.01s, +34d24m55.9s Equ J2000

Precovery

All-Sky Radius:

Valid range between: 0.000278 deg and 5 deg

Proposal Constraints

Observation Constraints

Instrument Constraints

FIFI-LS Configuration:

Spectral Element:

Wavelength (um): From: To:

Data Product Constraints

Processing Level: Level 0 Level 1 Level 2 Level 3 Level 4

Observation Type:



Dario Fadda
SOFIA data school - 16 Apr 2023



Select and download

Select all AORs.
Prepare download.

Skip the PNG to
download less stuff.

The screenshot displays the IRSA SOFIA data interface. At the top, there are navigation links for IRSA, DATA SETS, SEARCH, TOOLS, and HELP. Below this is a 'Prepare Download' section with a table of AORs. The table has columns for AOR ID, Mission ID, Target Name, NAIF ID, ra (deg), and dec (deg). A green arrow points to the 'Prepare Download' button. A 'Download Options' dialog box is open, showing a warning that the data is proprietary. The dialog includes fields for Title (SOFIA-0), Include File Type (PNG and Region), Zip File Structure (Structured (with folders)), Save as (FIFI-LS_Files Working...), and File Location (Send to background and Cancel). At the bottom, a progress bar shows 'FIFI-LS_Files-part1.zip' with 18s left and a download speed of 114 MB/sec. On the right, a 'Coverage' window shows a 2MASS color J, H, K image with a grid and various coordinates.

AOR ID	Mission ID	Target Name	NAIF ID	ra (deg)	dec (deg)
75_0045_13	2019-11-14_FI_F640	NGC7331_C2-1-2		339.3180000	34.4120000
		B1_C2-1-2		339.3180000	34.4120000
		B1_C2-1-2		339.3165000	34.4120000
		B1_C2-1-2		339.3165000	34.4120000
		B1_C2-1-2		339.3165000	34.4120000
		B1_C2-1-2		339.3165000	34.4120000
		B1_C2-1-2		339.3180000	34.4120000
		B1_C2-1-2		339.3165000	34.4110000
		B1_C2-1-2		339.3165000	34.4110000
		B1_C2-1-2		339.3180000	34.4110000
		B1_C2-1-2		339.3165000	34.4110000
		B1_C2-1-2		339.3180000	34.4120000
		B1_C2-1-2		339.3180000	34.4110000
		B1_C2-1-2		339.3165000	34.4090000
		B1_C2-1-2		339.3165000	34.4090000
		B1_C2-1-2		339.3180000	34.4120000
		B1_C2-1-2		339.3180000	34.4110000
		B1_N2		339.2585569	34.4430000
		B1_N1		339.2558882	34.4230000
		B1_C2-1-2		339.2596002	34.4100000
		B1_S4		339.2649007	34.3790000
		B1_C2-1-2		339.2821123	34.3860000
		B1_C2-1-2		339.2646980	34.3970000
		B1_C2-1-2		339.2630331	34.3970000
		B1_C2-1-2		339.2687876	34.4040000
		B1_C2-1-2		339.2762464	34.4040000
		B1_C2-1-2		339.2828943	34.4040000

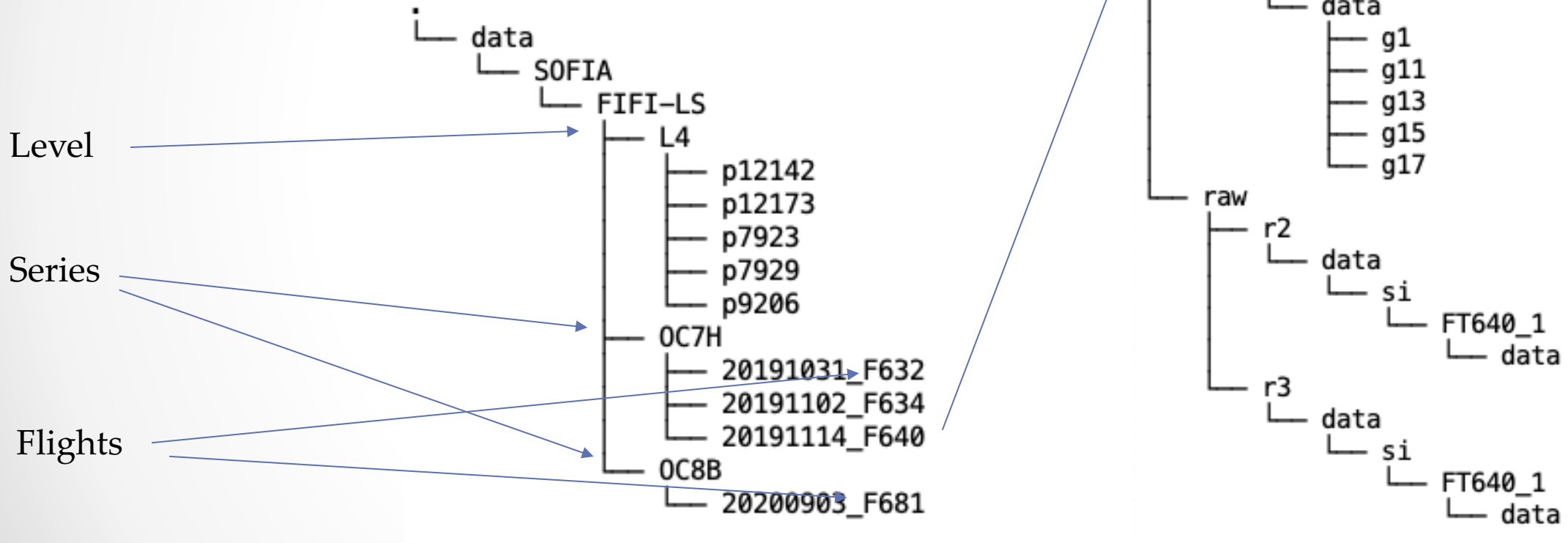


Dario Fadda
SOFIA data school - 16 Apr 2023



Exploring the data tree

In this case, the whole set of data is 3.0 GB. We are interested in the raw files only. It is possible to get all the files in one directory only, but then it is not clear which flights are they from.



Each flights contains processed and raw data directories. For our reduction we will consider only raw data.

Raw files

Raw files from the archive have some information in their names:

```
00415_072207_00013_NGC7331_S2_act4_-2_-5_A_lw.fits
00416_072245_00014_NGC7331_S2_act4_-2_-5_B_lw.fits
```

Progressive run number	UTC time acquisition	Observation number	Target name	Part of observation	Offsets Nod	Channel
---------------------------	-------------------------	-----------------------	----------------	------------------------	----------------	---------

Each file contains the readouts for a certain number of grating positions for one channel only.

For each telescope pointing there is a new file.

Why using the pipeline ?

The IRSA archive contains level 4 data ready to be used for science. Nevertheless, there are advantages to use the pipeline to reduce the raw data:

- The processing of data has been done in different times with different versions of the pipeline and of calibration (flats, flux)
- You can be interested in combining data of the same object from different series or bigger regions which are reduced in pieces
- You would like to selectively reject some files (which have high noise), or you want to resample the final data with a different smoothing parameter

Installing the pipeline

- Documentation: https://sofia-usra.github.io/sofia_redux/index.html
- Repository: https://github.com/SOFIA-USRA/sofia_redux
- Installation: https://github.com/SOFIA-USRA/sofia_redux/blob/main/docs/install.rst

Suggested installation:

- Install Anaconda Python: <https://www.anaconda.com/products/distribution>
- Download the environment file: https://raw.githubusercontent.com/SOFIA-USRA/sofia_redux/main/environment.yml
- Create and activate the pipeline environment:
`conda env create -f environment.yml`
`conda activate sofia_redux`
- Clone the repository:
`git clone https://github.com/SOFIA-USRA/sofia_redux.git`
- Install in edit mode:
`pip install -e sofia_redux`

Updating the pipeline

Before starting any new reduction, it is a good habit to update your local repository in the case there is a new version of the pipeline.

If you cloned the repository, just go inside it and pull the new version.

```
cd sofia_redux  
git pull
```

Since you installed the pipeline in 'edit' mode, the new pipeline will be immediately activated.

If you want to try the pipeline, you can opt for a direct conda installation:

```
conda install -c sofia-usra -c astropy sofia_redux
```

Contributing to the pipeline

The screenshot shows the GitHub repository page for SOFIA-USRA / sofia_redux. The repository is public and has 6 stars and 0 forks. The 'Issues' tab is selected, showing a search bar with the query 'is:issue is:closed'. There are 9 labels and 0 milestones. A green 'New issue' button is visible. Below the search bar, there are two closed issues:

Issue Title	Author	Projects	Milestones	Assignee	Sort
✓ Add direct link to raw environment file #2 by darioflute was closed on Jul 29, 2021					
✓ Suggestion to link environment.yml in the installation instructions #1 by darioflute was closed on Jul 29, 2021					

At the bottom of the issues list, there is a ProTip! indicating that issues updated in the last three days are [updated:>2022-04-22](#).

Any issue, comment, or proposal for new features can be written in the github page of sofia_redux. You will need a github account (you can open it for free). Once logged in, click on the 'Issues' tab. At this point, this window will appear showing open and closed issues. You can open a new issue by clicking on the green button.

Adding the library of atmospheric models

- The library of atmospheric models (ATRAN) should be downloaded separately through the URL:

https://sofia-downloads.s3-us-gov-west-1.amazonaws.com/atran_fifi-ls_wv.tgz

This can be unarchived in a directory which will be then referenced in the parameter file used for the reduction.

Explore data through a Python notebook

The raw files store all the ramps of each observation. It is possible to look into the details of the data before starting the reduction.

This can help individuating files to reject, especially those corrupted.

A companion Jupyter notebook can help with this exploration.

The notebook can be uploaded from:

<https://www.sofia.usra.edu/sites/default/files/2022-04/FIFI-LS.ipynb>

Running the pipeline

- The pipeline is written in sequential modules
- Each module has options
- Products from each step can be saved and inspected
- Some modules can be run in multi-threading to speed them up

The pipeline can be run via a GUI (redux).

See this presentation to know how:

<https://www.sofia.usra.edu/meetings-and-events/events/pipeline-week-introduction-sofia-processing-pipeline>

I will show how to run it through a command line by specifying options and modules to run in a configuration file.



Logo of the REDUX pipeline

Running the pipeline

After entering the `sofia_redux` environment:

```
conda activate sofia_redux
```

You will need a configuration file and the list of input files.
Once these are created, you can run the pipeline with:

```
redux_pipe -c redux_param.cfg redux_infiles.txt
```

Where `redux_param.cfg` is the configuration file and `redux_infiles.txt` contains the list of the input files.

The list of input files

To create a list of files, after unzipping the data, a directory called 'data' is created and in the same path one can create another directory to reduce the data.

First, one need to create a list of files:

```
mkdir reduced
```

```
cd reduced
```

```
find ../data -name '*lw.fits' > redux_infiles.txt
```

Then, one needs a configuration file with all the parameters needed for the reduction: `redux_param.cfg`

Pipeline configuration: basic modules

Let's examine the parameters contained in the `redux_param.cfg` file

```
Redux parameters for FIFI-LS instrument in IFS mode
# Pipeline: FIFI_LS_REDUX v2_6_1_dev0
recipe = checkhead, split_grating_and_chop, fit_ramps
[1: checkhead]
  abort = True
[2: split_grating_and_chop]
  save = False
[3: fit_ramps]
  save = False
  parallel = True
  s2n = 10.0
  thresh = 5.0
  badpix_file = ""
  remove_first = True
  subtract_bias = True
  indpos_sigma = 3.0
[4: subtract_chops]
  save = False
[5: combine_nods]
  save = False
  b_nod_method = nearest
  offbeam = False
```

List of module to run (optional)

Checking headers

Separate grating and chop positions

Slope of the ramps

Subtract sky from source signal

Combine the two nods

Pipeline configuration: calibration

```
[6: lambda_calibrate]
    save = False
[7: spatial_calibrate]
    save = False
    rotate = True
    flipsign = default
[8: apply_static_flat]
    save = False
    skip_flat = False
    skip_err = True
[9: combine_grating_scans]
    save = False
    bias = True
```

Apply wavelength calibration

Add astrometry

Apply NE orientation

Apply flats

Do not apply flat error to error propagation

Combine different grating scans in a spectrum

A part deciding to save the files, the only parameter we should mention is the possibility to rotate the field in the usual NE direction or to leave it along the array coordinates. This is done with rotate.

Pipeline conf: tellurics & flux calibration

```
[10: telluric_correct]
    save = False
    skip_tell = False
    atran_dir = ""
    cutoff = 0.6
    use_wv = False
[11: flux_calibrate]
    save = False
    skip_cal = False
    response_file = ""
[12: correct_wave_shift]
    save = True
    skip_shift = False
```

Atmospheric correction

Path to ATRAN models

Cutoff in atmospheric transmission

Calibration in flux

Barycentric correction

Atmospheric correction with 'use_wv' = True is now possible for all the data, since we archived values of the precipitated water vapor for all the observations.

The cutoff is the level under which data should not be considered since the atmospheric transmission is too low. It is safe to put it at 0.6.

Pipeline: cube resampling

The resampling step put all the information together into a spectral cube with 2 spatial and 1 spectral dimensions.

```
[13: resample]
  save = True
  parallel = True
  max_cores = ""
  check_memory = True
  skip_coadd = False
  interpolate = False
  error_weighting = True
  fitthresh = -1
  postthresh = -1
  negthresh = -1
  append_weights = False
  skip_uncorrected = False
  scan_reduction = False
  save_scan = False
  scan_options = ""
  detector_coordinates = False
  xy_oversample = 5.0
  xy_pixel_size = 3.0
  xy_order = 2
  xy_window = 2.0
  xy_smoothing = 0.5
  xy_edge_threshold = 0.5
  adaptive_algorithm = none
  w_oversample = 8.0
  w_pixel_size = ""
  w_order = 2
  w_window = 0.8
  w_smoothing = 0.3
  w_edge_threshold = 0.5
```

Parameters for multi-threading

Fluxes weighted by variance

Rejection not activated

Parameters for OTF mapping

Oversampling

Output pixel size (arcsec)

Order of polynomial fit

Size of window and kernel for spatial smoothing

Size for limiting border effects

Options: none, scaled, shaped

Same meanings for the spectral dimension

general

spatial

spectral

Running the pipeline

Once all the parameters are chosen and the list of files is defined, once in the sofia_redux environments, one can launch the reduction with the command:

```
redux_pipe -c redux_param.cfg redux_infiles.txt
```

If all the files are usable, the pipeline will run through all the modules and write the final spectral cube as well as any intermediate products which were required.

We saved the single frames before resampling them into the cube to show where they are taken.

Notes about resampling

The resampling involves some choices.

Since the data taken are sparse, the smoothing allows one to have a continuous cube across space and wavelengths.

Spatially, the size of FIFI-LS pixels are not much smaller than the PSF. This means that dithering is used to recover the PSF.

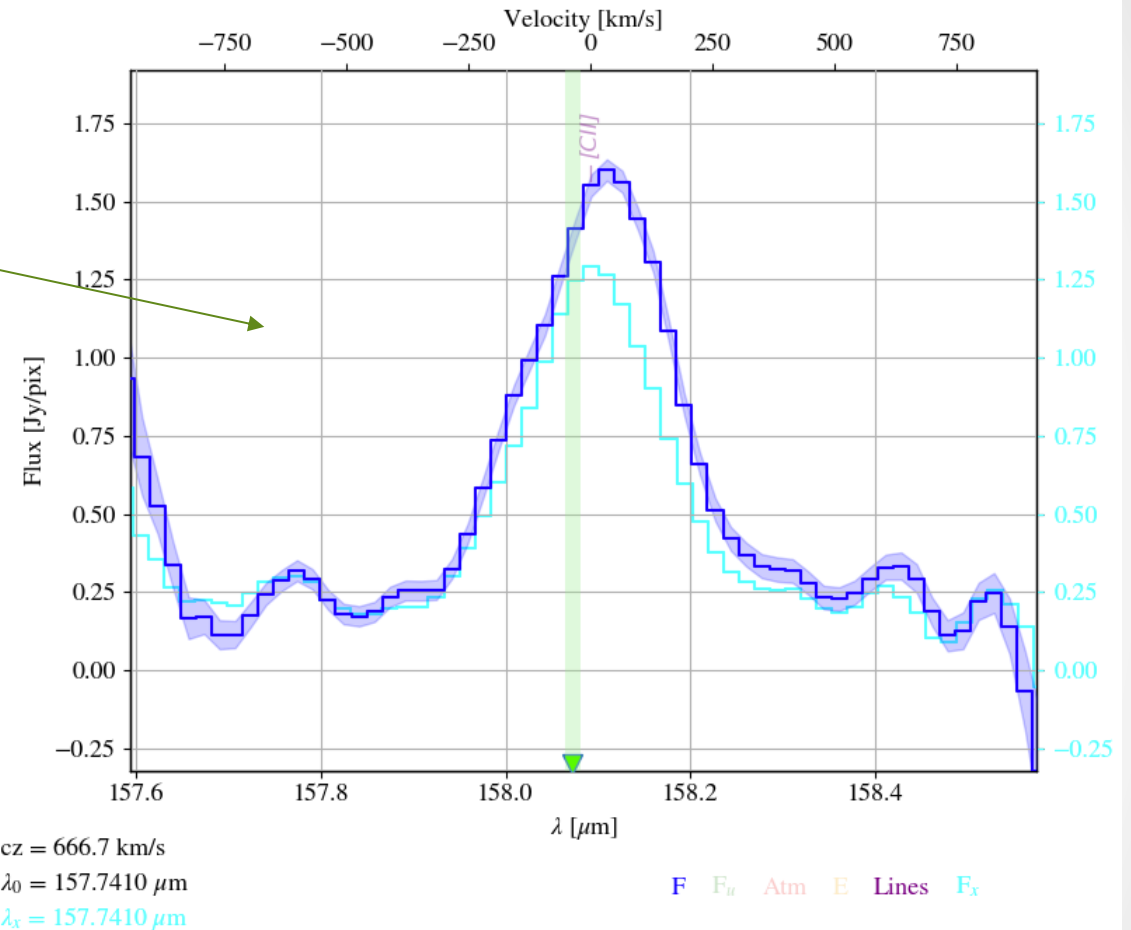
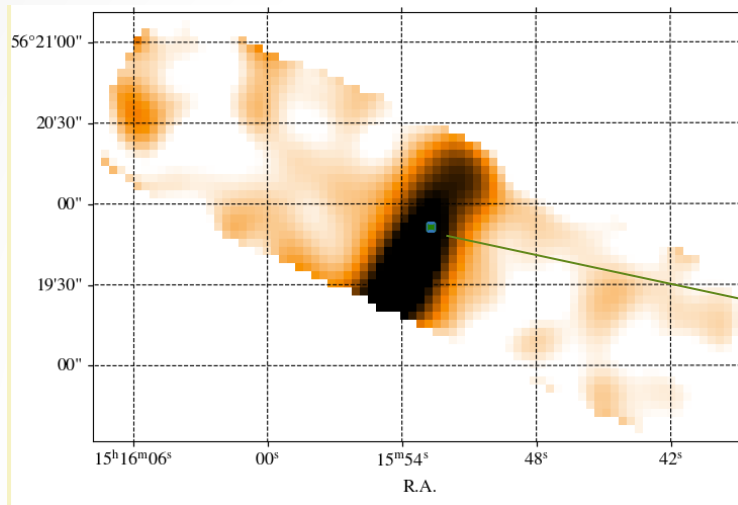
Clearly, over-smoothing will return a cube with broad point sources.

The peak flux on the top of the source will be depressed.

On the other hand, too little smoothing returns a noisy cube.

If the field contains diffuse emission and point sources, a solution is applying the 'adaptive algorithm' where the kernel is not fixed all over the cube but varies according to the flux. This is not used for the archival products.

Effects of over-smoothing



The danger of over-smoothing peaks of emission is clear along the center of the edge galaxy NGC 5907.

By choosing a kernel with a width corresponding to the FWHM of the spatial PSF, there is a loss of flux in the central ridge of the galaxy. This means that apertures corresponding to the PSF will contain less flux than what observed.

The choice of an adaptive kernel in this case leads to a smoothing very close to the blue curve, which has smoothing = 0.5 x FWHM.

Blue: xy_smoothing 0.5. (conserves spatial resolution)
Cyan: xy_smoothing 1.0 (oversmoothed)

Pipeline products

Several files are produced by the pipeline.

Each file has a different label indicating its type. Some useful files are listed here:

- *_RP[01]_*.fits results of ramp fitting (on and off positions)
- *_WAV_*.fits wavelength calibrated files
- *_XYC_*.fits spatially calibrated files
- *_FLF_*.fits flat-fielded files
- *_SCM_*.fits combined scans files
- *_CAL_*.fits flux calibrated files
- *_WSH_*.fits barycentric correction applied to wavelengths
- *_WXY_*.fits final cube resampled

WXY files contain several extensions. E.g., they contain also the cubes before applying the atmospheric corrections.

Exploring the data

I developed two codes to explore the data.

The codes are in Python and use the anaconda distribution.

To use them, you can create a conda environment and install them.

```
conda create -n fifils python=3.10
```

```
conda activate fifils
```

```
conda install -c conda-forge lmfit reproject fitsio
```

```
conda install -c darioflute sospex fifipy
```

I created versions for linux, windows, and the two flavors of mac-osx with the latest version of anaconda which uses Python 3.10

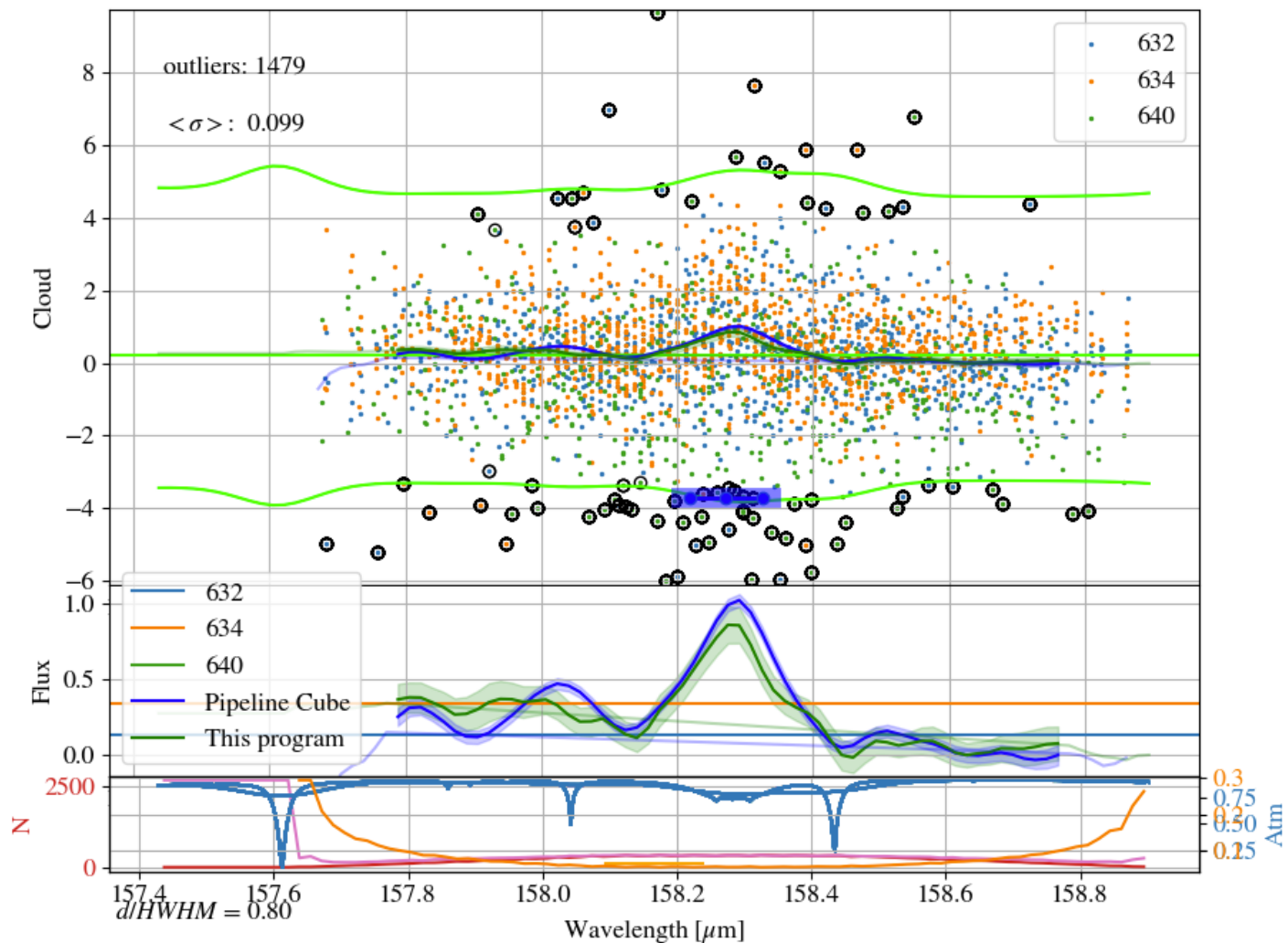
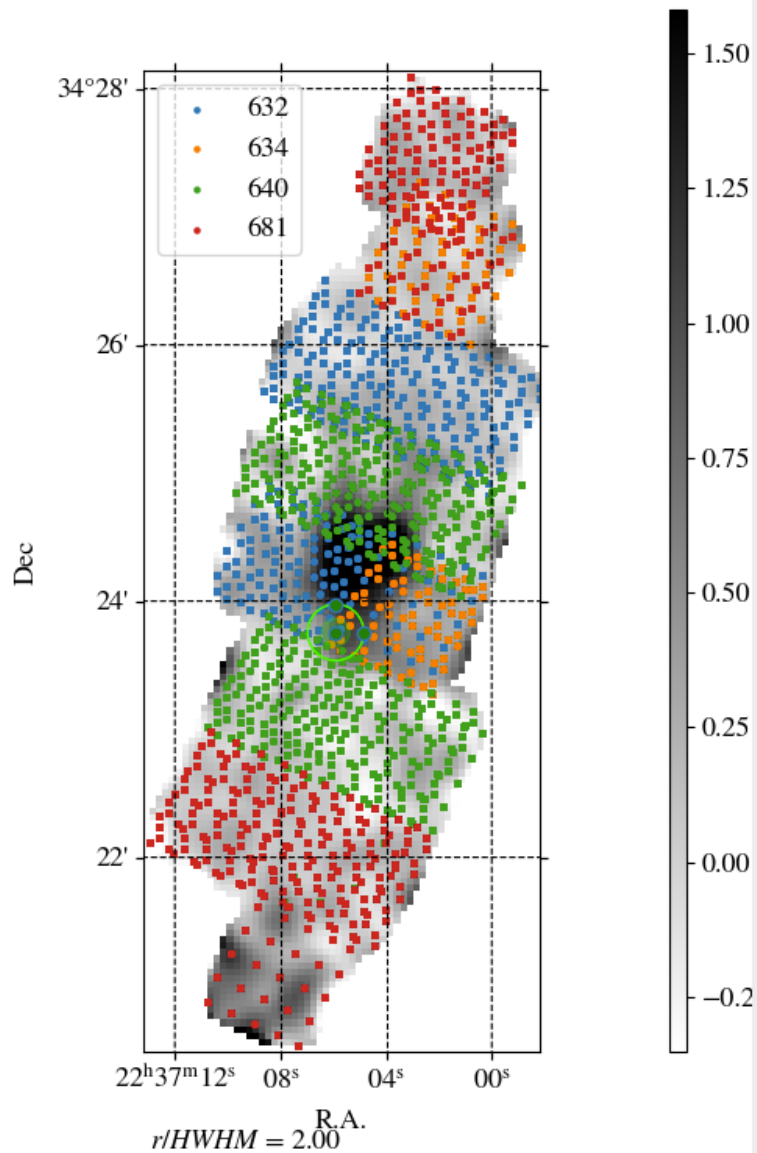
Explore the resampling with cubik

To explore the data used to obtain the resample cube and to have a better sense of the kernels used for the resampling, one can use cubik which comes with the library fifipy.

Lunch from a terminal: cubik

Then select the *WXY* file, the spectral cube. To work, the files *WSH* should be present in the same directory.

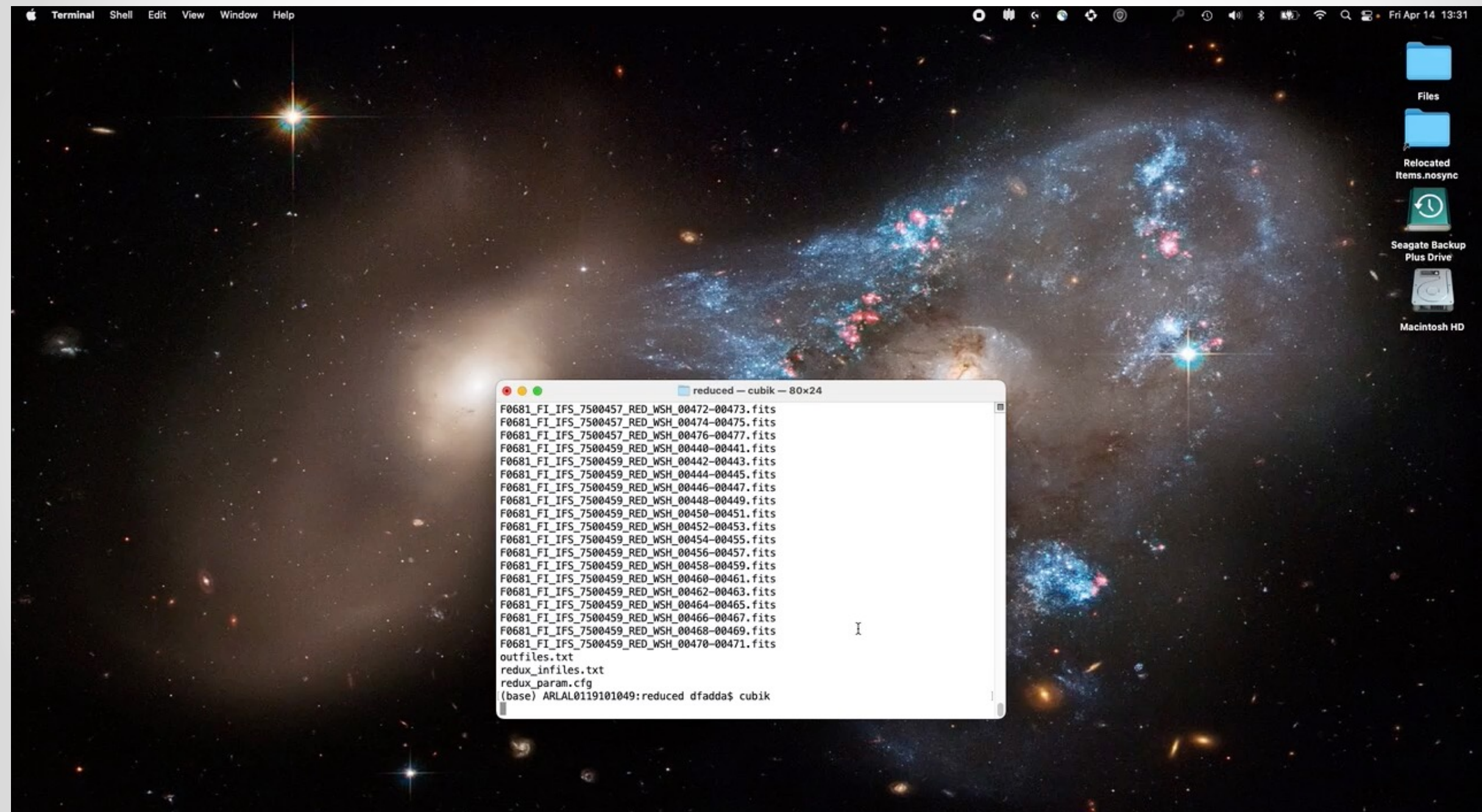
The code will show the spatial and spectral kernels used to resample the data as well as the cloud of points used for the resample which are included in the circular aperture on the left panel. The pipeline spectrum is shown overlapped on the cubik resampled data. Cubik can create a new cube or simply a new error cube based on the distribution of the data.



Cubik: the movie



Dario Fadda
SOFIA data school - 16 Apr 2023



```
reduced - cubik - 80x24
F0681_FI_IFS_7500457_RED_WSH_00472-00473.fits
F0681_FI_IFS_7500457_RED_WSH_00474-00475.fits
F0681_FI_IFS_7500457_RED_WSH_00476-00477.fits
F0681_FI_IFS_7500459_RED_WSH_00440-00441.fits
F0681_FI_IFS_7500459_RED_WSH_00442-00443.fits
F0681_FI_IFS_7500459_RED_WSH_00444-00445.fits
F0681_FI_IFS_7500459_RED_WSH_00446-00447.fits
F0681_FI_IFS_7500459_RED_WSH_00448-00449.fits
F0681_FI_IFS_7500459_RED_WSH_00450-00451.fits
F0681_FI_IFS_7500459_RED_WSH_00452-00453.fits
F0681_FI_IFS_7500459_RED_WSH_00454-00455.fits
F0681_FI_IFS_7500459_RED_WSH_00456-00457.fits
F0681_FI_IFS_7500459_RED_WSH_00458-00459.fits
F0681_FI_IFS_7500459_RED_WSH_00460-00461.fits
F0681_FI_IFS_7500459_RED_WSH_00462-00463.fits
F0681_FI_IFS_7500459_RED_WSH_00464-00465.fits
F0681_FI_IFS_7500459_RED_WSH_00466-00467.fits
F0681_FI_IFS_7500459_RED_WSH_00468-00469.fits
F0681_FI_IFS_7500459_RED_WSH_00470-00471.fits
outfiles.txt
redux_infiles.txt
redux_param.cfg
(base) ARLAL0119101049:reduced dfadda$ cubik
```

Analyze the cube with sospex

Spectral cubes can be easily analyzed using sospex, a Python GUI which can be installed with anaconda:

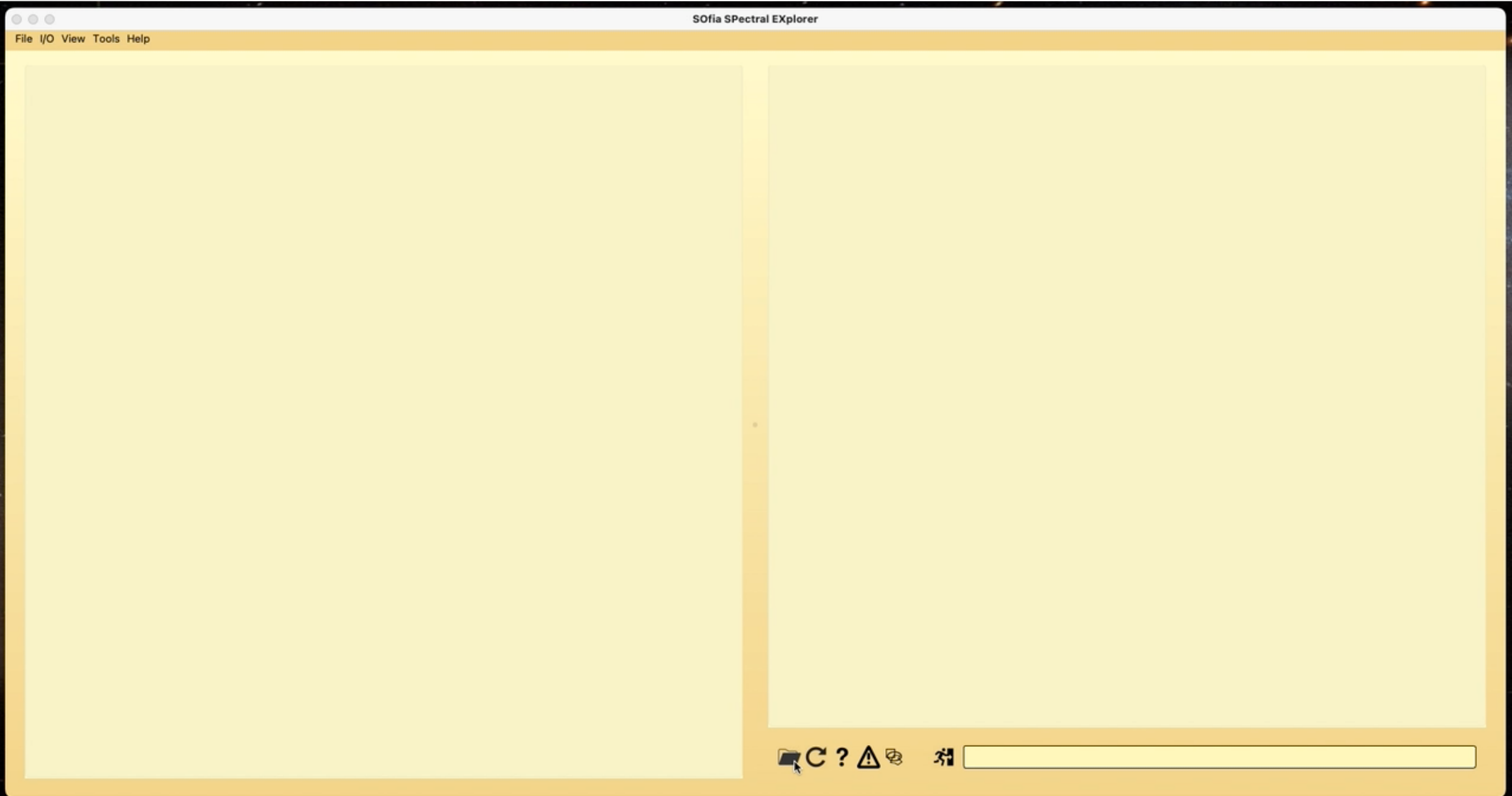
- <https://github.com/darioflute/sospex/blob/master/INSTALL.md>

The github repository is:

- <https://github.com/darioflute/sospex>

The GUI allows one to explore the cube produced by the pipeline and to do scientific analysis of the data (extract apertures, fit one or more lines, fit the entire cube, compare with other images/spectral cubes).

sospex: the movie



Dario Fadda
SOFIA data school - 16 Apr 2023

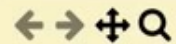
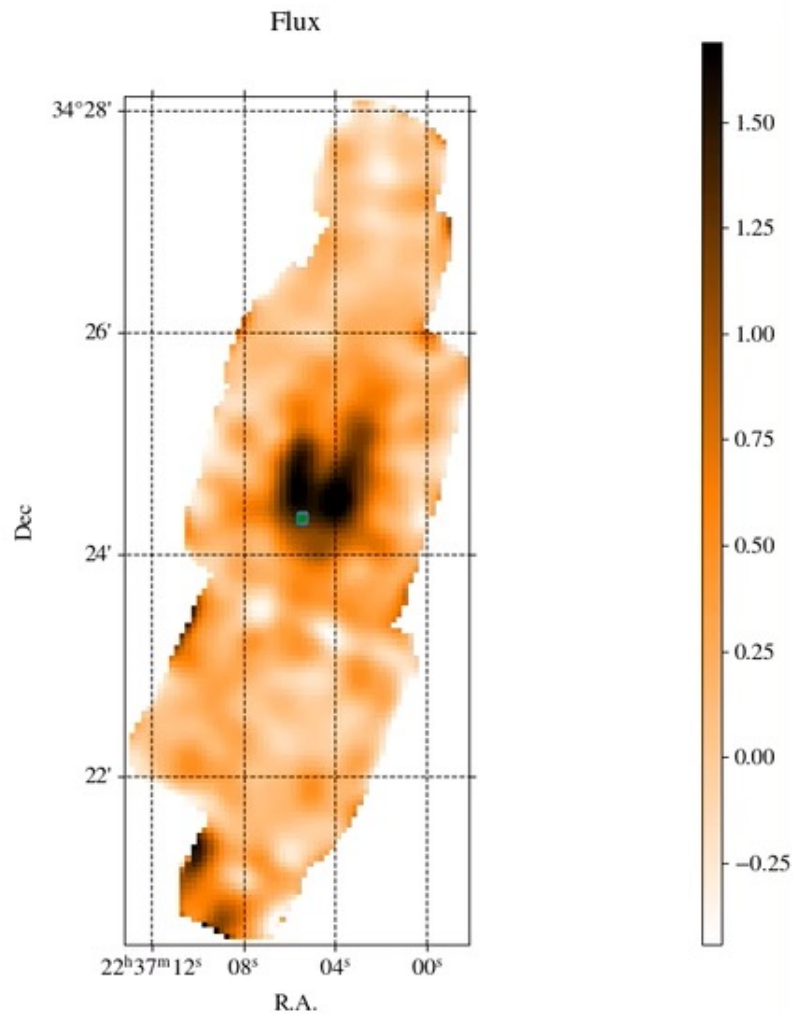


sospex: moments

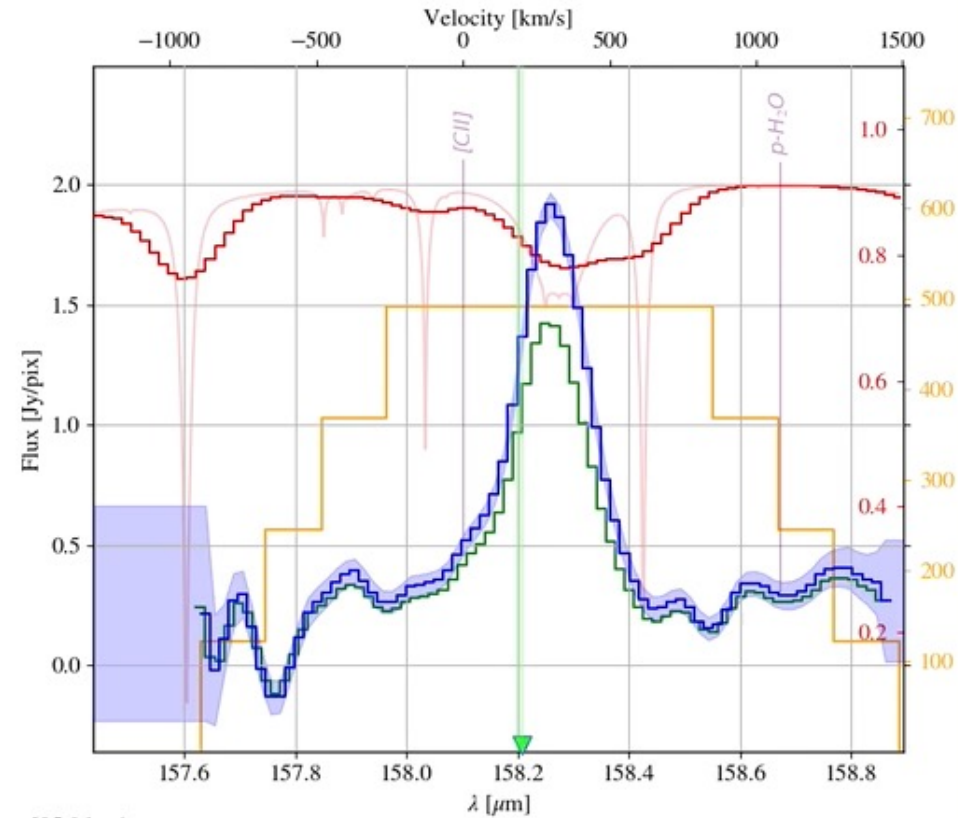
Sofia Spectral Explorer [NGC7331_N0 - FIFI-LS - Obs: FIFI-LS team - File: F0632_FI_IFS_7500451_RED_WXY_00028-100325.fits]

File I/O View Tools Help

F × F_u × E

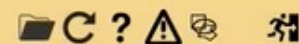
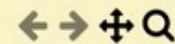


All Pix

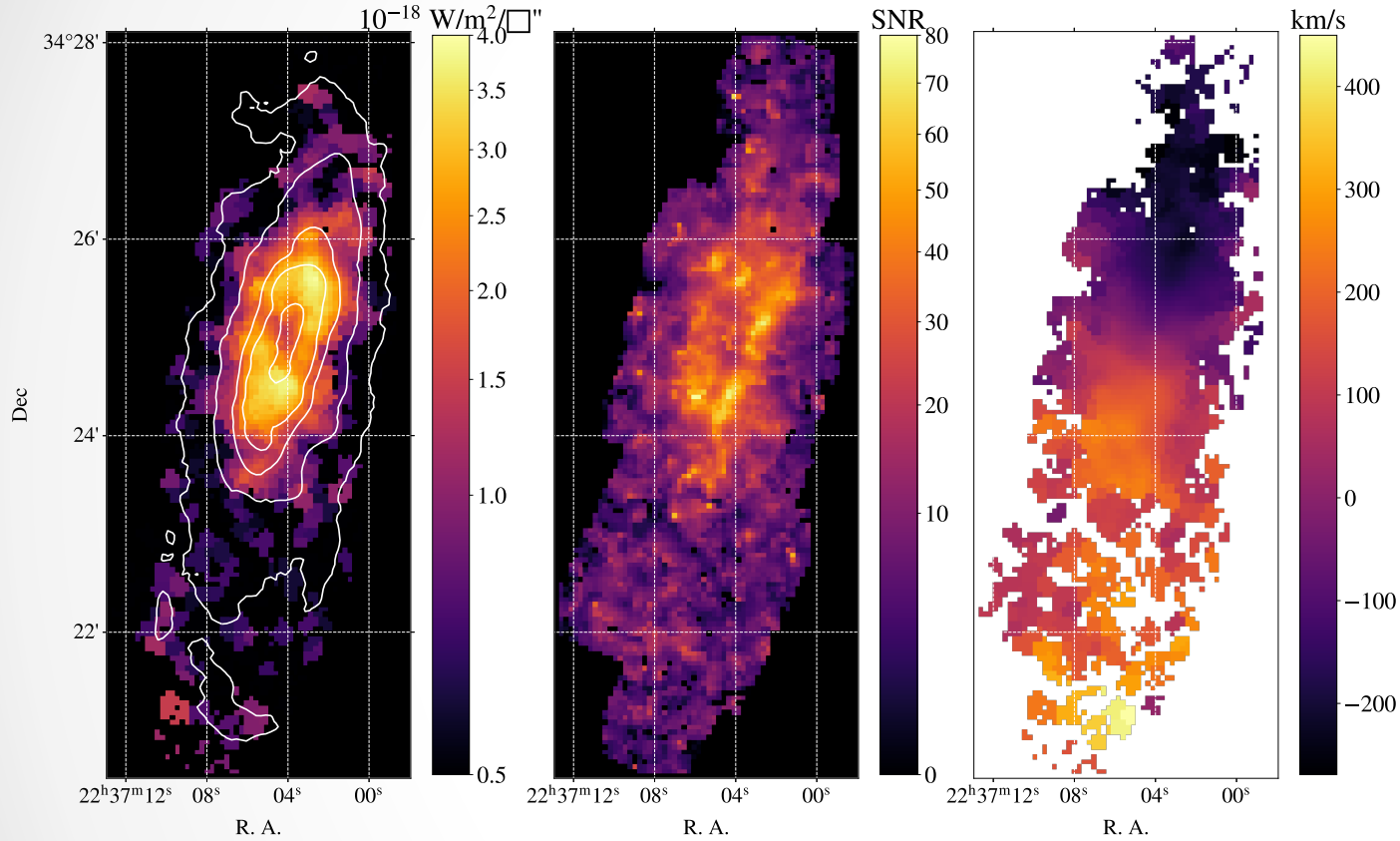


cz = 685.0 km/s
 $\lambda_0 = 157.7410 \mu\text{m}$

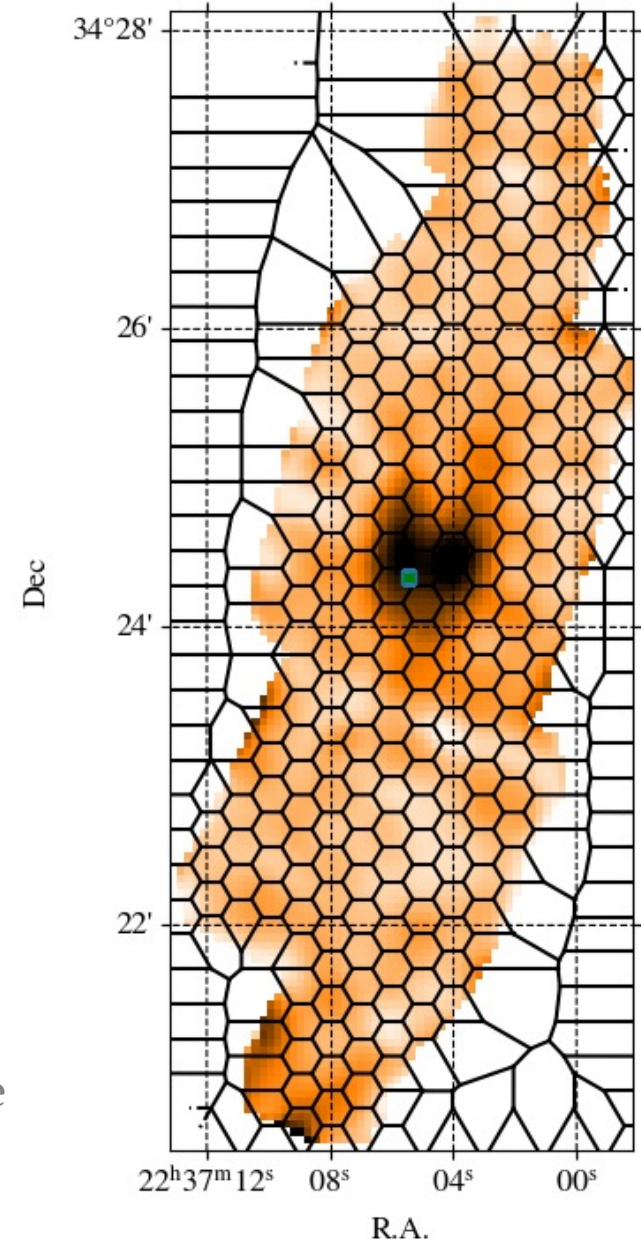
F F_u Atm E Lines



Integrated flux



Better maps are obtained by specifying the initial conditions for the line fitting in many cells, as in this example of NGC 7331.



More about sospex

sospex allows one to make even more things, such as analyze Herschel/PACS cubes and spectral cubes from other instruments (i.e. ALMA).

It is possible to directly compare two spectral cubes (using the velocities in the x-axis for direct comparison of two lines).

Please, contact me if you have troubles using it, you want to leave a comment/suggestion, or you find bugs by submitting a ticket on github: <https://github.com/darioflute/sospex/issues>

The recent version (1.0.2) has been released to use Python 3.10 and the new version of matplotlib (3.7) which introduced many modifications.