FIFI-LS Pipeline Tutorial

This tutorial is designed to help guide users through the steps of processing FIFI-LS data, from raw data files (Level 1) to the final, calibrated data product. The information in this tutorial can also be found in more detail FIFI-LS User's Manual.

Retrieve the Data

For this sample reduction, we'll use [CII] observations of M82. These data were obtained as part of the FIFI-LS team's Guaranteed Time Observations in 2016. To retrieve the raw data that we will process from the SOFIA Science Archive at IRSA, either:

1. Enter the following criteria in the main SOFIA search page (https://irsa.ipac.caltech.edu/applications/sofia):

   Spatial Constraints → enter M82 as the object name, and increase the Radius to 10 arcminutes
   Observation Constraints → enter 2016-02-25_FI_F280 as the Mission ID
   Data Product Constraints → Select Level 1 only (de-select Level 3 and Level 4)

or

2. Follow this link:
https://irsa.ipac.caltech.edu/applications/sofia/?api=search&spatialConstraints=position&WorldPt=148.96969;69.67938;EQ_J2000&sr=10m&execute=true&processingLevel=LEVEL_1&planId=70_0408

On the search results page, click on the FIFI-LS tab. If you used the link above to search for the data, then click the down-arrow in the Mission ID column, select the check box next to Mission ID 2016-02-25_FI_F280, and click 'filter.' Each Mission ID corresponds to a specific flight.

We now have raw FIFI-LS data for M82 from Flight 280, but at two wavelengths. To narrow the results down to the [CII] data only, enter '>155' in the Min Wavelength column and then press return to filter out the short wavelength observations. This selection is made because only Level 1 files from the same spectral channel can be processed together. You should now have 191 FITS files in your search results. Select all of these files by selecting the box next to the AOR ID column header.

Click on Prepare Download to open a download dialog box, and then select Prepare Download within the dialog box. After a couple of minutes, you will be able to download a set of 5 zip files, containing the 191 FITS files (total size ~ 1.2 GB). Save these files wherever you wish (either individually by clicking on them, or by using the Download Script provided by IRSA), unzip them (if not using the IRSA Download Script), and navigate the subfolders to locate the 191 FIFI-LS FITS files. You may either leave these files in the directory or move them to another location of your

choice; the folder containing the 191 FITS files will be referred to as the 'input' directory for the remainder of the tutorial.
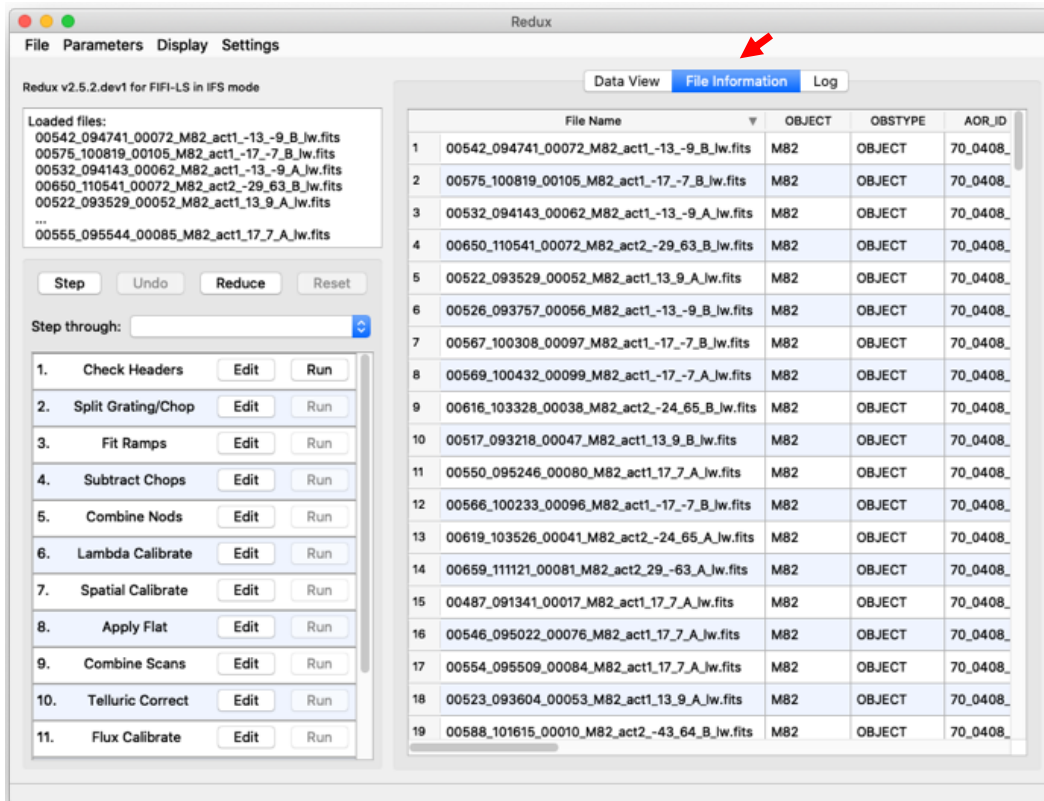
Download and Install the Pipeline
Instructions on how to download and install the *sofia_redux* pipeline package can be found at https://github.com/SOFIA-USRA/sofia_redux. If you need assistance, you can contact us through GitHub or the SOFIA Helpdesk (sofia_help@sofia.usra.edu).

Start the GUI
The pipeline can be executed in either Automatic Mode (on the command line) or in Manual Mode (through a GUI). In this tutorial, we will focus on running the pipeline through the GUI. To begin, launch the GUI by opening a terminal and typing:
 > redux
In the Redux GUI, load the M82 data by selecting File → Open New Reduction, and navigating to the input directory where you saved the data, select all 191 FITS files by clicking on the first file and shift-clicking on the last file, and then click Open. After clicking Open, the software reads the FITS headers and extracts relevant information, such as the instrument name for determining which pipeline the software will call (note that 'Redux [version] for FIFI-LS in IFS mode' is now displayed above the list of loaded files). Some additional FITS header information is shown on the File Information on the right side of the GUI.

When the pipeline runs, it is possible to automatically display the intermediate data using SAO DS9 if the connection between the pipeline/python and DS9 is set up properly (see http://ds9.si.edu/ for download and installation instructions; the ds9 executable must be

available in the PATH environment variable for the pyds9 interface to be able to find and control it). Alternatively, the automatic display can be disabled (each pipeline module can optionally produce FITS file output, in case a different FITS viewer is preferred). Options for displaying images in DS9 are found under Display Settings on the Data View tab.
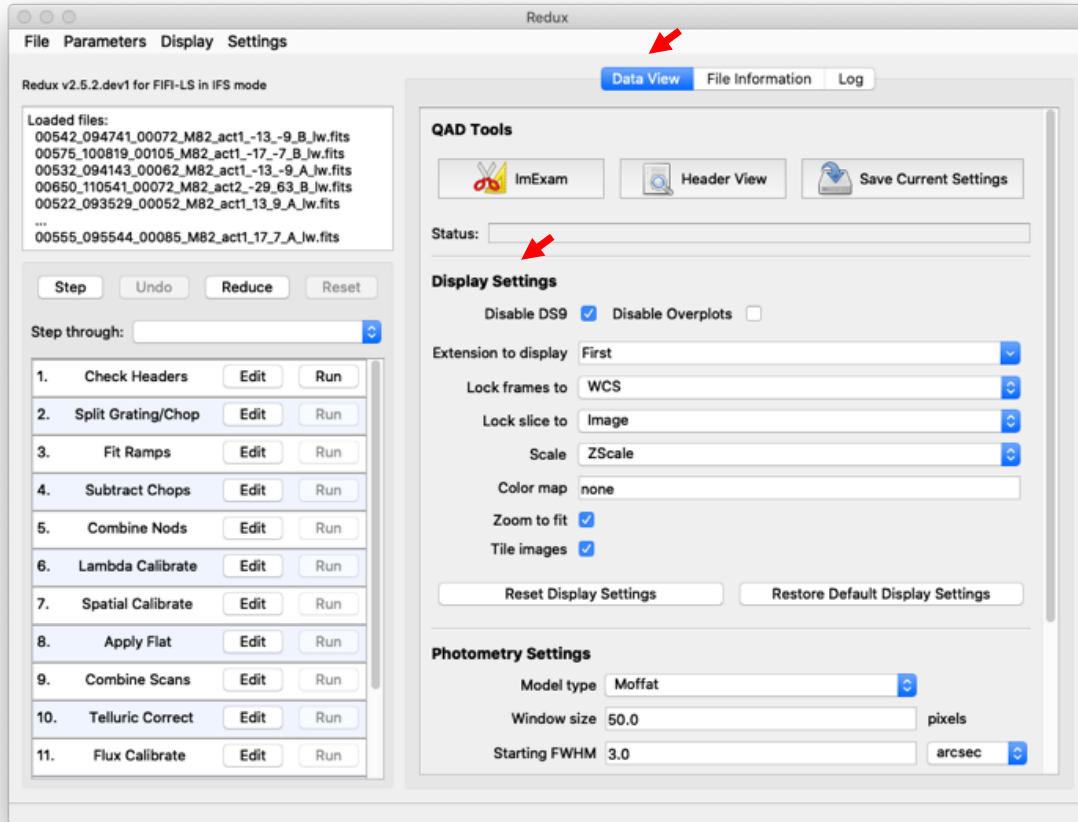

Run the Pipeline

There are 14 modules in the FIFI-LS pipeline, shown in the bottom-left panel of the GUI. Each pipeline module can be configured by clicking the 'Edit' button for that module. The entire pipeline can be executed by clicking the 'Reduce' button. Alternatively, the modules can be run individually by clicking on the 'Run' button for that module, or in a small group by selecting an end point in the 'Step Through' dropdown menu, then clicking 'Step.' In this tutorial, we'll run the modules one by one and examine their output.

1. Check Headers: This module examines the fits headers to make sure that they have the keywords required by the pipeline.

Options: Click on the Edit button next to Check Header module. A popup window for Edit Parameters for Check Headers will open up. In the popup window if the 'Abort reduction for invalid headers' option is not selected, the pipeline may fail in unexpected ways (as always: garbage in, garbage out!).
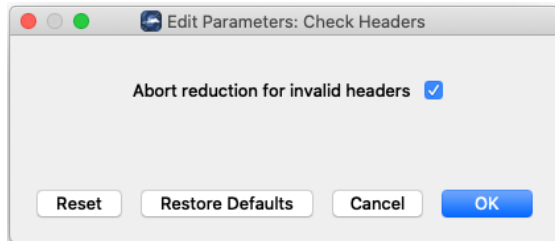
Output: This module creates no output; it allows the data set to pass through to the next modules in the reduction.

2. Split Grating/Chop: In NMC mode (nod-match-chop mode; information stored in the NODSTYLE keyword in the FITS header), the FIFI-LS raw data format contains data from both chop positions and data from each grating scan within a single file. This pipeline module separates each chop position into its own file, with each grating scan as an extension within the file.

Options and Output: The only option for this module is whether to save the output. For this example, saving the output would result in 191*2=382 files. Before running this module, you may want to make sure the 'Disable DS9' box is selected (On the Data View tab, in the Display Settings section); otherwise, the 382 files will all be opened in DS9.
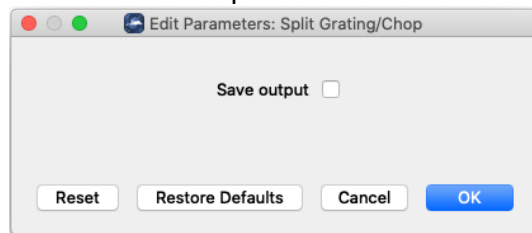

3. Fit Ramps: The flux measured in each pixel is reconstructed from the readout frames by fitting a line to the voltage ramps. The slope of the line corresponds to the flux value. This module converts the data from the raw data format into a 25 (detector pixels) x 16 (wavelength bins) flux array.

Options: The options in this module help identify and remove ramps that should be discarded. The default values are usually adequate, but they can be adjusted if too many (or too few) ramps are discarded. A default bad pixel mask is automatically applied, but a different one may be specified.
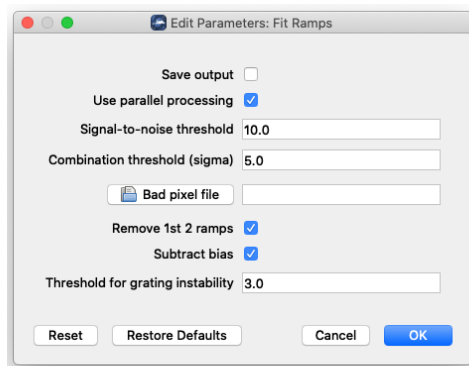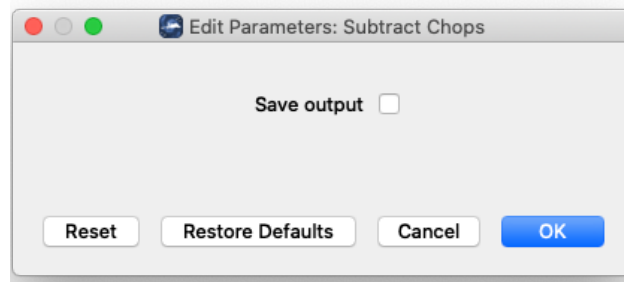
Output: The output of this module is not saved by default, but one sample image is shown below. If the output for this module is saved, it produces 382 FITS files.
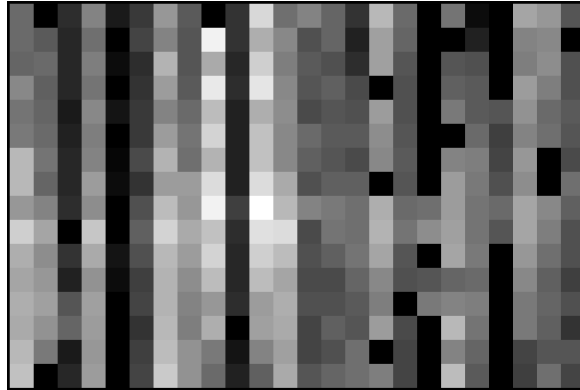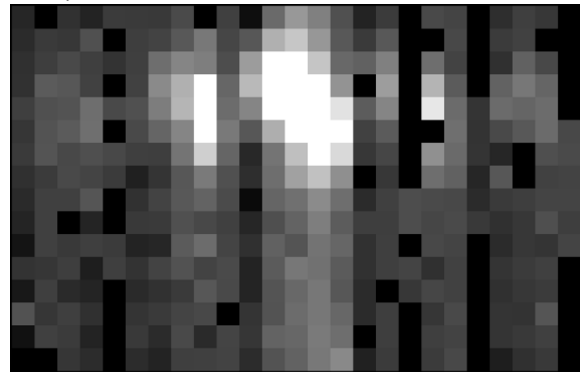

First extension of F0280_FI_IFS_7004081_RED_RP0_100471.fits. The spectral dimension is along the y-axis.

4. Subtract Chops: This module removes the instrument and sky background emission by subtracting the off-source chop data from the on-source chop data.
Options: The only option for this module is to save the output (off by default).


Output: The output of this module is the on-off chop subtracted flux image. A sample is shown below. Since this module combines the on- and off-source data (previously stored in separate files), the FITS output, if saved, results in 191 files.
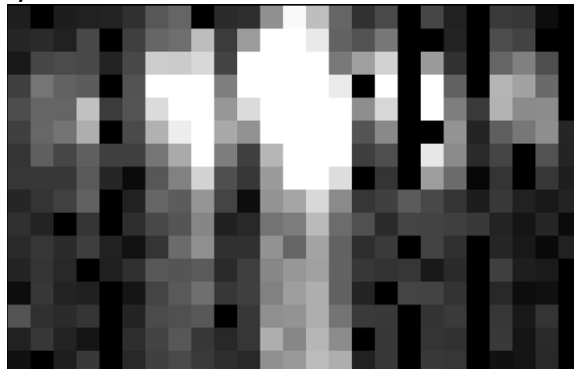

First extension of F0280_FI_IFS_7004081_RED_CSB_100471.fits. The spectral dimension is along the y-axis.

5. Combine Nods: This module, similarly to the Subtract Chops module, removes background emission from the data.

Options: The 'Method for combining off-beam images' option is used for data taken in C2NC2 or OTF mode, and is not relevant for this example. The 'Propagate off-beam image instead of on-beam' option can be used to make the map at the off position instead of the on position. This is mostly used for testing, so we will leave it unchecked.
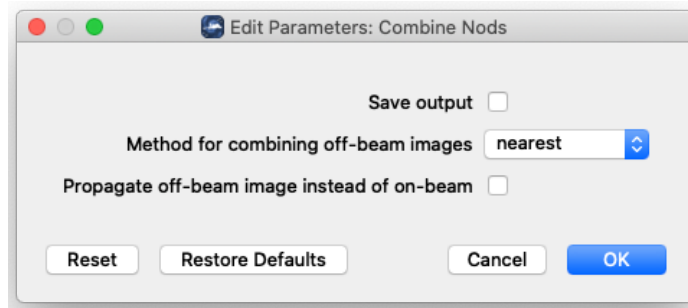

Output: Because this module is combining nods that are in different files, the number of output files (if saved) is reduced by a factor of two to 95.
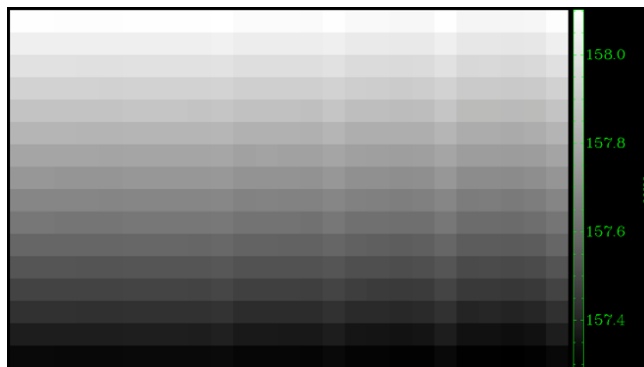

First extension of F0280_FI_IFS_7004081_RED_NCM_100471-100472.fits. The spectral dimension is along the y-axis.

6. Lambda Calibrate: During this module, the pipeline calculates the observed wavelength of each pixel based on the grating position at the time of the observations, an optical model of the instrument, and measurements of the positions of known spectral lines.
Options: The only option for this module is to save the output (off by default).
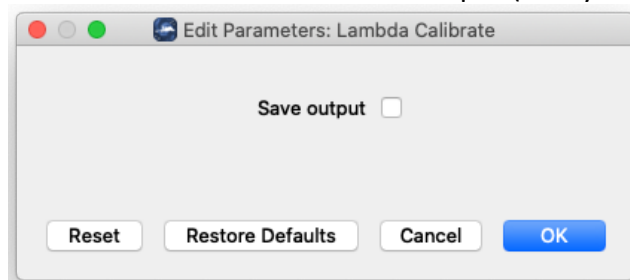

Output: The output of this module is an additional extension in the file containing a wavelength value for each pixel.

Sixth extension of F0280_FI_IFS_7004081_RED_WAV_100471-100472.fits. This shows the wavelength at each pixel.

7. Spatial Calibrate: This module calculates the sky position of each pixel at each wavelength, taking into account the nonuniform distribution of the pixels across the detector, the wavelength being observed, and any dithers used during the observations.

Options: Default settings for this module are to select 'Rotate by detector angle' and 'default' for 'Flip RA/DEC sign convention. This combination of settings will result in a final data product with the standard convention of North up and East to the left. If the final data product shows the wrong convention, it can be corrected in this module.
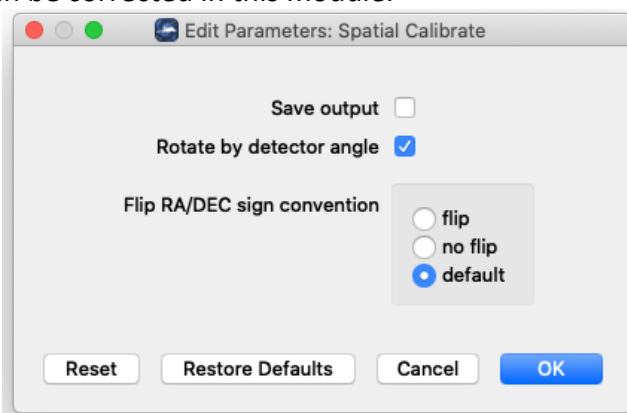

Output: The calculated x and y coordinates for each pixel are stored in two 25-element arrays for each grating scan (XS_G$_i$ and YS_G$_i$ for each grating scan i). If the fits file output is saved, the files will have XYC in their file names.

8. Apply Flat: This module flat-fields the data. The pipeline automatically calculates the values to use based on the detector used in the observations (red channel or blue channel), as well as the wavelength range of the data.

Options: This module can be skipped, though that is not recommended. There is also an option to 'Skip flat error propagation'; this option is on by default, and includes the systemic error in the CALERR keyword instead of propagating it into the flux error plane in the final data (if this option is de-selected, the CALERR keyword is still propagated, and the systemic error is included in the flux error plane).
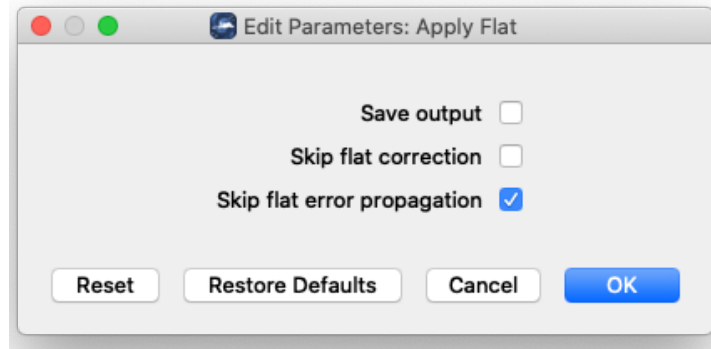
Output: Module 8 outputs 95 FITS files (if the 'Save output' option is selected).
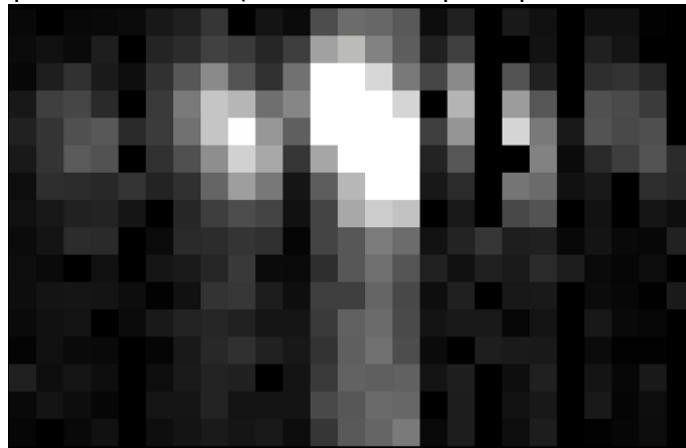
First extension of F0280_FI_IFS_7004081_RED_FLF_100471-100472.fits. A careful inspection shows differences between this image and the output image shown for Module 5.

9. Combine Scans: This module combines the grating scans within each of the input files.
Options: Aside from saving the output, the only other option for this module is to 'Correct bias offset.' This option corrects for any bias offset between scans as measured in the overlapping wavelength regions.
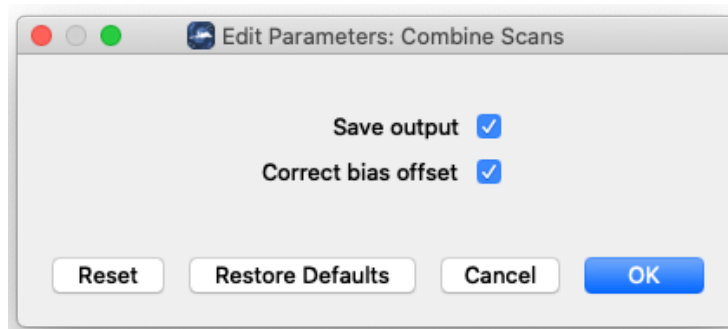
Output: Here is example output from this module. There are now more pixels in the spectral (y) direction because different grating positions have been combined.
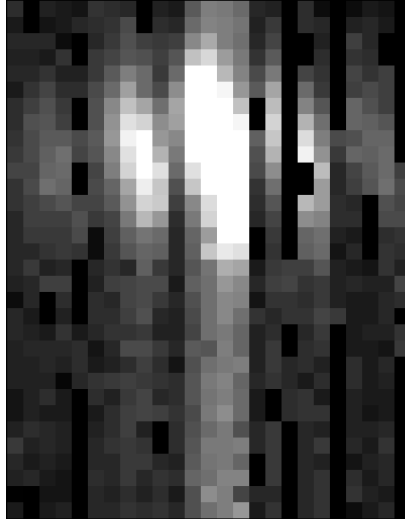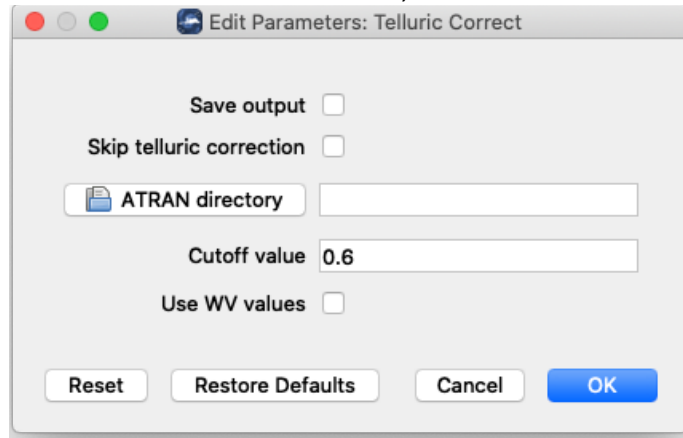
First extension of F0280_FI_IFS_7004081_RED_SCM_100471-100472.fits. Because two different gratings have been combined in this module, there are now more pixels in the spectral dimension along the y-axis.

10. Telluric Correct: This module corrects for the opacity of the atmosphere using ATRAN models of the atmosphere (courtesy of Steve Lord: https://atran.arc.nasa.gov/cgi-bin/atran/atran.cgi). Due to size constraints in the git repository, the default pipeline installation contains only a single atmospheric model. It's very likely that this file will be insufficient to properly correct for the atmosphere, so you will need to download a library of ATRAN files for a range of flight altitude, telescope elevation, and water vapor. There are two libraries available for download on the https://www.sofia.usra.edu/science/data/data-pipelines page: standard models, which use an assumed water vapor column based on flight altitude and telescope elevation, and specific models, which are calculated for every combination over a range of zenith water vapor values, telescope elevations, and flight altitudes. In this tutorial, we will use the standard models (1) because we have no measurements of water vapor during the observation, and (2) to save download time and disk space. Take the time now to download the standard models, and note their destination directory (if you have trouble finding or downloading the files, please contact the SOFIA Helpdesk at sofia_help@sofia.usra.edu).

In recent FIFI-LS observations, the water vapor in the atmosphere is measured by FIFI-LS at the beginning and end of each flight leg, and when the observatory changes altitude. These values are propagated into the FITS headers via the WVZ_OBS keyword and can optionally be used to correct the data using the specific models described above.

Options: The output of this module is not saved by default, and the correction can be skipped if desired (this is not recommended). To use the standard atmosphere models, click on the 'ATRAN directory' button and browse to the directory containing the models you downloaded. The 'Cutoff value' is an atmospheric transmission value that used to mask spectral regions with low transmission. This value can be lowered if too many regions are masked, but low transmission values tend to result in poor corrections. The 'Use WV values' option will use the values in the WVZ_OBS FITS keyword to correct the data provided that the specific model files (see above) are

located in the specified ATRAN directory. In our M82 data, the WVZ_OBS keyword does not exist, so we will leave the 'Use WV values' box unchecked, as shown below.


Output: The corrected flux is contained in extension 1 of the output, and the smoothed ATRAN model is contained in extension 8.
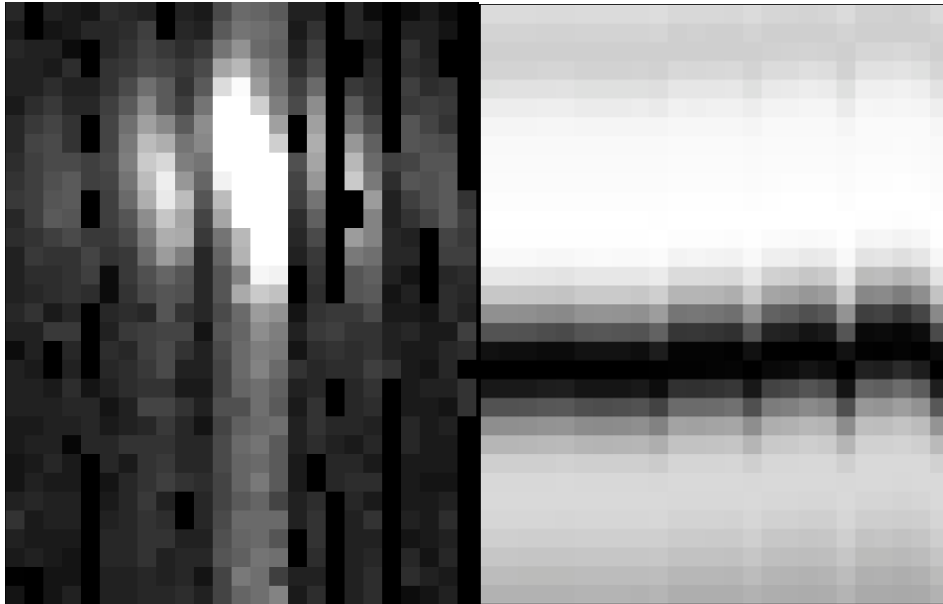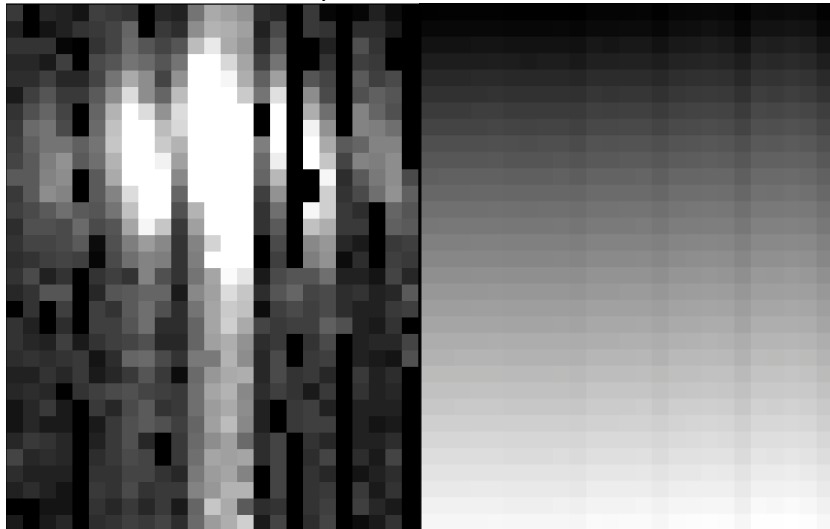

Extension 1 (corrected flux, left) and Extension 8 (atmospheric transmission, right) from F0280_FI_IFS_7004081_RED_TEL_100471-100472.fits.

11. Flux Calibrate: In this module, the instrumental response, derived by FIFI-LS scientists based on calibration observations, is applied to the data.

Options: Aside from the 'Save output' option, the options in this module are to skip the flux calibration module entirely, or to use a different response/calibration file. The response files are updated over time, but the appropriate response file for the date of the observations is automatically used.
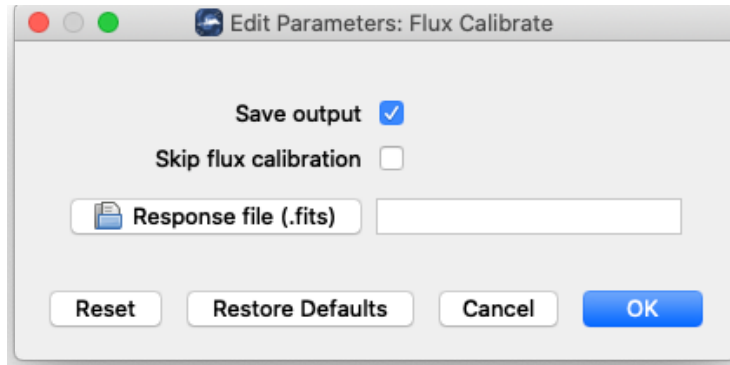
Output: The units in the flux array (extension 1) are now Jy/pixel, and an extension with the instrument response is added to the output file.

Extension 1 (corrected flux in Jy/pixel, left) and Extension 9 (instrument response, right) from F0280_FI_IFS_7004081_RED_CAL_100471-100472.fits.

12. Correct Wave Shift: This module shifts the wavelengths in the data from the topographic reference frame to the solar barycentric reference frame. The shift to the local standard of rest (LSR) frame is calculated and stored in the FITS header (LSRSHIFT) but is not applied to the data. Options: The main option in this module is skip the wavelength shift entirely. Since we're looking at an extragalactic source, we want to make sure the final data are shifted to the solar barycentric frame and will leave the box unchecked (so that the correction will be applied).
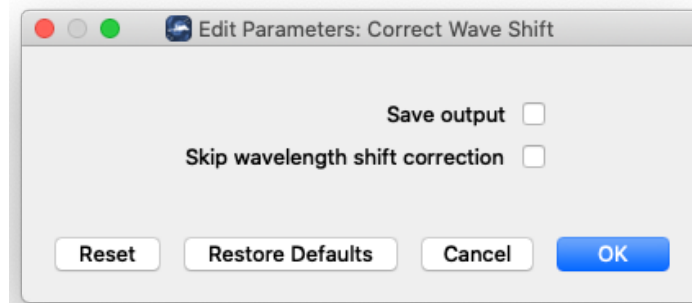
Output: The output fits file contains the corrected wavelength data in extension 5, and the uncorrected wavelength data in extension 6.

13. Resample: This module resamples the flux in each pixel, across all input files, onto a regular 3-dimensional grid (RA, Dec, wavelength). The parameters for this module are in three sections: General, Spatial, and Spectral.

General parameters:
- 'Skip coadd' creates a cube for each input file. These individual cubes can be inspected to determine which, if any, have negative effects on the final combined map
- 'Interpolate instead of fit' is an option to not resample, which might be beneficial in the case of poorly dithered data.
- 'Weight by errors' allows one to consider data errors when resampling, thereby giving less weight to data affected by large uncertainty
- The 'Fit rejection threshold' rejects data based on the local dispersion of the data. Although it can be effective in some cases, in the presence of point sources it can lead to the rejection of observations with higher flux, thereby affecting the final flux of the sources. The pipeline is typically run with is parameter set to -1, inhibiting the rejection.

Spatial parameters:
The values of the spatial resampling parameters depend on the dithering pattern used for the observation and on the objects observed. A good dithering pattern with multiple sampling points will allow a better PSF reconstruction, so that a finer grid can be used to resample the data. This is ideal when observing point sources that will be smoothed using a coarse grid. On the other hand, when observing diffuse emission, it is less important to have a fine grid and it can be convenient to resample with a coarser grid.
The resampling is done using a Gaussian kernel inside a window that is parameterized by the FWHM. A choice of 1.5 times the FWHM for the spatial fit window allows one to reproduce the original PSF. If a higher value is chosen (such as 3), more data will be considered to estimate each value and more smoothing is expected. The other parameter to consider is the smoothing radius, which is connected to the dispersion of the Gaussian kernel. Choosing a value of half the FWHM will result in data that conserves the angular resolution of the instrument. Larger values will result in more smoothed data with lower angular resolution. Therefore, if finer angular resolution is desired, a more pointed (smaller) kernel has to be used in the cube resampling.
The adaptive smoothing option should solve this problem by changing the kernel size depending on the amount of signal. However, this feature has not yet been fully tested.

Spectral parameters:
Observations are usually oversampled in the wavelength direction. Therefore, smaller kernels can be used to recover the spectral resolution completely.
The resampling algorithm creates border artifacts because it assumes continuity of the data beyond the border of the map. For this reason, some data are masked by the algorithm using the 'Spatial edge threshold' parameter. This threshold can be lowered (to 0.5 for instance) to recover data with strong emission that happen to be at the border of the observation. The user should be aware that some of the structures appearing at the limit of the observation are artifacts created by the resampling algorithm.

Edit Parameters: Resample

General Parameters

|	
Save output	☑
Use parallel processing	☑
Skip coadd	☐
Interpolate instead of fit	☐
Weight by errors	☑
Fit rejection threshold (sigma)	-1
Positive outlier threshold (sigma)	-1
Negative outlier threshold (sigma)	-1
Append distance weights to output file	☐

Spatial Resampling Parameters

Spatial oversample (pixels per mean FWHM)	5.0
Output spatial pixel size (arcsec)	3.0
Spatial surface fit order	2
Spatial fit window (factor times FWHM)	3.0
Spatial smoothing radius (factor times FWHM)	1.0
Spatial edge threshold (0-1)	0.7
Adaptive smoothing algorithm	none

Spectral Resampling Parameters

Spectral oversample (pixels per mean FWHM)	8.0
Output spectral pixel size (um)	
Spectral surface fit order	2
Spectral fit window (factor times FWHM)	0.5
Spectral smoothing radius (factor times FWHM)	0.25
Spectral edge threshold (0-1)	0.5

Reset   Restore Defaults   Cancel   OK

Output: The output fits file now contains the final M82 data cube (in RA, Dec, wavelength) for this example.
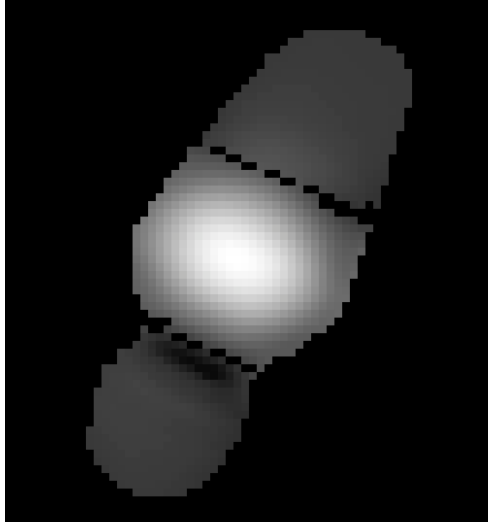
The 157.816 mm frame of F0280_FI_IFS_7004081_RED_WXY_100471-100472.fits.

14. Make Spectral Map: This last module is used to generate a png image file from the final data product.
Options: These options control which slice of the cube will be displayed in the final image, the spatial pixel from which the displayed spectrum will be extracted, and the overall look of the final png image. To get the image shown below, the 'Override wavelength slice' option was used with a value of 33, to display slice 33.

Output:

Object: M82, Channel: RED

Filename: F0280_FI_IFS_7004081_RED_WXY_100471-100660.fits
Wavelength slice at 157.82 um
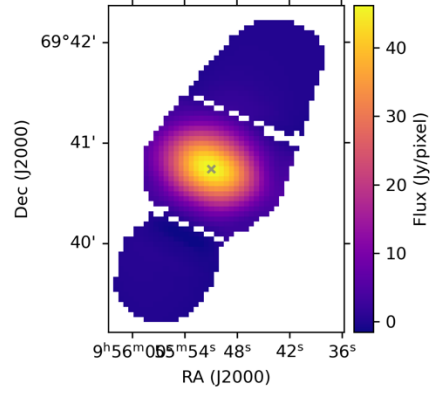
Spectrum at RA=9h55m50.958s, Dec=69d40m44.4636s