OSIRIS Hackathon Final Report

March 28-31, 2017, UCLA
Edited: 2017-05-04

Authors: Anna Boehle, UCLA, Sam Chappell, UCLA, Devin Chu, UCLA, Anna Ciurlo, UCLA, Tuan Do, UCLA, Mike Fitzgerald, UCLA, James Larkin, UCLA, Kelly Lockhart, IfA, Jim Lyke, Keck Observatory, Jamie Ryan, UCLA, Andrey Vayner, UCSD, Greg Walth, UCSD, Shelley Wright, UCSD

Summary

- Flux assignment investigations - we developed a new metric to characterize how the pipeline algorithm is distributing fluxes. We applied this to a large set of data with range in dates, filters, and scales. The correct lenslet and spectral location has the correct flux. The artifacts from flux assignment (+- 32 spectral channels) have both positive and negative values that sum up to about zero. The typical artifact is about 6-7% of the peak flux in the correct lenslet. This effect is relatively insensitive to scale or filter from 2010-2016. Both skies and arc lines show this effect. We believe this is due to the asymmetry in the shape of the PSF on the detector. This primarily affects narrow emission line sources. Continuum sources and broad line sources will likely not show this effect.
- At 100 mas there is a small shift between the location of the spectra on the detector from the white light scan and the science data. The FWHM is also different. This likely occurs because the 100 mas scale does not have a cold pupil. A remedy was attempted with the KcX filters, but due to misalignment, they should be used with caution. A long term remedy may be to re-align the KcX filters. This may be reason that very compact objects like QSOs have flux allocation issues. Shifting and changing the FWHM of the data may help. Additional investigations will be necessary to develop a fix. At the moment, we do not recommend using the 100 mas plate scale.
- Spatial rippling investigations - we developed a metric to characterize the spatial rippling based on the maximum deviation of the integrated flux of OH lines divided by the mean of integrated OH lines. We obtain an average value of 0.5 for 2012 (moderate rippling), 0.8 in 2013, 1.0 in 2015 (high rippling) and 0.3 in 2016 (low or no rippling) in Kn3, Kbb, Hbb bands (Jbb is not well represented by the metric). Using this metric, we find that between the data from 2013-2015 has significantly more rippling that either before the grating upgrade or after the new detector. The rippling varies very rapidly across a filter. The rippling changes (shapes and strength) from filter to filter but not significantly from scale to scale. This effect may be the reason for elevated noise from 2013-2015. Individual channels of OH lines show rippling in all data because of sampling effects. For data taken after 2015, individual channels show rippling primarily due to sampling effects, but the integrated line flux is not affected.

- New cosmic ray module - we have developed new cosmic ray modules and have begun testing them. There is a promising one, but more testing is necessary for a replacement module.
- Scaled sky module - we identified a potential issue with the scaled sky module in how it deals with scaling the continuum. We have a proposed solution and default, but requires further testing.
- 2D OSIRIS model - we created a model of the OSIRIS spectra on the 2D detector by using the rectification matrix, lenslet mapping, and wavelength solution. We will use this tool to study spectral extraction in the future, including 2D PSF extraction, etc.

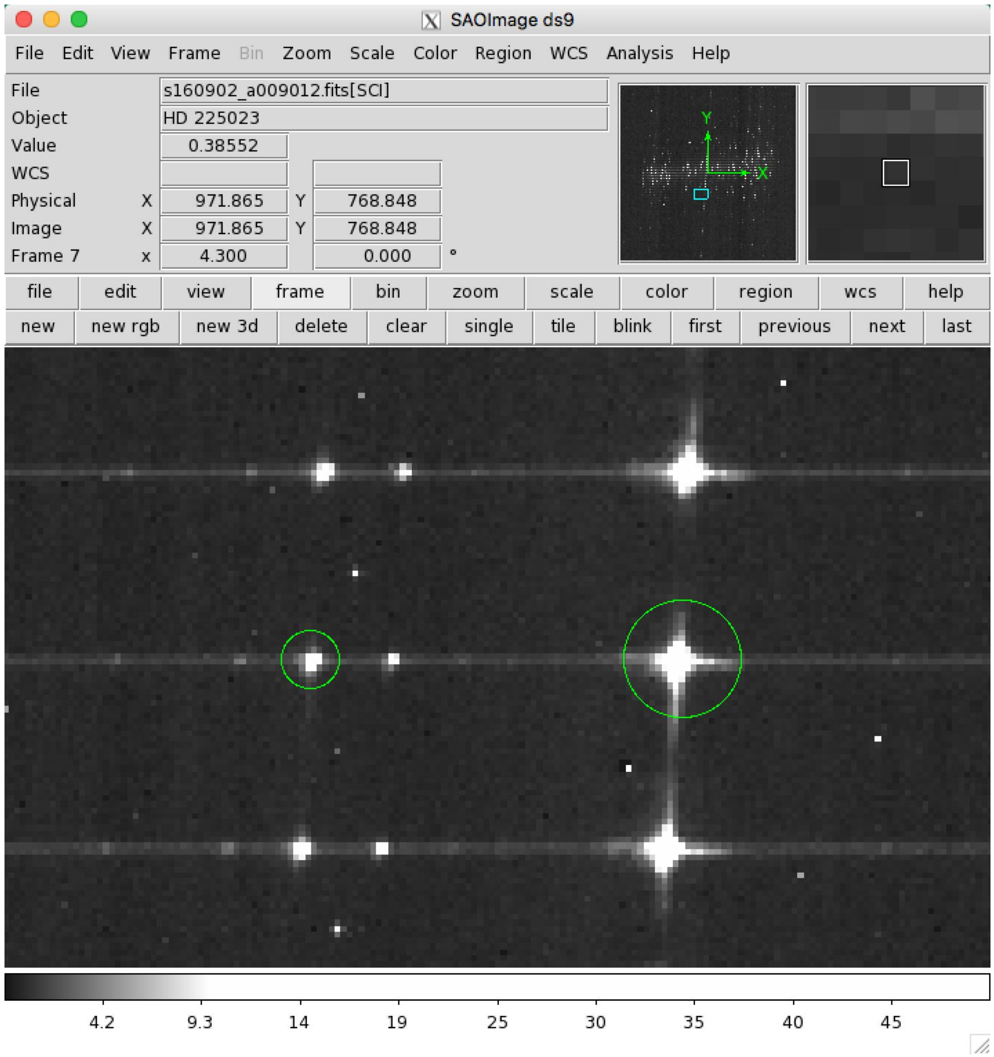
1 Flux assignment investigations

1.1 Characterizing flux assignment: flux conservation tests

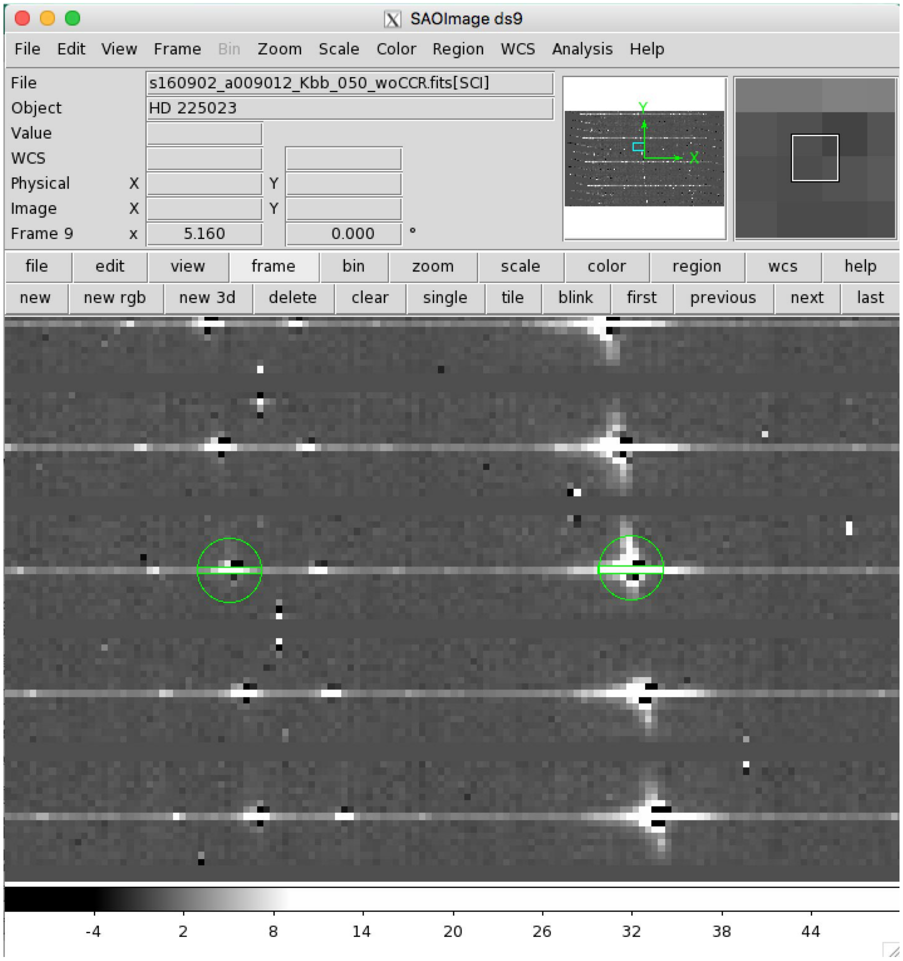
We began using a new method for investigating flux misassignment - the Swap Channels module. Swap Channels (<module Name="Swap Channels for Test" />) is run after spatial rectification (Extract Spectra) and Assemble Data Cube is not run. It orders the extracted spectra according to spectrum number (i.e. in the same order as in the 2D detector), so spectra that are next to each other in the raw data are still next to each other since the wavelength shift has not been applied. So columns in swapped channel frames are still columns from the original data, but the y-axis now represents lenslet number. Since data are extracted one column at a time, it is easier to see what happened on each extraction. This provides another way to visualize the flux misassignment, and allows us to more easily test for flux conservation.

1.1.1 Arc lamp data

We wanted to compare the flux in the original 2-d data with the flux in the extra spectrum. For the extracted values we want to see if the lenslet gets assigned the correct amount of flux, or if flux is actually lost to the neighboring lenslets. We began with testing on single-column Kbb/50 arc lamp data from 09/2016 (new detector). Tested two different arc lines in the same spaxel, as the bright line peaked at half the saturation level. Raw 2D detector image, with an aperture placed on a bright emission line:



And the Swap Channels output, for the same spectrum, with an aperture placed at the same spectrum and emission line as above. In addition, a rectangular aperture is placed over just the illuminated spaxel. With respect to the raw frame above, the vertical axis is flipped and now represents lenslets instead of pixels.



From the channel swapped data you can see that much of the misassigned flux appears in extended streaks above and below the correct lenslet which corresponds to the diffraction spikes of the pupil images. Since the spikes are rotated slightly compared to dispersion, they can't be captured in the 1-column (1D) reductions. But the total flux in these spikes are a tiny percentage of the original flux. More concerning are the p-cygni like spectra produced in the lenslet just above and just below the correct lenslet (white and then black pixels. These effects are shown graphically in the three panels below which are the specta from the correct lenslet and from the ones above and below:

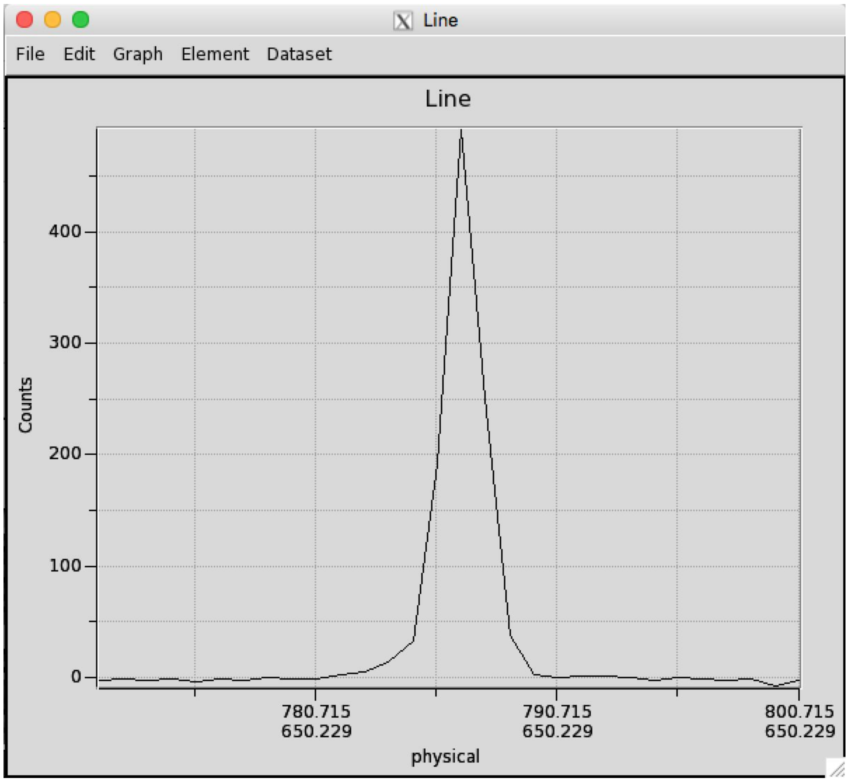


Figure - spectrum in the correct lenslet showing a sharp spectral line. Actually from 2015 data, but very similar to 2016.

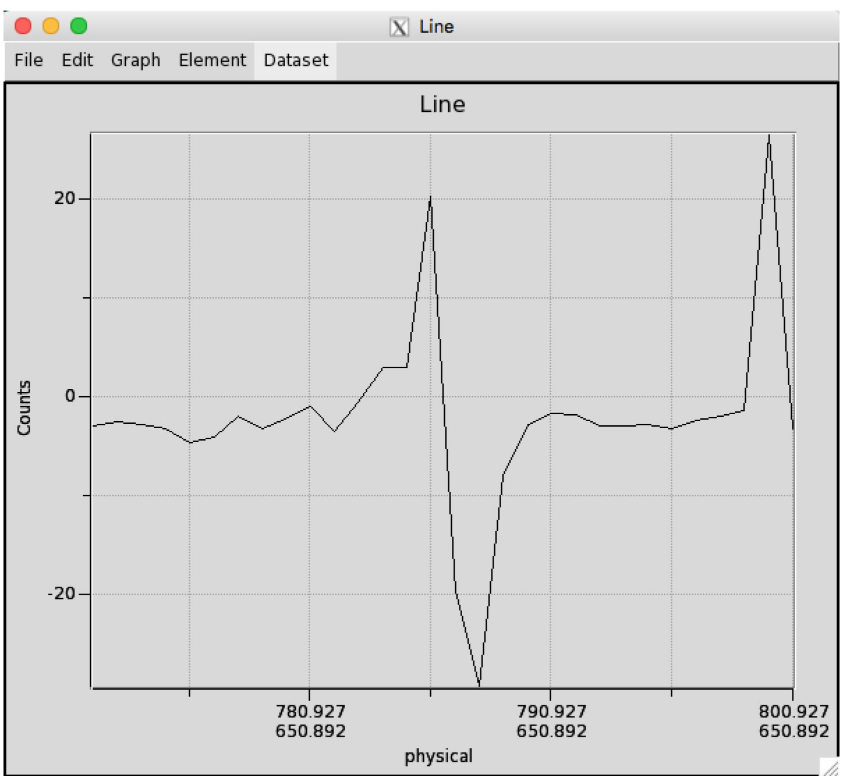


Figure - Spectrum of lenslet above correct lenslet which should have no flux. Actually from 2015 data, but very similar to 2016.

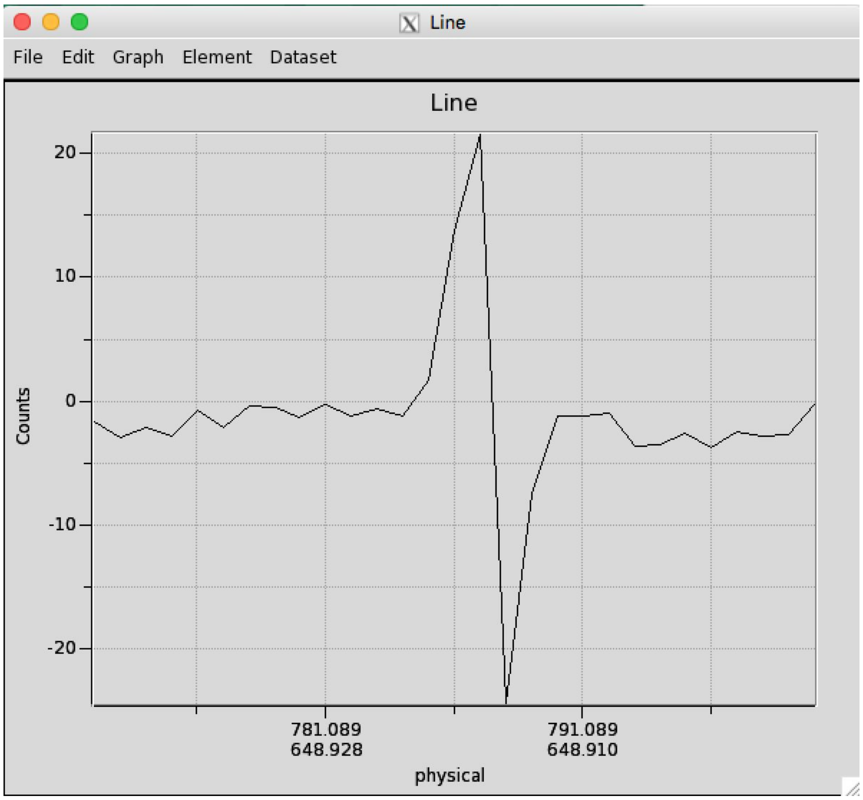


Figure - Spectrum of lenslet below correct lenslet which should have no flux. Actually from 2015 data, but very similar to 2016.

Since only 1 lenslet column is illuminated, all of the flux should have been gathered into a single row (per 19 rows) and the neighboring rows should be dark. (The 3 row wide gray regions correspond to spaxels that have been masked because they do not fall completely on the detector.) However, there is flux visible in the spaxels/rows above and below the illuminated spaxel, indicative of the flux misassignment. Our goal is to quantify the amount of misassigned flux.

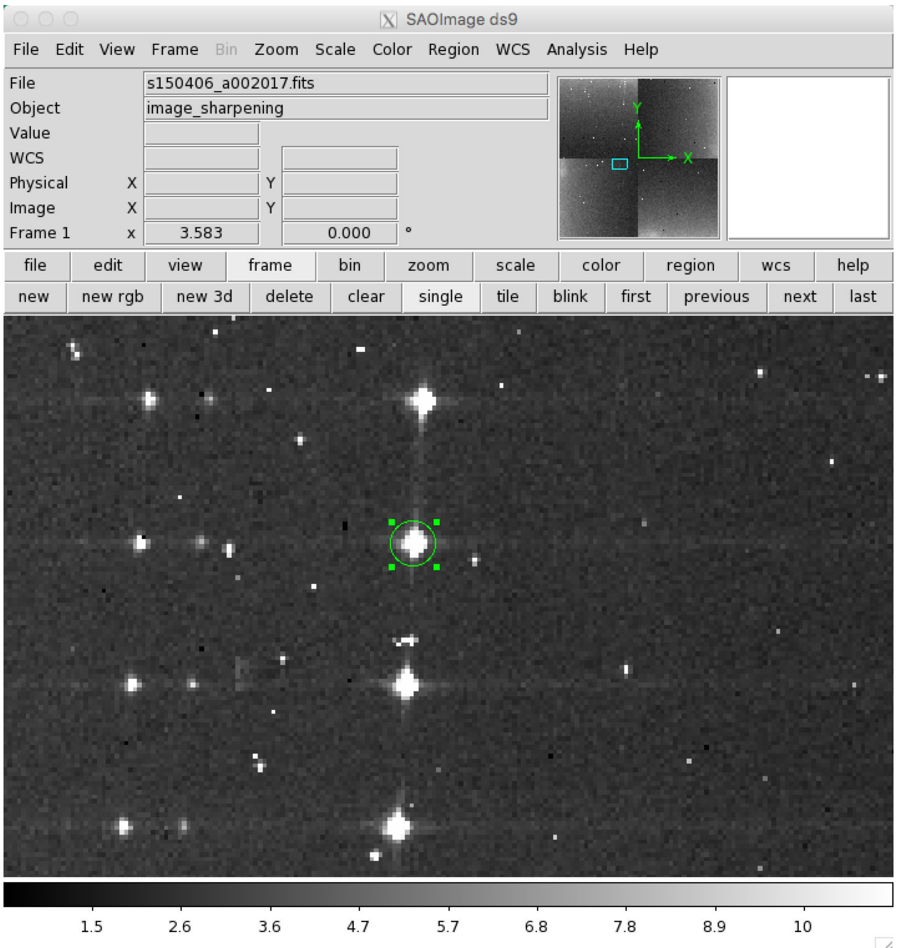
Brighter line in swapped channel frame:
Total counts, circular aperture: 2960 DN/s
Total counts, rectangular aperture (correct lenslet): 2809 DN/s
Peak value in correct spectrum: 1578

Peak deviations, spaxel above: 48, -117 (7%)
Integral of spectrum above: -2 DN/s
Peak deviations, spaxel below: 13, -8
Integral of spectrum below: 34 DN/s

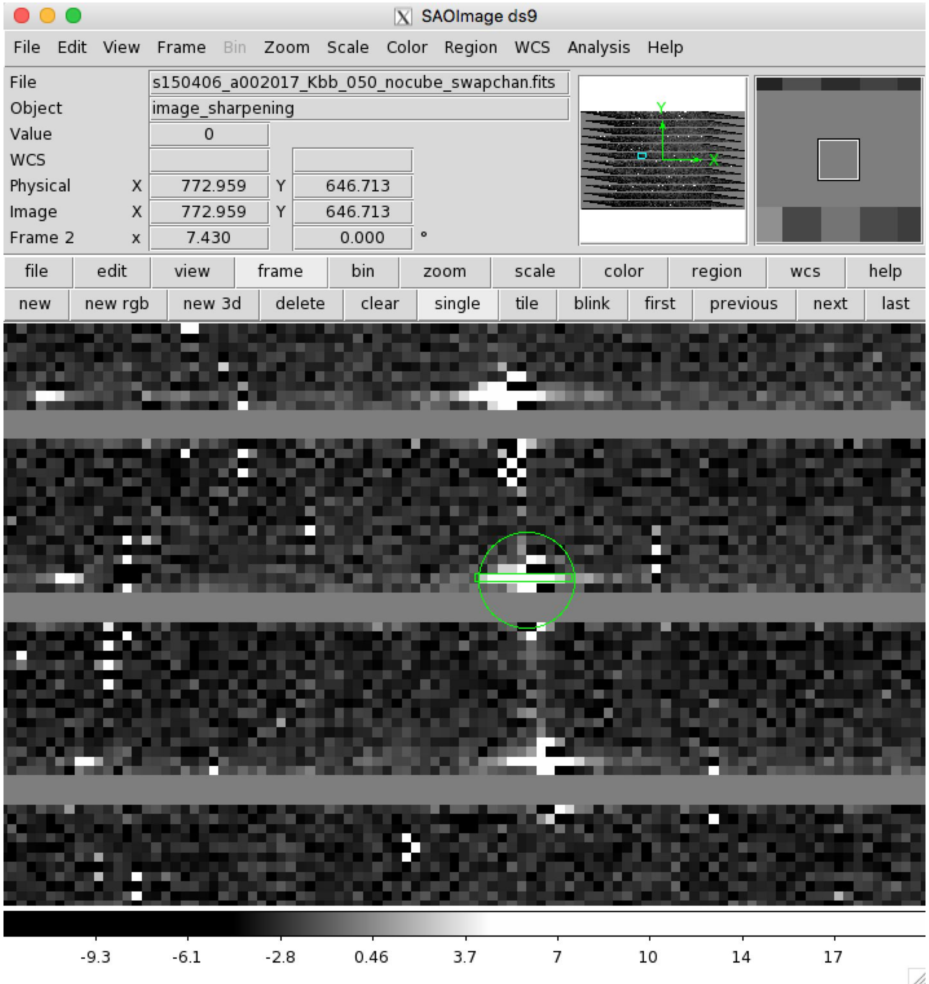
Fainter line:
Total counts, circular aperture: 350 DN/s
Total counts, rectangular aperture (correct lenslet): 328 DN/s
Peak value in correct spectrum: 194

Peak deviations, spaxel above: 8, -12 (6%)
Integral of spectrum above: -2 DN/s
Peak deviations, spaxel below: 0, -2
Integral of spectrum below: 3.5 DN/s

We then looked at similar Kbb/50 single column arcs from 04/2015 (old detector but new grating). Raw frame, with the circular apertures overlaid:



Swap channels frame, with apertures:



Total counts in raw frame, circular aperture: 1288
Total counts in swap channels frame, circular aperture: 939
Total counts in swap channels frame, rectangular aperture (1 lenslet): 1033

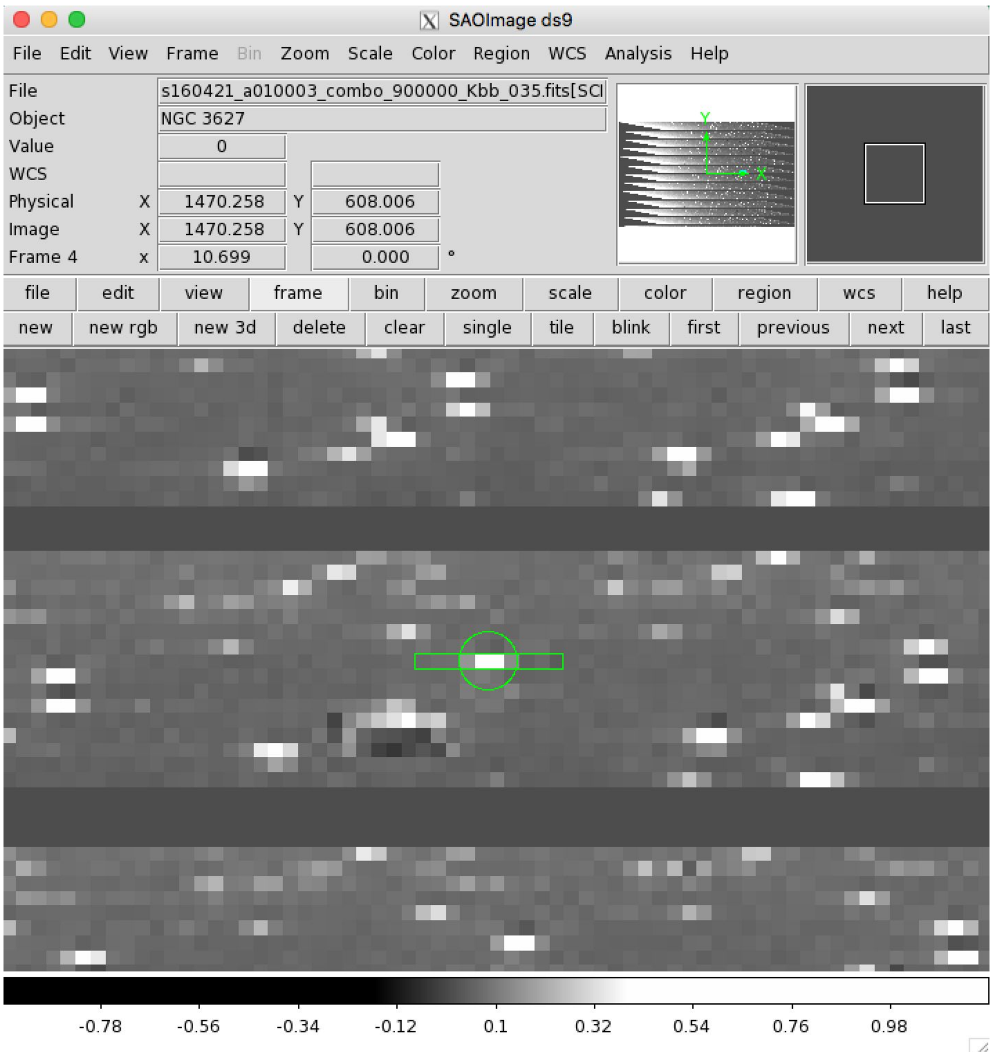
The spatial rectification code divides the raw flux by a factor of 1.28 (for historical reasons), so the flux conservation isn't a one-to-one correspondence, but in this case if we correct for the 1.28 factor, the values roughly match.
Adjusted counts in swap channels frame, circular aperture: 1202
Adjusted counts in swap channels frame, rectangular aperture: 1322

Summary for Kbb 50 from 2015 and 2016
In all of these examples, the correct lenslet has approximately the same amount of flux as a circular aperture placed over many lenslets implying that signal is not being lost due to flux misassignment. This is confirmed with integrals of just the lenslets above and below the correct lenslet typically contains very little flux (consistent with 0). But the neighboring lenslets do typically have a p-cygni profile or small deviation from flux misassignment. In the 2016 data we found that the maximum deviation was about 7% of the peak of the line flux or about 3% of the integral of the line flux.

In these examples, the positive misassigned flux corresponds with the flare seen on the left side of the 2D PSF, and the negative misassigned flux corresponds with the narrower right side of the 2D PSF. This may be due to a mismatch between the 2D PSF and the averaged PSF provided by the white light scan

1.1.2 Sky data

Also tested on sky data - Anna made super skies, which are sets of 6 combined skies (each 900 s, for a total of 1.5 hours). We did the same analysis on the Kbb/35 sky as on the arcs:



Kbb/35: one issue is strong background of roughly 0.05 dn/pixel
Total counts, circular aperture: hard to measure but consistent with line flux
Total counts, rectangular aperture (correct lenslet): 1.2 after subtracting continuum
Peak value in correct spectrum: 0.45 (subtracted)

Peak deviations, spaxel above: 0.02, -0.02 (4%)

Peak deviations, spaxel below: 0.05 (10%), 0

1.2 Metric for flux investigation

So a reasonable metric that will now be pursued is to look at reduced cubes and calculate the peak of the line flux in the correct location. Then look 32 spectral channels before and after the line and measure the maximum deviation (absolute value). We can then report the percentage of this deviation compared to the line peak similar to the 7% found for the 2016 data.

For the lower S/N sky data, we first take the median spectrum across all rows in a single lenslet column before performing this analysis. For the higher S/N arc data, we perform the analysis for each spaxel along the column, then take the median of the results.

1.3 Summary

The following table is a summary of the flux allocation metric for different periods of OSIRIS, different scales and type of data:

Data summary:
https://docs.google.com/spreadsheets/d/11eRiCkkMsVHFysIhho-BjqCFQrWMmtAebX4lf2Cfu0Y/edit#gid=0

Dataset to gather:

Reduced cubes

Post-detector upgrade (2016 - present):

Date	Type	Filter	Scale	Person	Avg. peak, right spaxel (DN s^{-1})	Abs. peak, -1 spax, -32 channels	Abs. peak, +1 spax, +32 channels
2016	QSO	Kn1	50	Andrey			
2016	QSO	Hn2	50	Andrey			
2016	QSO	Hn3	100	Greg			
2016	Sky	Kbb	100		3.5	1.10%	5.30%
2016	Sky	Jbb	50		0.1	5.40%	7.89%
2016	Sky	Hbb	50		1.1	1.35%	2.09%
2016	Sky	Kbb	50		0.8	2.09%	5.99%

2016	Sky	Jn2	35		0.1	15.69%	8.84%
2016	Sky	Kbb	35		0.2	3.05%	6.42%
2016	Sky	Kbb	20		-	-	-
2016	Arcs	Kbb	100		4586.6	1.54%	5.50%
2016	Arcs	Kbb	50		1171.7	1.56%	6.64%
2016	Arcs	Kbb	35		585.2	2.88%	5.89%
2016	Arcs	Kbb	20		227.9	2.64%	5.44%
2016	Whitelight		100				
2016	Whitelight		50				
2016	Whitelight		35				
2016	Whitelight		20				
2016	Nova	Many	35	Jim			
2014	Nova		35	Jim			
2011	Nova		35	Jim			
2014	QSO	Hn3	100	Greg			

Post-grating upgrade, pre-detector upgrade (2012- 2015):

Date	Type	Filter	Scale	Person	Avg. peak, right spaxel (DN s^{-1})	Abs. peak, -1 spax, -32 channels	Abs. peak, +1 spax, +32 channels
2012	Arc	Hbb	20		16.2	2.64%	2.53%
2015	Arc	Kbb	100		457.7	10.04%	7.57%
2013	Arc	Kbb	50		400.0	13.39%	5.74%
2015	Arc	Kbb	50		130.2	7.49%	3.91%
2015	Arc	Kbb	50		427.4	3.97%	7.31%
2015	Arc	Kbb	35		63.3	5.45%	4.98%
2015	Sky	Kn3	100		2.1	1.71%	4.35%

2015	Sky	Hbb	50		0.5	3.10%	1.75%
2013	Sky	Kn3	50		0.9	3.71%	4.10%
2013	Sky	Kn3	35		0.3	2.14%	2.03%
2012	Sky	Kn3	35		0.2	1.00%	2.49%
2012	Sky	Kn3	20		0.1	2.72%	3.02%
2014	Nova		35	Jim			
2014	QSO	Hn3	100	Greg			

Before grating upgrade (pre-2012):

Date	Type	Filter	Scale	Person	Avg. peak, right spaxel (DN s^{-1})	Abs. peak, -1 spax, -32 channels	Abs. peak, +1 spax, +32 channels
2010	Sky	Kc3	100		0.6	0.80%	2.77%
2011	Sky	Kcb	100		0.3	8.60%	4.97%
2011	Sky	Kbb	100		0.1	6.50%	3.80%
2010	Sky	Kbb	50		0.2	6.99%	3.95%
2011	Sky	Kn3	35		0.1	3.95%	3.18%
2011	Nova		35	Jim			

2 Examine 1D Spectral Extraction

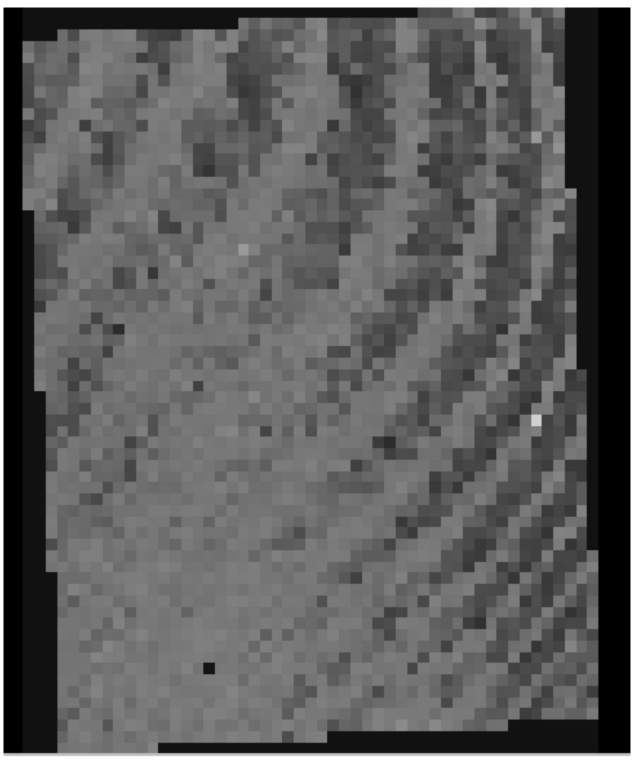
At the October hackathon, James made a revised version of the extraction routine repairing a minor issue with normalizing the rectification matrices. In the process it was possible to add new parameters to make the extraction more or less aggressive and with some smoothing. Kelly tested the new routine with several parameters on the same arc lamp data and this always resulted in the same amount of flux misassignment. So our conclusion is that the deconvolution (extraction) algorithm doesn't matter significantly but it's the mismatch of the PSF from the lamp scans to the other data that creates the flux misassignment. So at least for now, the 1D extraction is optimized and we're not planning more changes.

3 Spatial Rippling Investigations

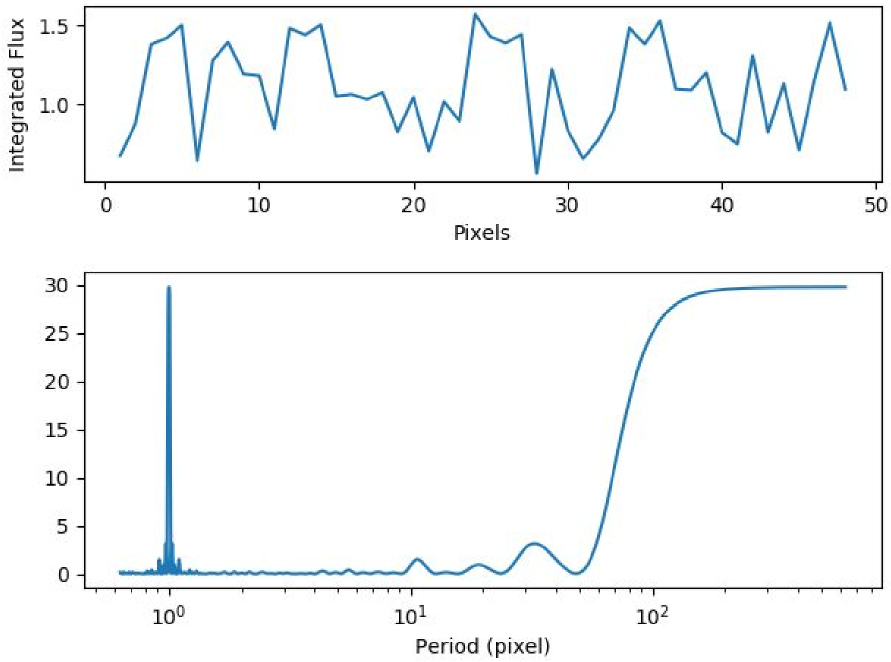
3.1 Metric for spatial rippling

Our goal is to characterize the spatial frequency of the rippling and the amplitude. In order to do this, we will a map of the total flux of a sky line by summing over several spectral channels.

We attempted to measure the frequency of ripples using a power spectrum, specifically Lomb-Scargle. Shown below is from a Kbb 0.035 scale sky image, integrated over the OH sky line at 2.196 microns.



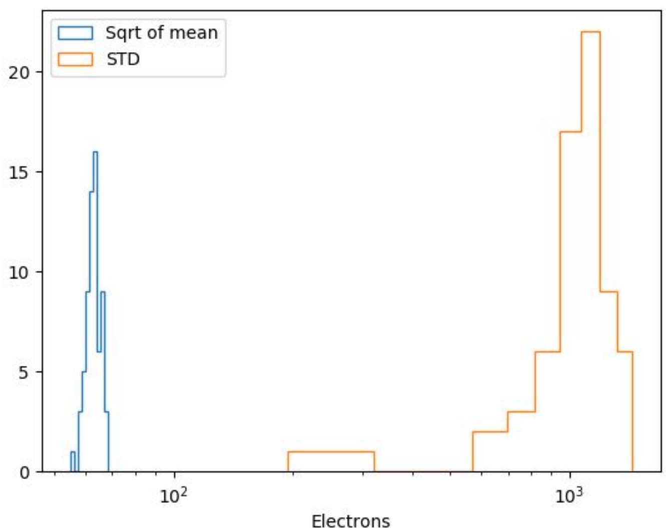
A specific row was then chosen in that flux integrated image. The top figure shows the integrated flux over that row and the lower figure shows the power spectrum as a function of period in pixels. The peaks in the Lomb-Scargle plot occur at roughly one pixel and pixels scales on the same order as the size of the image. This is suggestive of that, while there is noticeable variability in flux, it does not have a characteristic frequency. We turned to other methods of quantifying the spatial rippling.

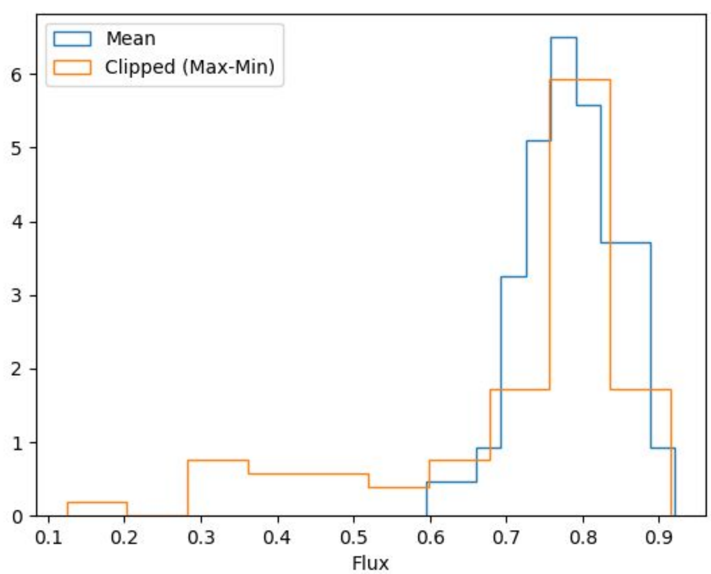


We then select a row and determine the following values:

1. Fringe visibility: (max(flux) - min(flux))/ mean(flux)
2. Standard deviation vs. noise: stdev(flux)/noise (noise = sqrt(# of electrons))

Can we pick single row to quantify rippling? Below shows the distributions of above values across an entire image, all of the rows (130514 Kn3 0.035 sky in 2.196 microns). Distributions appear to be peaked, can use single rows.

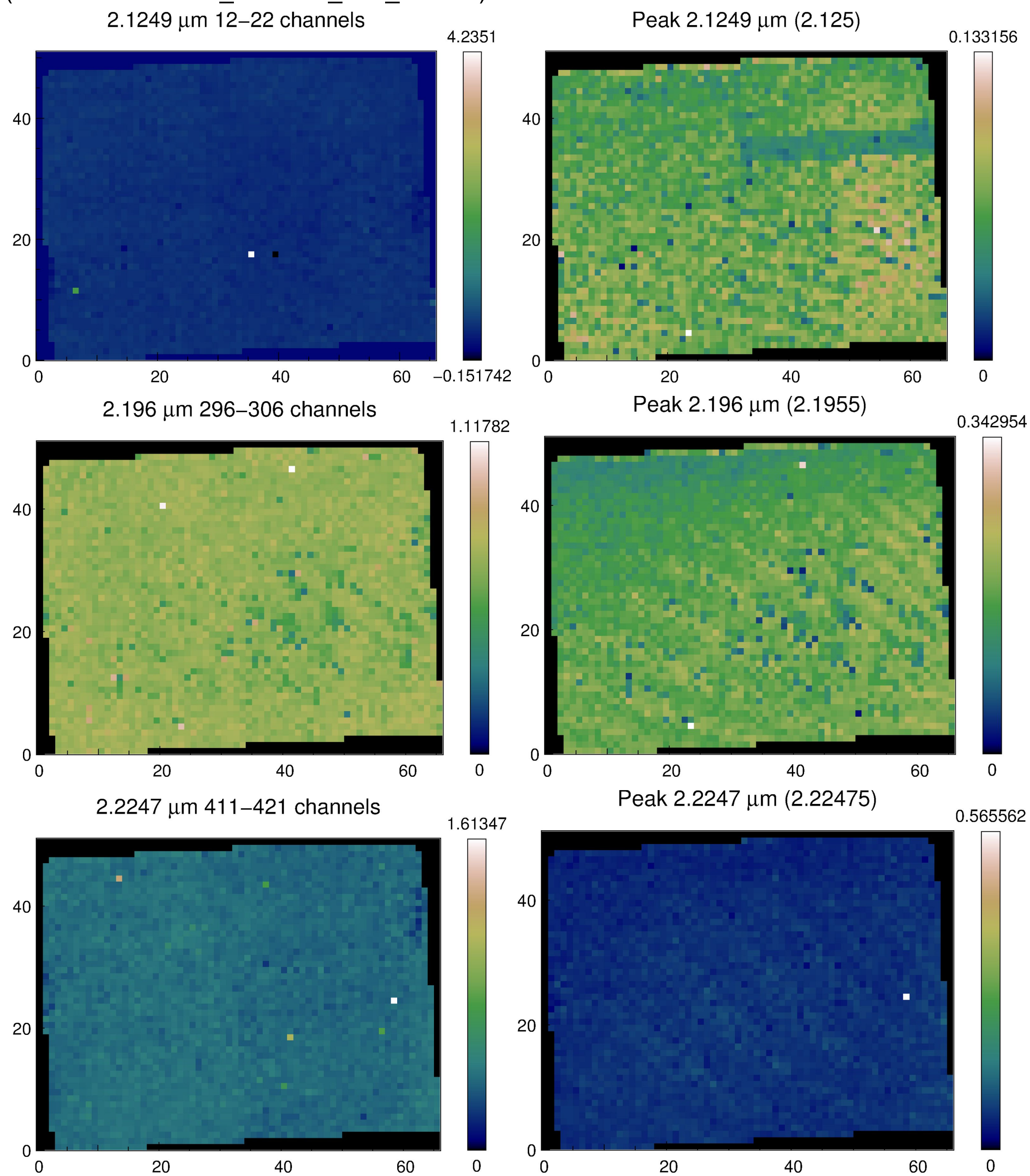

Sqrt of mean
STD
0
5
10
15
20
10^2
10^3
Electrons


Mean
Clipped (Max-Min)
0
1
2
3
4
5
6
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
Flux

3.2 Characterize the spatial rippling in different scales and times

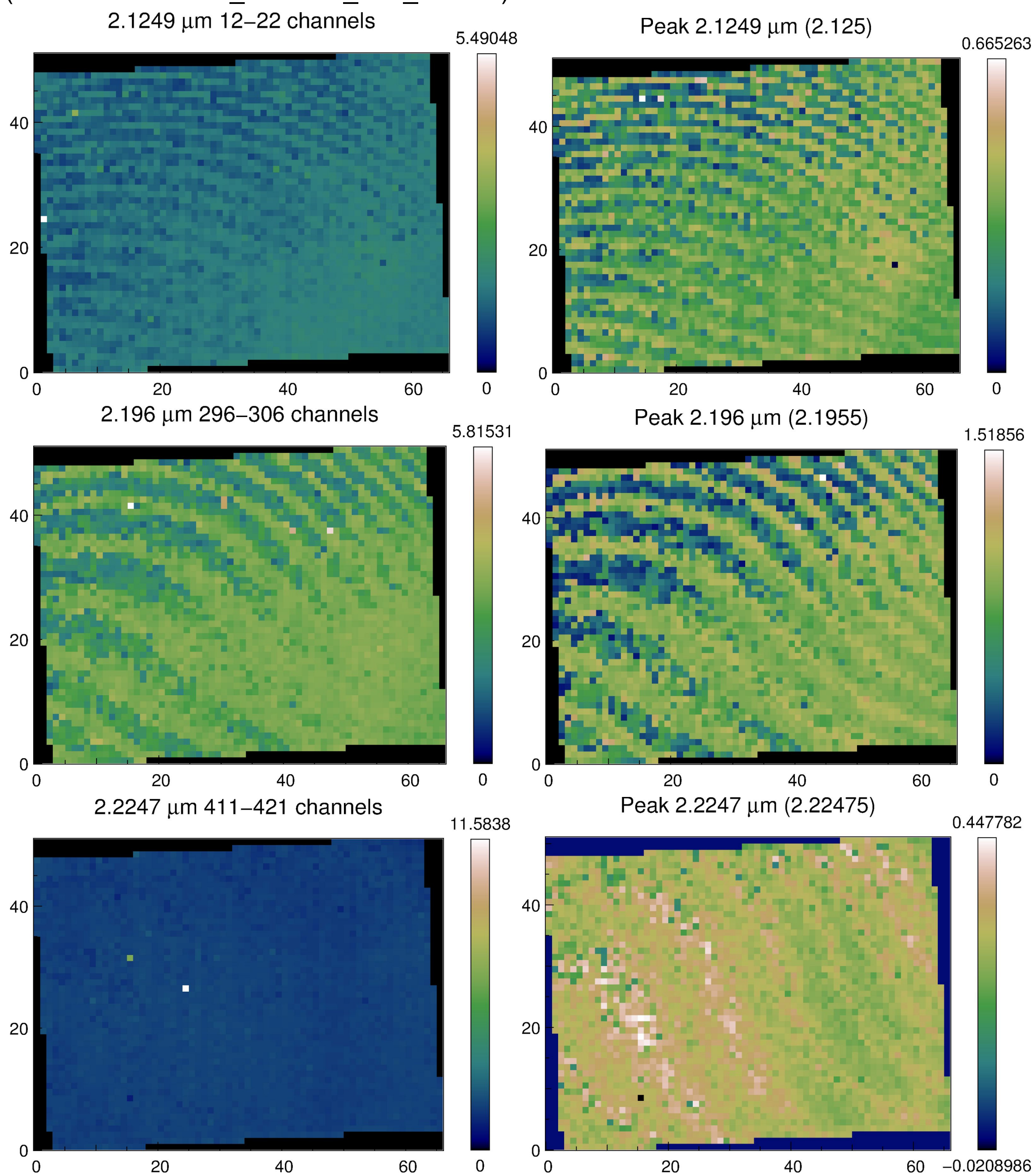
3.2.1 2012 sky data, Kn3 band, scale 35 (new)

(file name: s120609_a024001_Kn3_035.fits)



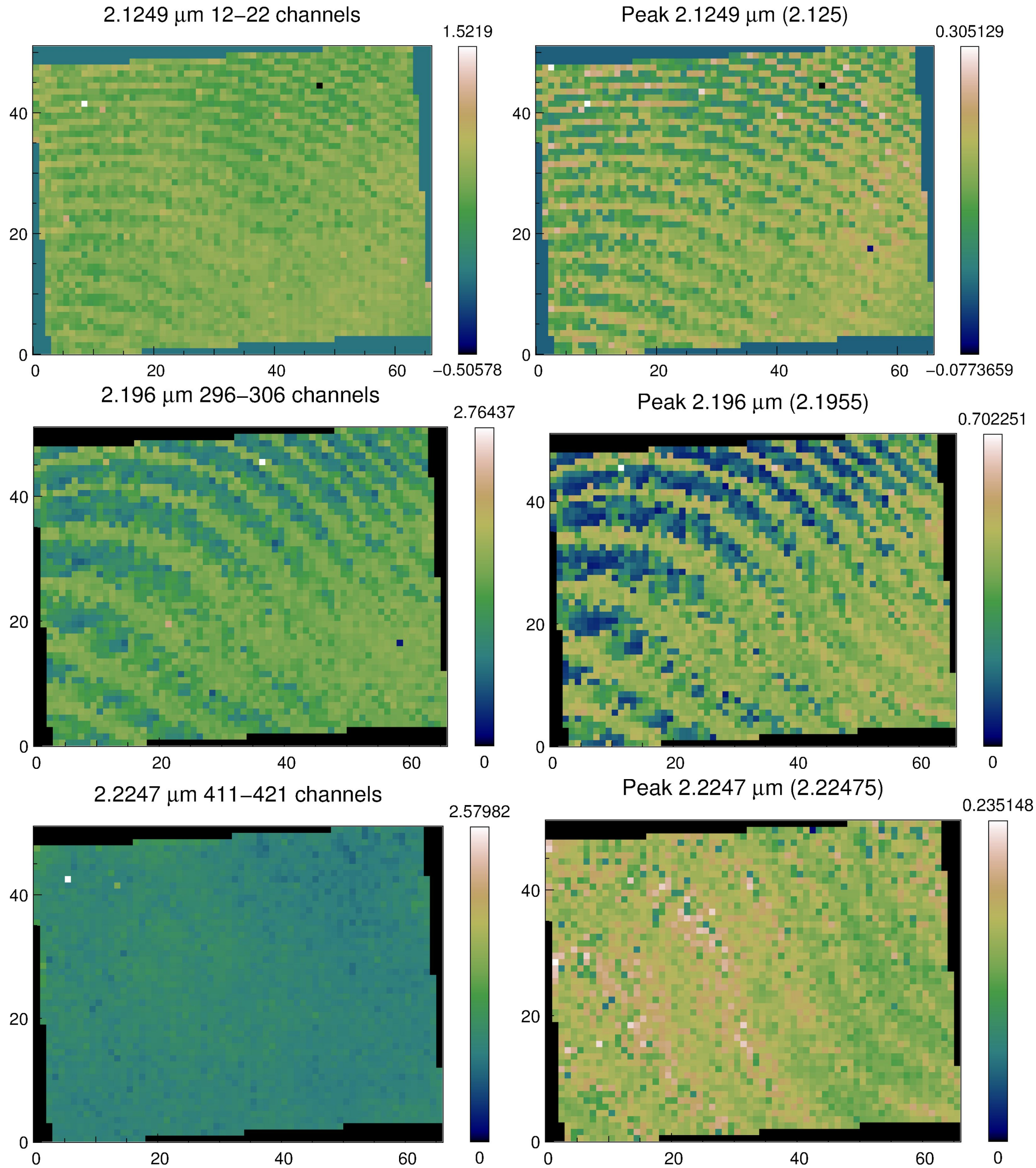
3.2.2 2013 sky data, Kn3 band, scale 50

(file name: s130514_a018001_Kn3_050.fits)



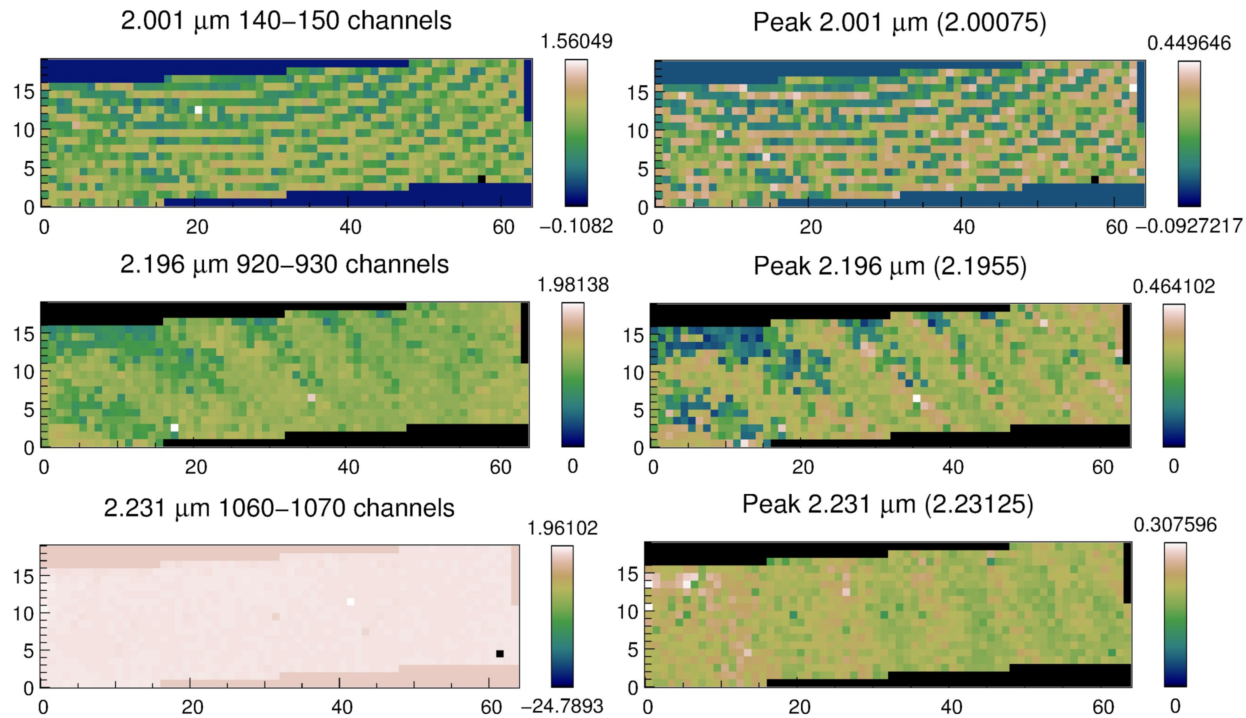
3.2.3 2013 sky data, Kn3 band, scale 35

(file name: s130514_a017001_Kn3_035.fits)



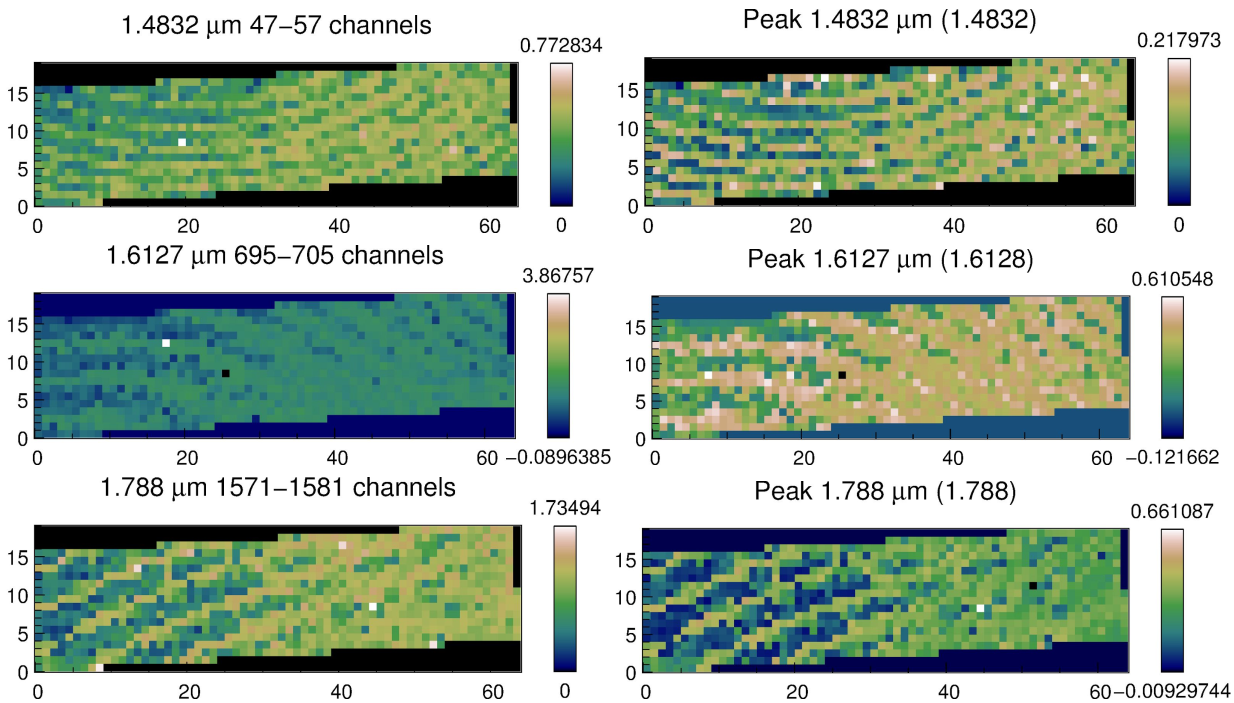
3.2.3 2013 sky data, Kbb band, scale 35 (new)

(file name: sky_130511_Kbb_35.fits)



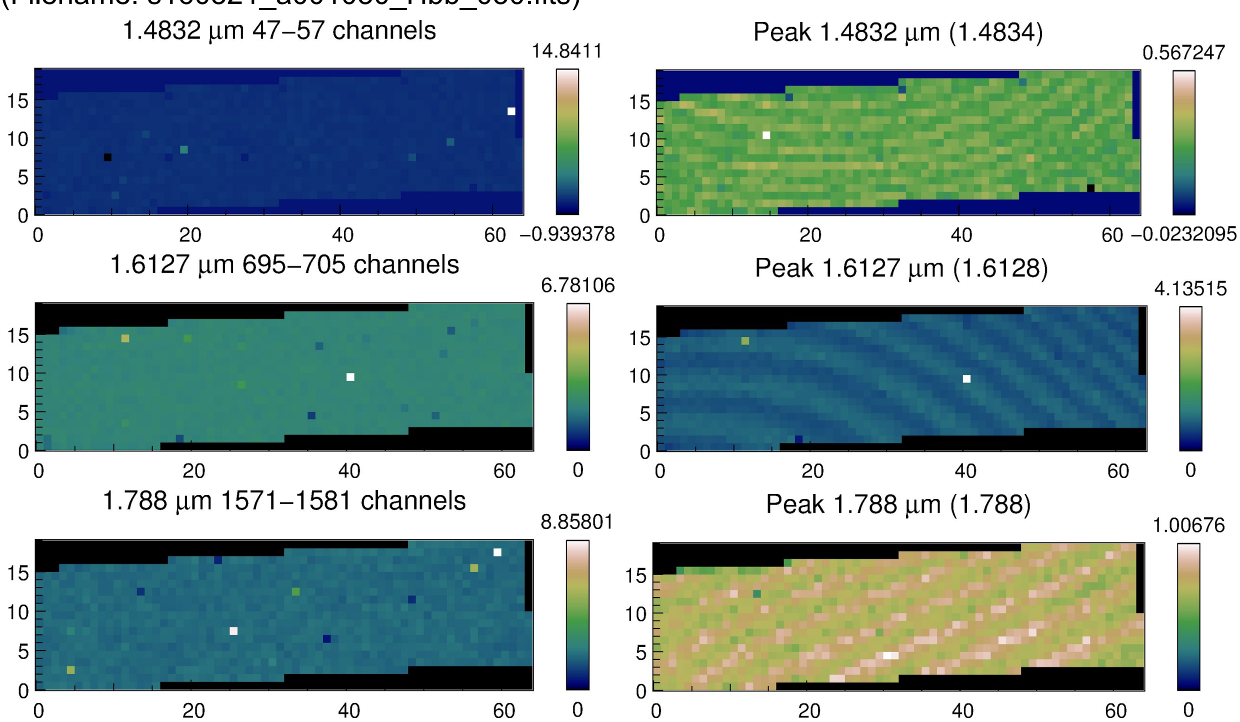
3.2.4 2015 sky data, Hbb band, scale 50

(Filename: s150722_a020001_Hbb_050.fits)



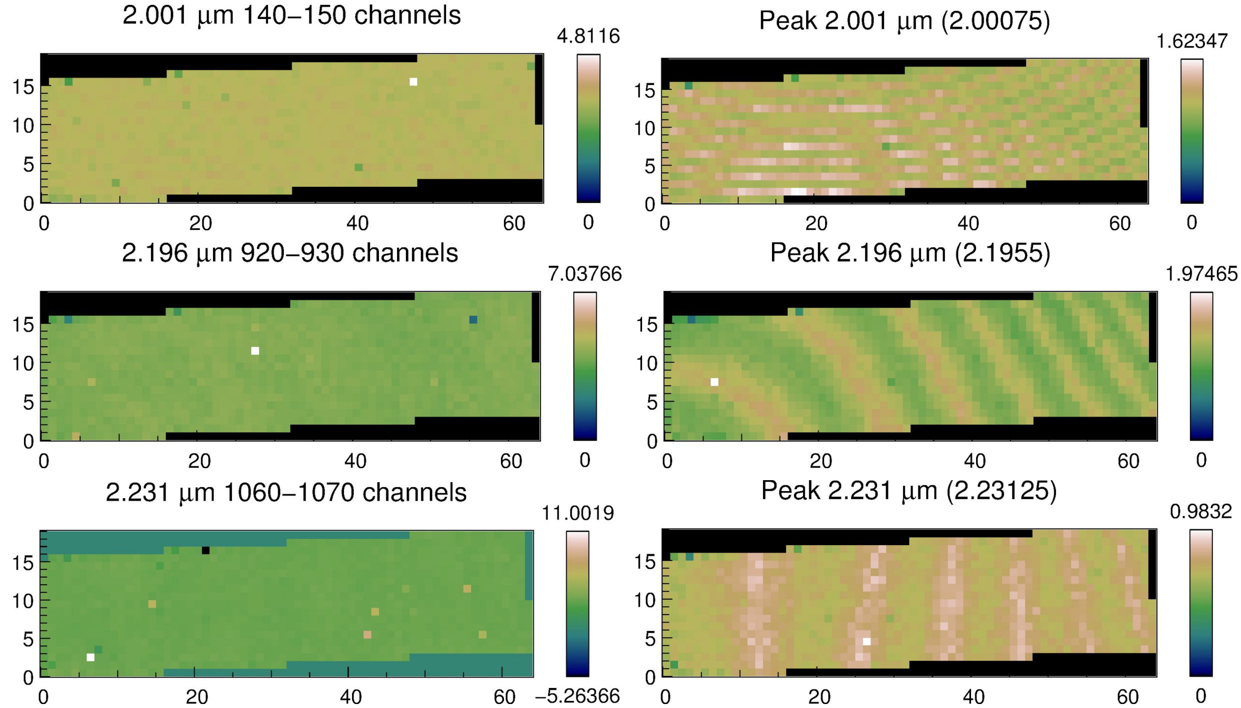
3.2.5 2016 data, Hbb band, scale 50 (new)

(Filename: s160321_a001030_Hbb_050.fits)



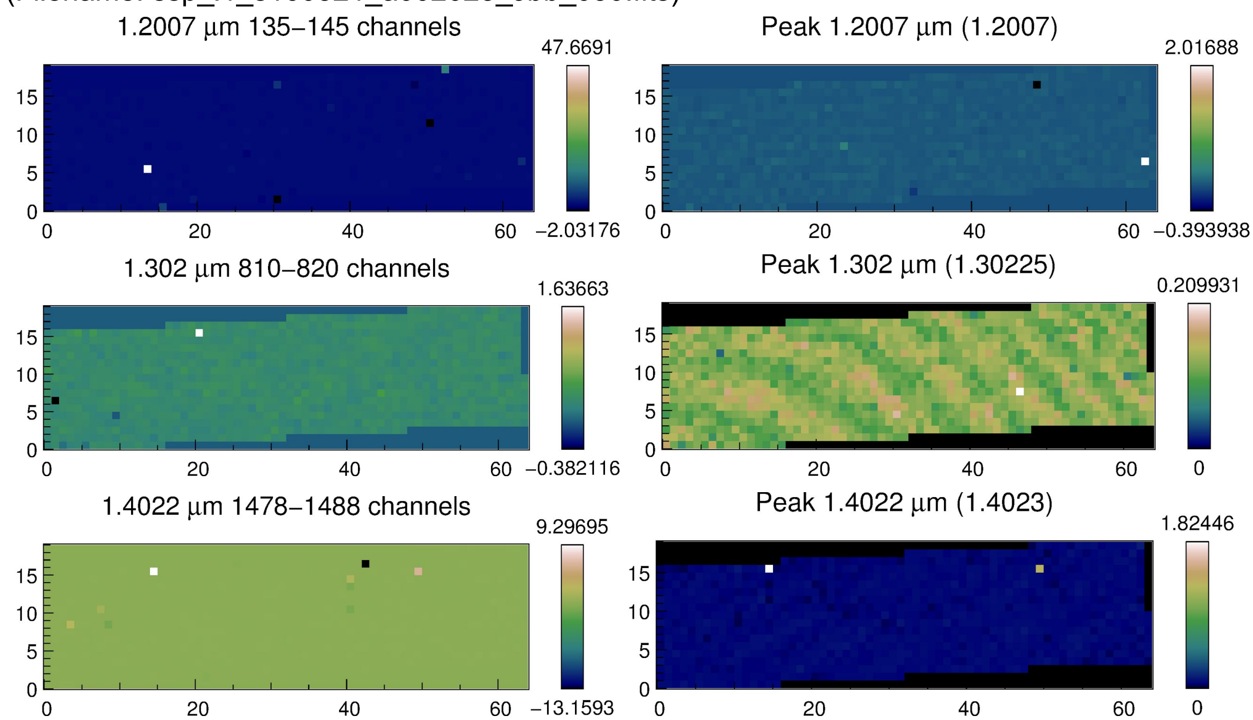
3.2.6 2016 data, Kbb band, scale 35

(Filename: s160514_a010002_Kbb_035.fits)



3.2.7 2016 data, Jbb band, scale 50 (new)

(Filename: ssp_I1_s160321_a002023_Jbb_050.fits)



Summary of qualitative analysis:

- The rippling changes from one side of the filter to the other in a not very symmetric way.
- The rippling shape slightly changes from filter to filter (but could also depend on how close to the border you are for different filters).
- No variations from scale to scale for a given filter.
- From year to year the rippling shape remains the same for a given filter.
- From year to year the effect of rippling on the integrated flux changes, with a smaller rippling for 2012 that 2013 and 2015. In 2016 the problem seems completely solved.

Fits file with the 2D images for each of the 3 selected OH line
(dimensions: [pxl_x, pxl_y, OH_line]) for both integrated ("int_") and peak ("peak_") values :

https://drive.google.com/drive/u/0/folders/0Bw9IjgefYXGjZjZVb0hfcDg5Ykk

3.3 Table

Metrics values integrating on 11 channels centered on a given OH line.

ATT! Here row and columns are inverted with respect to section 3.2 (as in Section 3.1)

				< Row	>	< Col	>
Date	Filter	Scale	λ(nm)	STD/mean	Max-Min/mean	STD/mean	Max-Min/mean
2012-06-09	Kn3	0.035	2125.0	4.14	0.48	4.92	0.63
2012-06-09	Kn3	0.035	2196.0	3.71	0.27	5.2	0.35
2012-06-09	Kn3	0.035	2224.7	3.58	0.4	4.32	0.37
2013-05-14	Kn3	0.035	2196.0	19.18	0.9	11.74	0.32
2013-05-14	Kn3	0.050	2196.0	30.79	1.08	11.65	0.38
2013-05-14	Kn3	0.035	2125.0	12.28	0.84	7.62	0.57
2013-05-14	Kn3	0.050	2125.0	13.9	0.65	9.02	0.33
2013-05-14	Kn3	0.035	2224.75	5.7	0.39	4.99	0.35
2013-05-14	Kn3	0.050	2224.75	6.84	0.34	4.96	0.23
2013-05-11	Kbb	0.035	2001.0	15.55	0.8	19.3	1.21
2013-05-11	Kbb	0.035	2196.0	13.72	0.61	8.68	0.46
2013-05-11	Kbb	0.035	2231.0	5.2	0.3	5.38	0.19
2013-05-11	Kbb	0.035	2125.0	5.16	0.37	5.74	0.55
2013-05-11	Kbb	0.035	2224.7	6.02	0.38	5.92	0.29
2015-07-22	Hbb	0.050	1483.2	11.61	0.91	11.57	1.08
2015-07-22	Hbb	0.050	1612.7	23.94	1.02	18.88	1.01
2015-07-22	Hbb	0.050	1788.0	32.36	1.68	24.52	1.2

2016-05-14	Kbb	0.035	2001.0	59.24	0.07	73.62	0.13
2016-05-14	Kbb	0.035	2196.0	77.46	0.08	78.89	0.1
2016-05-14	Kbb	0.035	2231.0	93.41	0.11	71.24	0.13
2016-05-14	Kbb	0.035	2125.0	106.24	0.08	76.18	0.16
2016-05-14	Kbb	0.035	2224.75	98.2	0.13	132.97	0.13
2016-05-14	Kbb	0.035	2247.0	311.94	0.13	81.87	0.14

2016-03-21	Hbb	0.05	1483.2	54.67	0.19	123.21	0.37
2016-03-21	Hbb	0.05	1612.7	69.75	0.16	66.13	0.15
2016-03-21	Hbb	0.05	1788.0	74.13	0.15	118.23	0.28

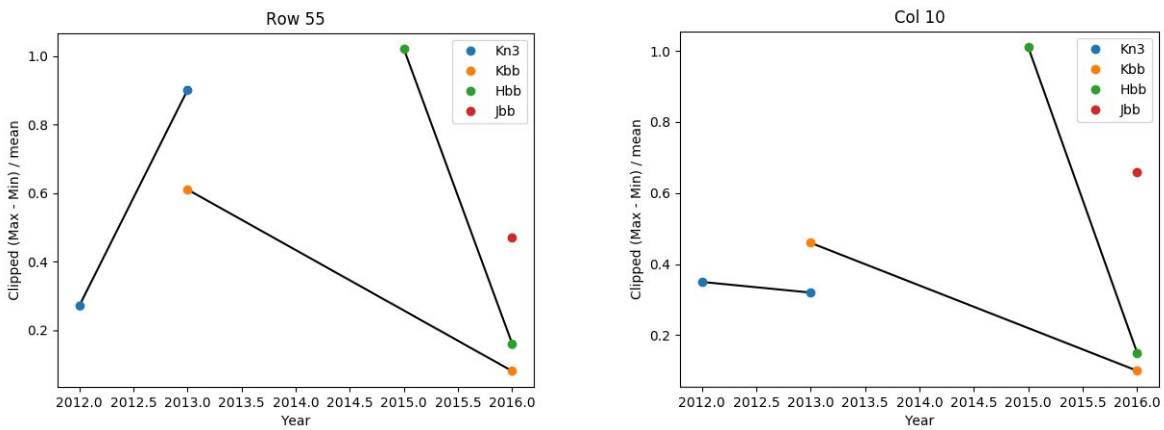
2016-03-21	Jbb	0.05	1200.7	230.61	1.7	224.1	2.01
2016-03-21	Jbb	0.05	1302.0	92.21	0.47	97.79	0.66
2016-03-21	Jbb	0.05	1402.2	172.62	1.54	416.24	1.24

Integrating over 5 channels

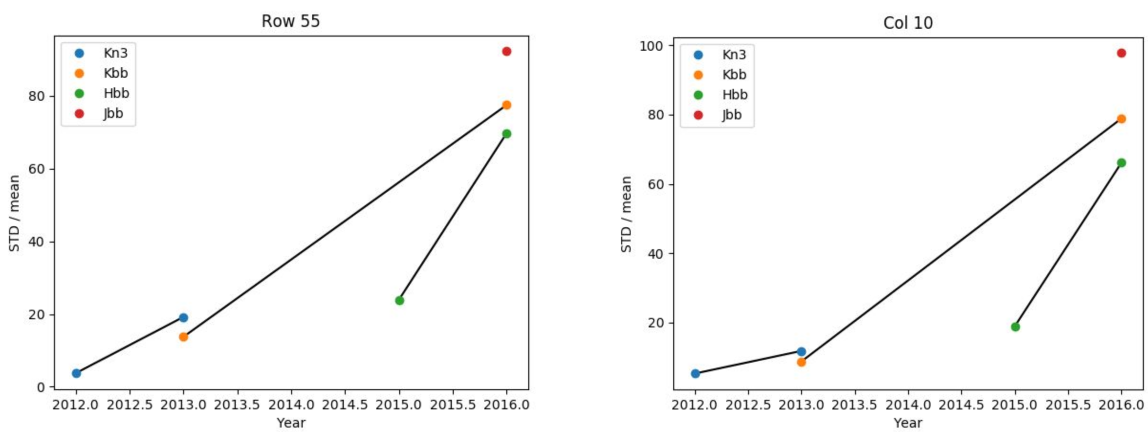
2013-05-14	Kn3	0.035	2196.0	19.33	1.08	11.4	0.4
2013-05-14	Kn3	0.050	2196.0	31.26	1.4	10.31	0.41
2016-05-14	Kbb	0.035	2196.0	104.49	0.14	110.34	0.15
2013-05-14	Kn3	0.035	2125.0	14.12	1.15	7.23	0.79
2013-05-14	Kn3	0.050	2125.0	14.66	0.88	8.48	0.42
2016-05-14	Kbb	0.035	2125.0	69.12	0.09	70.66	0.19
2016-05-14	Kbb	0.035	2247.0	118.22	0.19	80.05	0.23
2015-07-22	Hbb	0.05	1483.2	9.57	0.86	10.76	1.13
2015-07-22	Hbb	0.05	1612.7	25.16	1.14	19.42	1.18
2015-07-22	Hbb	0.05	1788.0	32.38	1.62	23.49	1.33
2016-05-14	Kbb	0.035	2001.0	76.59	0.14	78.46	0.15
2016-05-14	Kbb	0.035	2231.0	66.45	0.11	67.63	0.13

Integrating over 3 channels

2013-05-14	Kn3	0.035	2196.0	10.92	1.06	7.23	0.35
2013-05-14	Kn3	0.050	2196.0	20.24	1.32	6.86	0.46
2016-05-14	Kbb	0.035	2196.0	78.89	0.15	93.66	0.23
2013-05-14	Kn3	0.035	2125.0	15.0	1.39	7.25	0.79
2013-05-14	Kn3	0.050	2125.0	15.45	1.15	8.24	0.52
2016-05-14	Kbb	0.035	2125.0	78.34	0.23	80.88	0.27
2016-05-14	Kbb	0.035	2247.0	85.9	0.18	72.29	0.32
2015-07-22	Hbb	0.05	1483.2	8.3	1.08	8.66	1.21
2015-07-22	Hbb	0.05	1612.7	23.99	1.37	17.7	1.31
2015-07-22	Hbb	0.05	1788.0	31.56	1.88	21.23	1.54
2016-05-14	Kbb	0.035	2001.0	99.53	0.18	106.2	0.19
2016-05-14	Kbb	0.035	2231.0	93.41	0.11	71.24	0.13



Above clipped max - min values divided by mean across a row (left) and across a single column (right). For multiple filters, across multiple years. For kn3 the 35 scaling was used. For Kn3, Kbb, Hbb, and Jbb, the values for wavelength 2196, 2196, 1612.7, and 1302 are what is used for the plots. Below are the same plots for the STD / mean value across the row and single column.



The clipped (max - min)/mean is the quantity that gives most information about how much the integrated quantity are affected. When this value is lower than 0.30 the dataset seems to be sufficiently unaffected.
However the most reliable way to check how much a dataset is affected is looking directly at integrated OH lines across the filter.

3.4 Summary Table

(Max - Min) / avg metric summary values (picking most significant between row and columns).

For vaWe obtain an average value of 0.5 for 2012 (moderate rippling), 0.8 in 2013, 1.0 in 2015 (high rippling) and 0.3 in 2016 (low or no rippling) in Kn3, Kbb, Hbb bands (Jbb is not well represented by the metric)

	Kn3			Kbb			Hbb			Jbb		
year	L	C	R	L	C	R	L	C	R	L	C	R
2012	0.5	0.3	0.4									
2013	0.6	0.9	0.3	1.2	0.6	0.2						
2015							1.0	1.0	1.2			
2016				0.1	0.1	0.1	0.3	0.2	0.2	1.7	0.5	1.2

ATT! This metric is not representative of Jbb rippling (the metric values are high but actually the images show no rippling).

4 Scaled Sky Subtraction

The current version of the Scaled Sky Subtraction module is coded to subtract the continuum from the sky and the object spectra, scale the sky lines in the sky spectrum to match the sky lines in the object spectrum, and then apply the best fit scale factor to the entire sky spectrum (including the continuum) before subtracting it from the object spectrum. The spectra used to perform the scaling are identified in the cube via the Min_Sky_Fraction and Max_Sky_Fraction keywords - the selected spaxels are combined into a single spectrum for each of the sky, S_{sky}, and object, S_{obj}, cubes. The fitting of the scale factor, f_{lines}, is performed with these integrated spectra, then the application of the scale factor and the scaled sky subtraction is performed on a spaxel-by-spaxel basis between the object and sky cubes. This can be represented by:

$$S_{obj,final} = S_{obj,initial} - (S_{sky} \times f_{lines})$$

The Scale K Continuum option can be used for K band data, where there is a strong sky continuum. In addition to the usual sky line scaling, it also determines a sky continuum scaling factor (actually, a scaling factor vector), f_{cont}, to apply to the sky spectrum before subtracting it from the object spectrum. However, this option requires that the object cube contains enough spaxels that are dominated by sky emission. This can be represented by:

$$S_{obj,final} = S_{obj,initial} - (S_{sky}[OH\ lines] \times f_{lines} + S_{sky}[continuum] \times f_{cont})$$

If there is a strong sky continuum, but there are not enough sky-dominated spaxels in the object cube to adequately scale the sky continuum to match, neither of these options produces the desired results. If Scale K Continuum is not run, the entire sky continuum is scaled using the same scaling factor used for the sky lines. If Scale K Continuum is run, the sky continuum is scaled to match the object-dominated continuum level. Either way, the sky continuum will be oversubtracted.

To resolve this, it's possible to separate the sky lines from the sky continuum, apply the sky line scaling factor, and then add back the original sky continuum. This can be represented by:

$$S_{obj,final} = S_{obj,initial} - (S_{sky}[OH\ lines] \times f_{lines} + S_{sky}[continuum])$$

A preliminary version of this code is on Github as pull request #62. Before folding into the main pipeline, the keyword and default behavior must be decided.

4.1 Tests of Scaled Sky Subtraction

We tested two data sets for scaled sky subtraction:

- M31 nucleus observations - this data does not have pure sky pixels in the science frame. Previous methods of scaled sky subtraction led to incorrect continuum fits
- Stars at the Galactic center - this data set have observations of faint stars with long exposures so that there are strong OH lines.

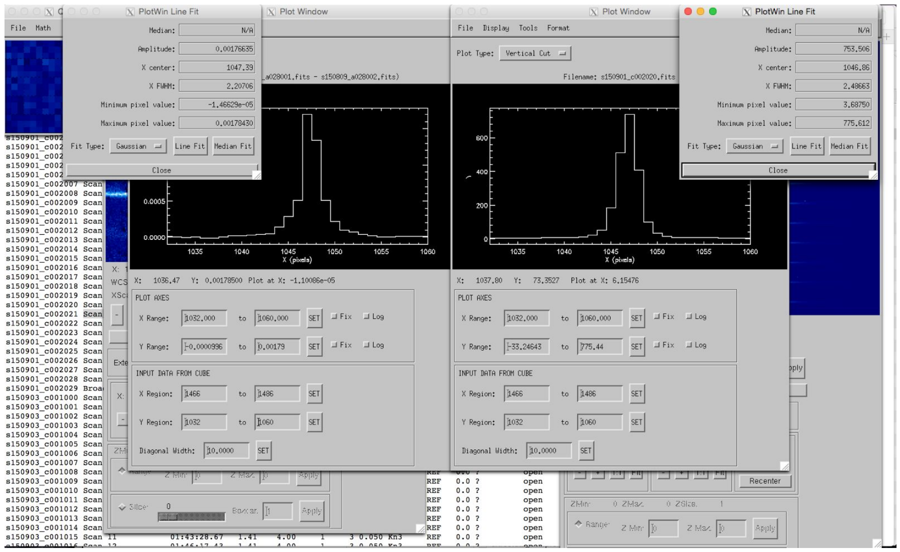
The goals of the tests are to:

- Determine how well the various options (including the new one) are removing OH lines
- Determine the effect on the continuum in science data
- Determine whether the new method leads to at least as good results as the current default.

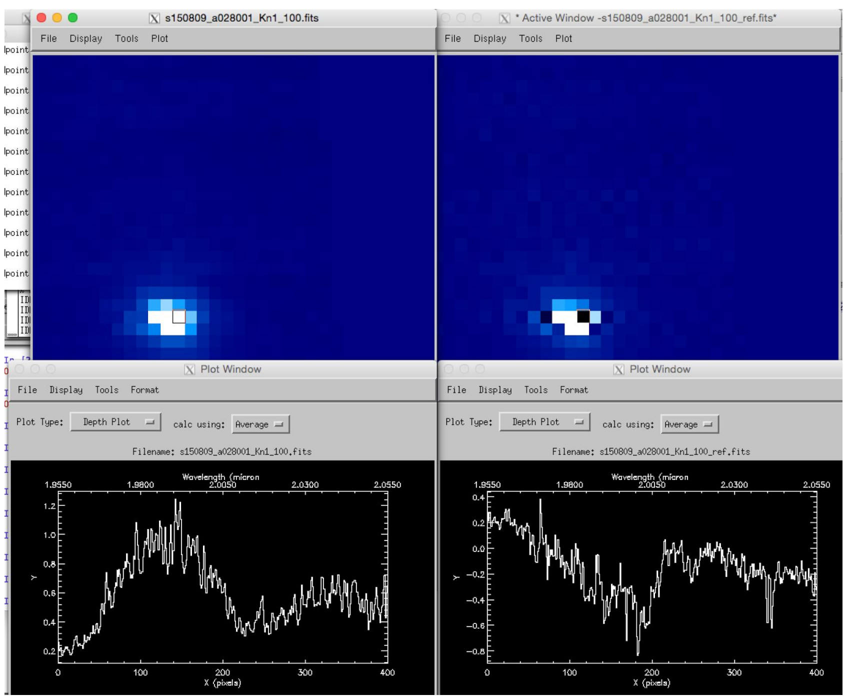
These tests have begun, but have not been concluded.

5 QSO - dark spaxels analysis:

Examining the white light data and comparing it to the raw QSO continuum data we found a spatial mismatch between the continuum centroid and the white light centroid for the same spaxel. In addition we found that the white light scan is broader than the quasar continuum. On the left is a plot of the vertical cut through the quasar continuum with a gaussian fit, and on the right is a vertical cut through the white light scan.



We applied a shift to the quasar data to match the centroid of the continuum to the white light and smoothed the data in the spatial direction and fully reduced the data.



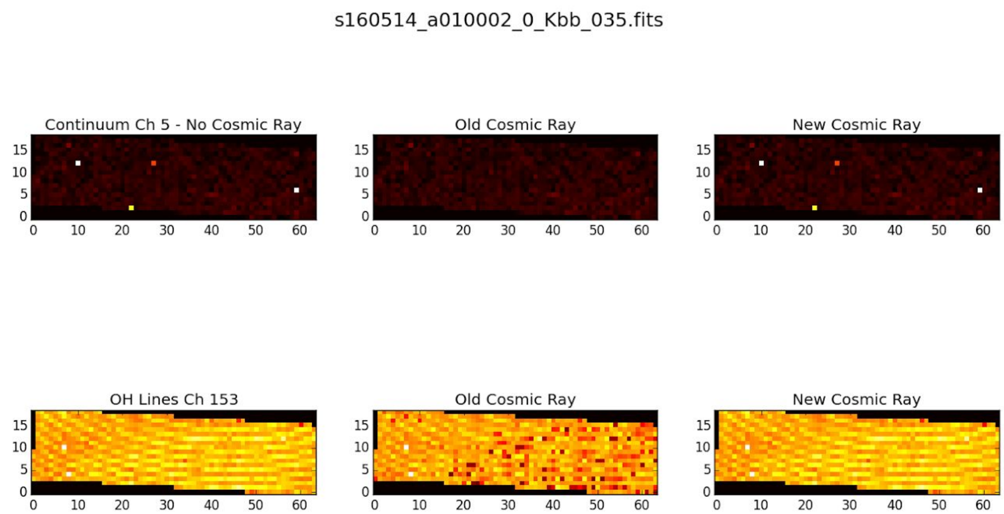
On the left is the shifted + smoothed data and on the right is the original. Below we plot the spectrum of a bad spaxel (20,33). We see a drastic change in the smoothed and shifted cube particularly the broad emission line of the QSO peaking near channel 120 is visible, and the shape matches much better to the peak spaxel (19,33).

6 Clean Cosmic Rays

There is concern that the old cosmic ray module was clipping OH emission lines starting with the new data in 2016. With the new detector upgrade, the OH emission lines appeared sharper on the detector, which led to some OH lines appearing like cosmic rays. This meant that some OH lines were then clipped in the old clean cosmic module in the pipeline.

Work was started before the hackathon to improve the clean cosmic module. A new clean cosmic module was made, and some initial testing on this module was carried out. Mike Fitzgerald took real cosmic rays from darks taken by the UCLA Galactic Center Group (GCC). These cosmic rays were then injected into a real sky file, also from the GCC. One sky file contained both natural cosmic rays and injected cosmic rays. Multiple sky files were made. These files provided the test data to test the cosmic ray module.

Devin Chu ran the sky data through the pipeline and compared the reduced data with the old and new cosmic ray modules. Initial results confirmed the clipping effects on OH lines from the old module. At the time, the new module was not removing cosmic rays, but it was keeping OH lines intact. (test plots and data can be found here: https://drive.google.com/open?id=0B7nvYZ_iBsXoMklzUFMxWVJSc2M)



This hackathon provided an opportunity to conduct more testing. The first approach was to take the new clean cosmic module and make the cut harsher to look for more cosmic rays. Different parameters within the clean cosmic modules to find more cosmic rays. However, the module was still not successfully picking up more cosmic rays, and the clipping of the sky lines appeared to be even worse. After this work, we decided to change the algorithm and also consider using lacosmic, another cosmic ray removal code.

Jamie Ryan has been developing a machine learning code to find cosmic rays.

Devin Chu worked with the idl version of lacosmic. The code lacosmic has tuneable parameters, and combinations were made to find the cosmic rays in the data. This process took some fine tuning to find cosmic rays. Ultimately cosmic rays were located as the cuts became harsher, but it also began to affect the OH lines. More work could potentially be explored here, if desired.

James made some changes to the clean cosmic module. This module was successfully finding real cosmic rays, although most of the injected cosmic rays from Mike appeared were not found. A more quantitative measure can be done on this. This new module kept the sky lines intact, especially compared to the old cosmic ray module.

This new module was then merged into the OSIRIS pipeline on github. More work can be done on this make more aggressive cuts, without affecting the OH lines. Additionally, James discussed changing the algorithm to search a line of pixels, rather than a square around the pixel.

7 2D Modeling of OSIRIS spectra - toward 2D extraction

Our goal is to model OSIRIS raw spectra on the detector in order to investigate different methods of spectral extraction. To do so we: (1) traced the rectification matrix slices in order to determine how the spectra are dispersed on the 2D detector for a 2015 Kbb 35 mas cube. (2) We translated the mapping of the lenslet positions to slices in the rectification from Jim Lyke into lookup tables that that are easily put into the models. (3) We currently model the traces of the spectra on the detector as Gaussians. (4) A more sophisticated PSF model is being developed.