





Jonathan McDowell (CXC Science Data Systems)





Chandra data and software

I will report on CIAO (user software) and the standard processing pipeline software, which come into being thanks to:

CXC Data Systems team

software design, development, operation, archive, etc.

CXC Science Data Systems team requirements, documentation, testing, helpdesk, interface with the science community





SAO Scientists:

Jonathan McDowellSDS lead, data model, coordinatesAntonella Fruscione (1/2 time)SDS SAO deputy, Docs and Release leadAneta SiemiginowskaSherpa, Astrostatistics, CatalogDoug BurkeScripts, Releases, Infrastructure, Sherpa, VisualizationFrank PriminiCatalog, Photometry, Source Detection, HRC

SAO IT Specialists:

Kenny Glotfelty Nick Lee Bill Joye

Helpdesk, scripts, docs, legacy expertise Helpdesk, scripts, docs ds9

MIT Scientists:

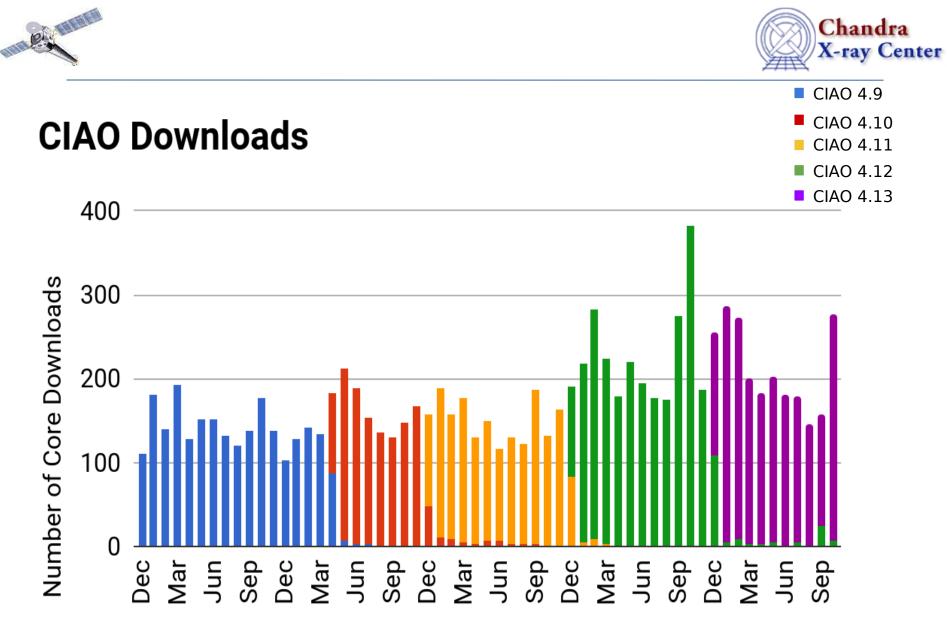
Dave Huenemoerder Moritz Guenther Dave Principe Melania Nynka Gratings, responses MARX, Sherpa Instruments and gratings Instruments and gratings





Community Support:

Downloads, Documentation, Helpdesk

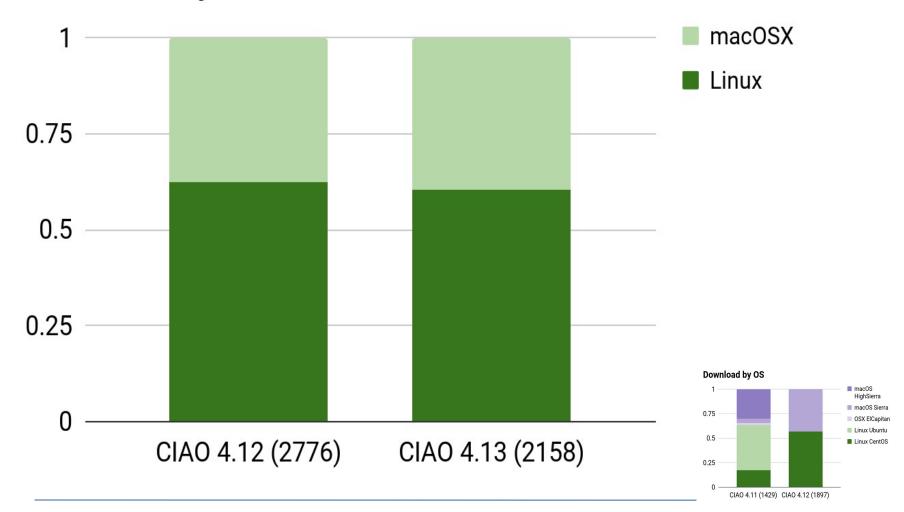


CIAO 4.12 switched from FTP to HTTPS. There are some differences in how the metrics are collected (bot filtering, incomplete d/l, internal d/l's).





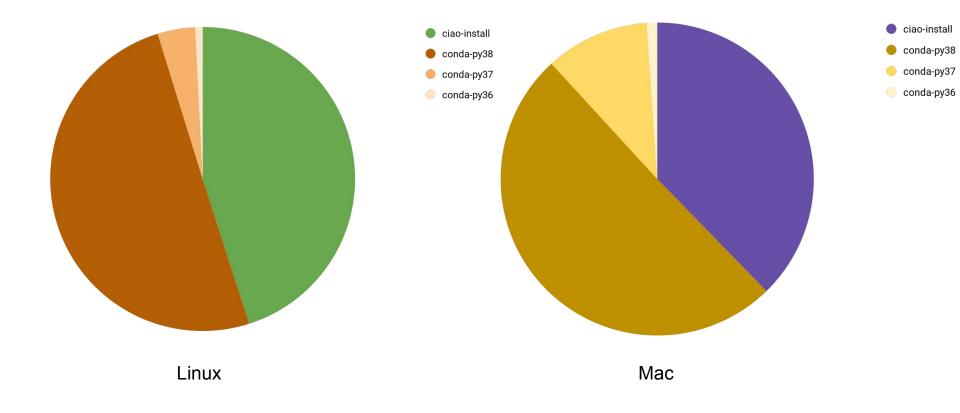
Download by OS







CIAO Installation Options







Downloads (lifetime)

OS	CIAO 4.12	CIAO 4.13
Linux	1734	1304
macOSX	1042	854
	2776	2158
Source	89	32
Total	2865	2190





Documentation





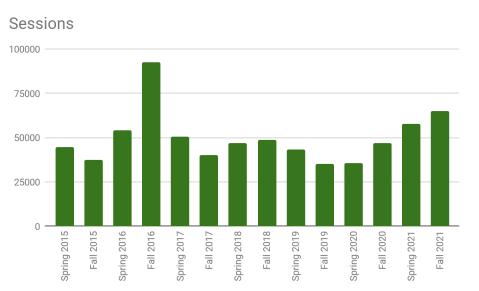
Documentation

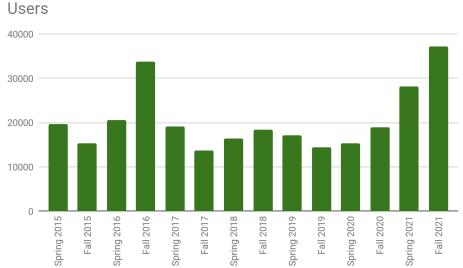
- CIAO 4.13 routine roll-out and updates
- Reworked download pages to encourage conda installation method.
 - Simplified conda download page, moved detailed information into a new conda installation thread.
- New multi-obi srcflux thread
- Posted a new caveat discussing subpixel resolution limitations due to dither.
- Many thread updates to match Repro-5 data products: file names, as well as physical coordinates.



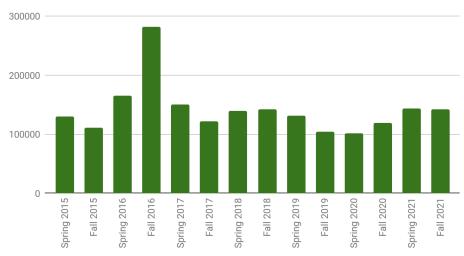


Long Term Website Trends





Page Views







caveat: dither effect on subpixel resolution

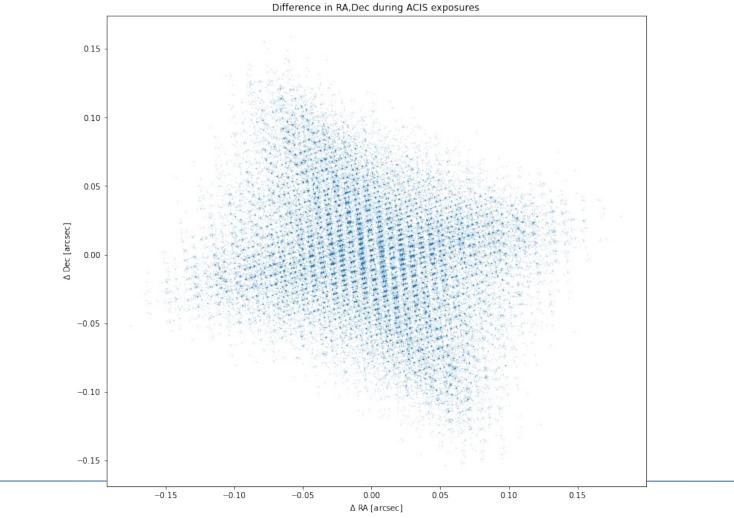
- The distance that Chandra dithers in a typical ~3sec exposure is small compared to the native ACIS pixel size (0.492 arcsec).
- However, users trying to exploit the highest possible spatial resolution using subpixel binning need to be cautious about dither on these scales.
- This caveat helps to quantify at what spatial scale users need to be concerned about dither effects

https://cxc.harvard.edu/ciao/caveats/acis_subpixel_dither.html





Uncertainty in RA/Dec due to ACIS exposure time







Bugs: ACIS frame store shadow bits

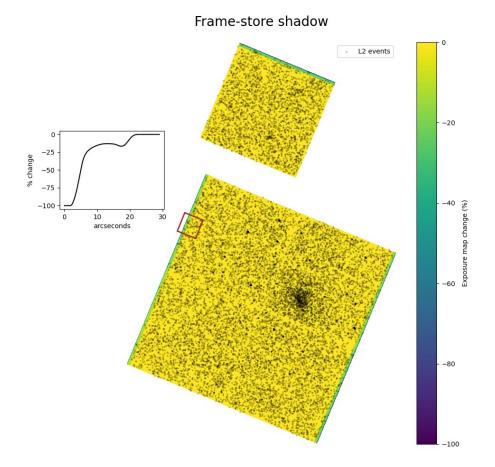
- Bottom few rows of each chip are now marked as bad using new (CIAO 4.12) frame store shadow bit (17).
- However, the response tools (specifically ardlib) was not updated to recognize the new bit.
- So while events have been filtered out, the responses are not correct.
- ciao-contrib 4.13.1 includes updates to scripts to correct this. Permanent fix in ardlib included in CIAO 4.14.

https://cxc.harvard.edu/ciao/caveats/acis_shadow_badpix.html





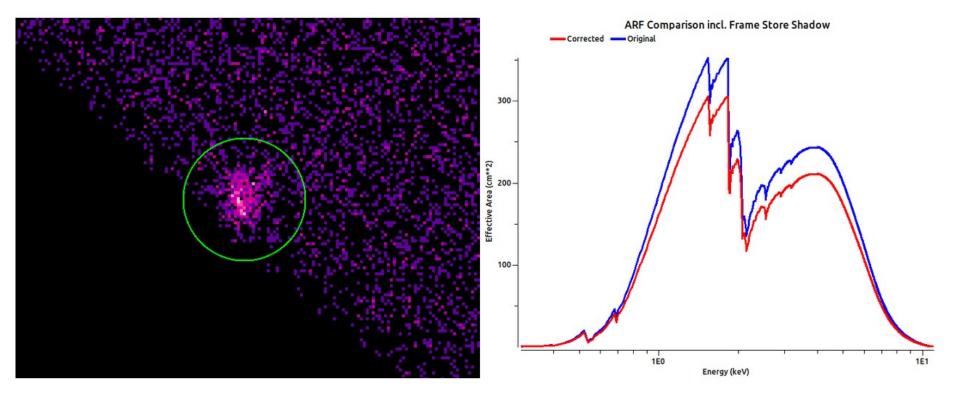
Exposure Map Example







ARF Example







Helpdesk





Helpdesk Spam Attack

- Late April 2021 CXC helpdesk was target of spam attack.
 - Malicious actor(s) submitted dozens of tickets with PDF attachments which contained inappropriate material via the helpdesk web interface.
 - User then submitted the URL to these PDF files to Google search engine.
 - Coming from "harvard.edu" URLs have higher search ranking.
- Helpdesk was taken offline and tickets deleted, but remained in Google's search cache.
 - Harvard IT and CXC helpdesk contacted by several users reporting spam material looking to claim Harvard's "bug bounty" (does not apply to CfA sites).
- Helpdesk website remained off line for ~1 month while new version of helpdesk software, osTicket, was installed and configured.
 - Users were still able to contact helpdesk via email.
 - New version allows for additional configuration options, including filtering file type (PDFs are no longer allowed).
 - Metrics derived from new system are different than before.





Helpdesk Stats

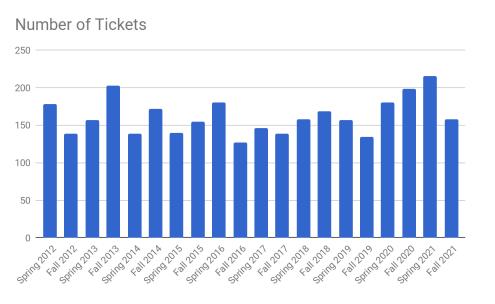
	2019-11-01 to 2020-10-31	2020-11-01 to 2021-10-31
Time period [months]	12	12
Number of Tickets	366	317
Median time to 1st contact [hrs]	1.22	1.53
Median time to close [hrs]	7.42	11.83
Maximum time to close [hrs]	2015.8	978.3

Longest ticket was for user who took a month to follow up on questions regarding fitting 2D PSF with sherpa.

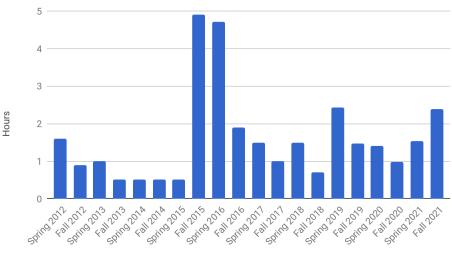




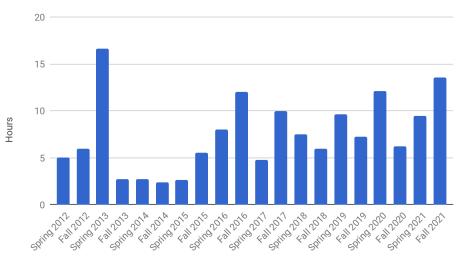
Long Term Helpdesk Trends



Time To Answer

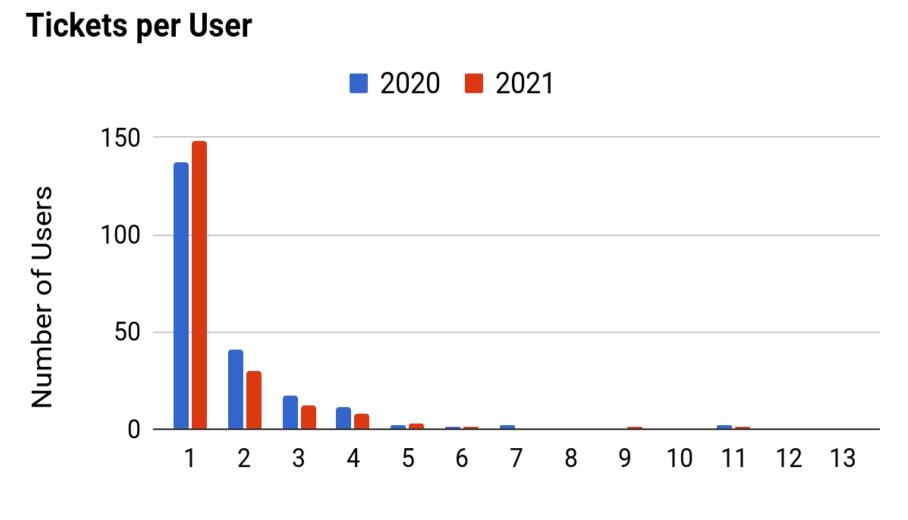


Time To Close







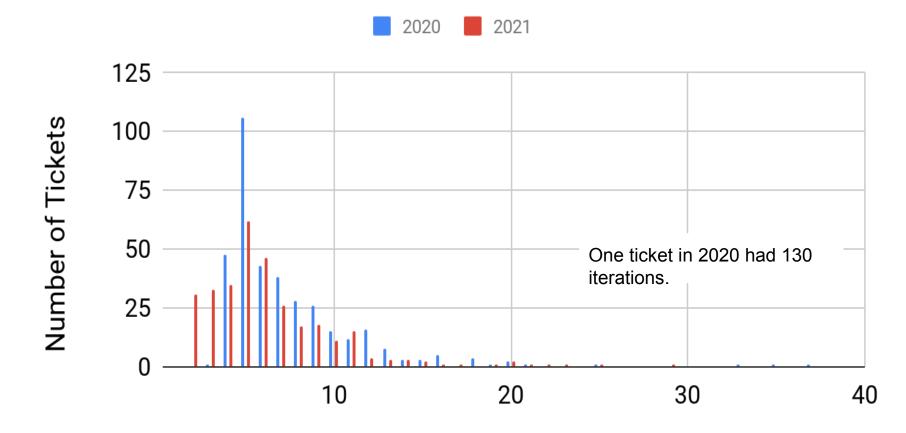


Number of Tickets





Iterations per Ticket



The number of iterations calculation changed in May 2021 with upgraded Helpdesk software.

Iterations





Helpdesk

• Examples of bugs

- Failure to include frame-store-shadow bad pixel bits in response files (exposure maps and ARFs).
- vtpdetect bug with negative coordinates in a large mosaic.
- Problem using pie() shaped regions with angles > 360°
- specextract error creating asphist (corner case)
- Problem running marx with asolfiles updated by asp_offaxis_corr
- dmextract using incorrect GTI when making lightcurves (last GTI instead of 1st)
- Coordinate issue when applying vignetting corrections in some tools including mkinstmap, mkarf, and mkgarf

• Example of documentation updates

- How to do an offline conda installation (CIAO files downloaded from alternative site)
- Platform support updates w.r.t. macOS 11 (Big Sur) and M1 (ARM64)
- New caveat for frame-store-shadow bug





Community





Community Outreach

- AAS 237 January 2021
 - (next slide)
- AAS 238 June 2021
 - Supported Chandra booth
 - Live Helpdesk
 - Presented ds9/dax demos during the "afternoon" breaks
- 2021 Chandra Data Science Symposium
 - CIAO well represented during science talks
 - Presented demos during special session
 - ds9/dax <u>https://youtu.be/bvoaxJkHec0</u>
 - sherpa <u>https://youtu.be/sZ7mEx142HQ</u>





CIAO Workshop #19 at Virtual AAS237 Jan 2021

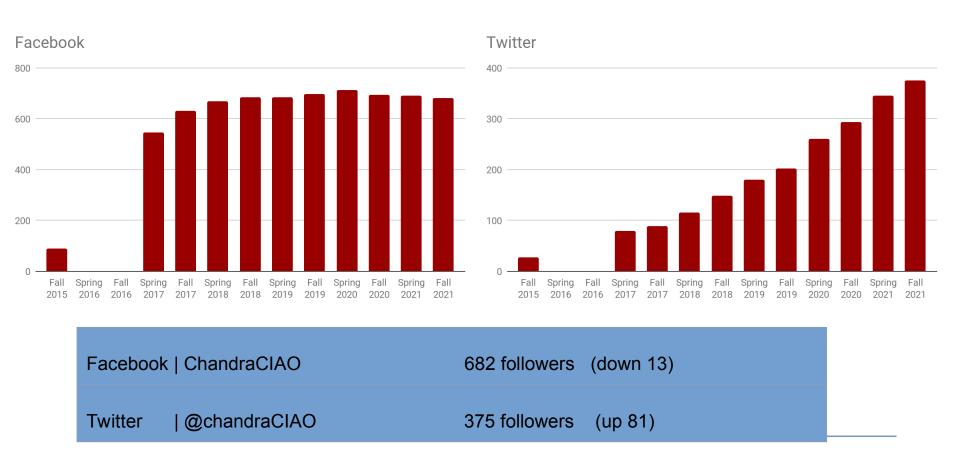


- Two day workshop before start of AAS
- Morning talks were well attended (~20 students). There was much less interest in the afternoon hands on sessions.
- Supported the Chandra booth during AAS meeting breaks.





Long Term Social Media Trends







CIAO 4.13-4.14 and Scripts Overview





CIAO 4.13 Highlights

- Primarily a bug fix release.
- Deprecated graphical applications
 - peg : parameter file editor
 - taskmonitor : run tasks in background and capture logs (stdout & stderr)
- Sherpa updates include
 - Improved Jupyter notebook integration (rich display elements)
- Included new asp_offaxis_corr tool used in Repro-5 to apply boresight corrections to aspect solution.
- wcs_match from CSC to support interactive source cross matching.
- Updates to srcflux to compute fluxes from merged|multiple observations.





CIAO 4.14 Highlights

- Bug fixes
 - ACIS frame store shadow badpixel bits
 - dmextract GTI issue (last vs. first)
 - removed obsolete "0 length polygon" warning
- Withdraw prism, the interactive file browser.
 - This is the last GTK application which greatly reduces OTS burden.
 - ds9 added prism functionality
- Sherpa updates described later in talk
- New adaptive binning tool: dmradar
- Updated support for scaled images and NULL pixel handling
- dmdiff improvements to array and WCS handling. Major code overhaul.



0.089

CUC Nov 2021

0.27

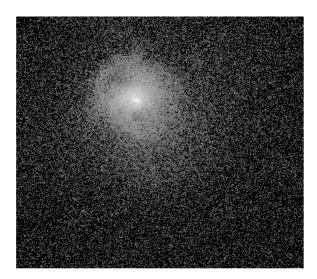
0.63

1.3



dmnautilus upgrade

- Current algorithm uses SNR as an **upper** limit (blocks divided **until** SNR falls below threshold).
- Updated algorithm allow SNR to be a **lower** limit (blocks divide **unless** the SNR falls below threshold).



2.8

5.6

11

23

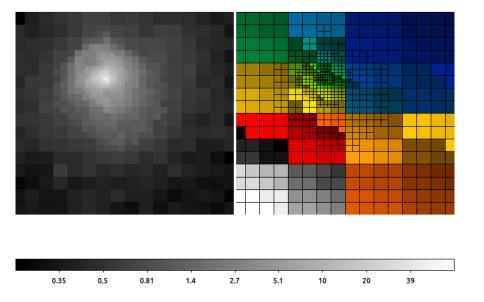
45

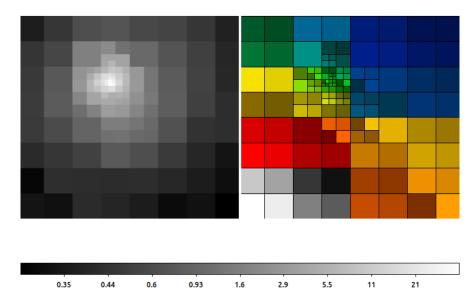
Example: HCG62 merged dataset, broad band, point sources have been removed.





dmnautilus example





Original algorithm. SNR=15, so each bin has at most 225 counts. (Left is binned image, right is map showing pixels belonging to same grid.)

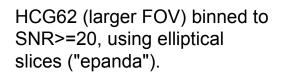
Updated algorithm. SNR=15, so each bin has at least 225 counts.

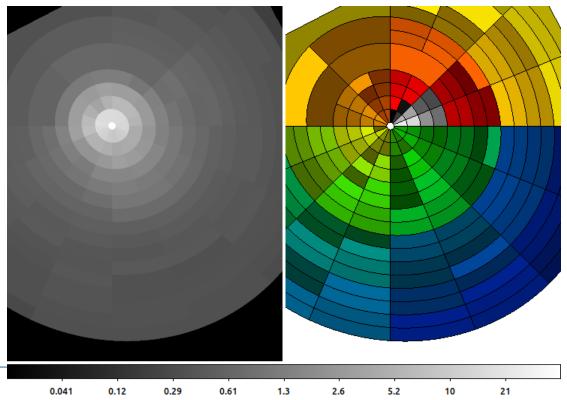




dmradar prototype

- Delivered spec and prototype for new dmradar tool.
- Similar "quad-tree" algorithm as dmnautilus, but instead of diving in cartesian X & Y, it divides in polar coordinates : radius and angle.
- Example:



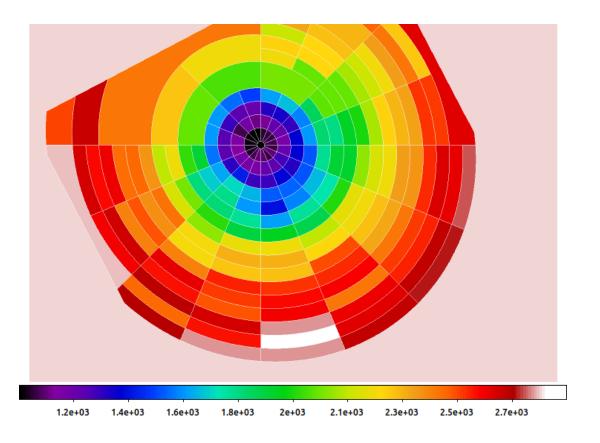






Simple use case

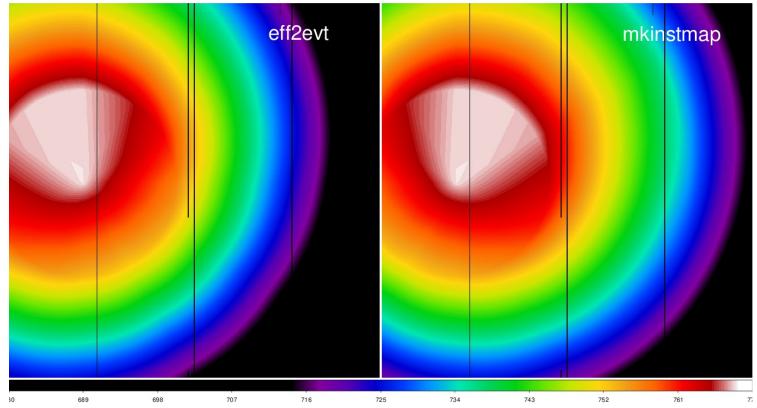
The adaptive binning output can then be used for things such as creating temperature maps, or as shown here a much more simple application which is just the mean energy (in units of eV) of the events in each of the bins.







Bugs: Coordinate sign error

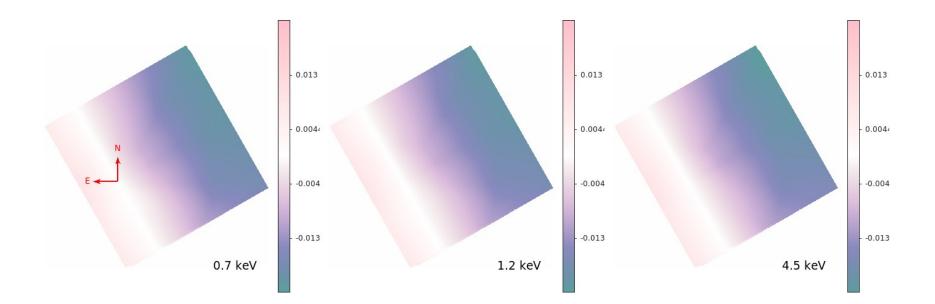


A local user reported an issue when looking at the Effective Area computed by different tools: eff2evt and mkinstmap. There is a rotation and flip of the vignetting component which has been tracked back to a sign error when computing the azimuthal angle, ϕ , used by mkinstmap, mkarf, and mkgarf. The image shows the effective area in Chip coordinates.





Effect on Exposure Maps



The vignetting component is energy dependent and nearly sinusoidal in ϕ . This image shows the fractional difference between the exposure map at different energies before and after the bug fix. The error is small, generally between -1.5 to 1.5%.

CAL team is investigating more deeply since contamination and HRC QE/U are derived from these tools.





Contributed Scripts

- CIAO 4.13 Updates
 - 4.13.1
 - Implemented work around for ACIS framestore shadow issue
 - dax: corrected energy limits when estimating fluxes from fit results
 - 4.13.2
 - specextract: fixes issue when using combine=yes with blank sky background files
 - 4.13.3
 - chandra_repro updated in coordination with CALDB to skip HRC-S/LETG PI filter for recently acquired data (high voltage gain uncertainty).
- CIAO 4.14
 - Rewrite of specextract including support for parallel processing
 - Color coded error messages (can be disabled)
 - Compatibility updates with CIAO and Sherpa 4.14.0



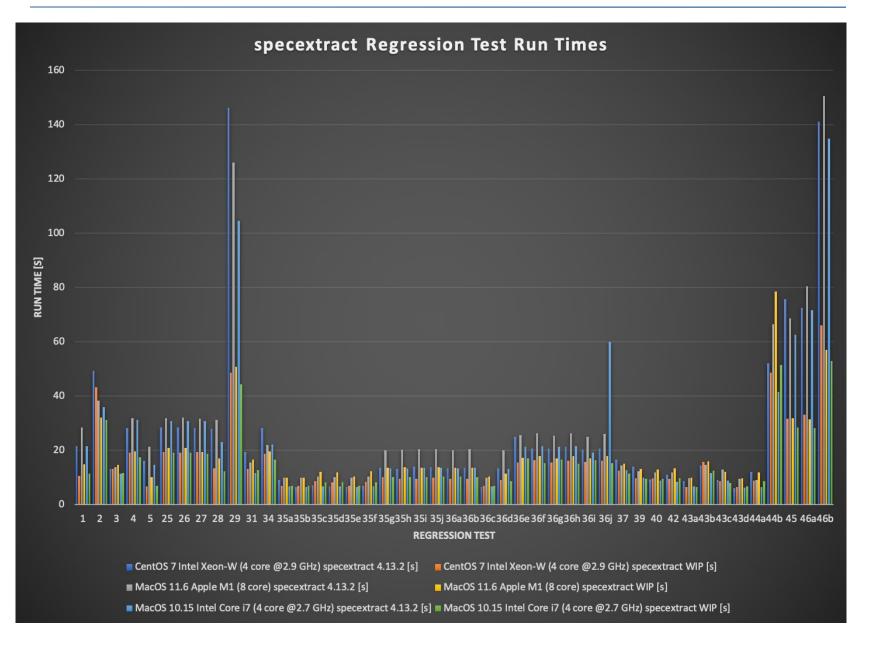


specextract rewrite

- changes transparent to user, transition to new version should be seamless
 - no changes in syntax usage
 - no changes to result
- code restructured and written to be more Pythonic for easier maintainability
- improve run time performance
 - parallelization where possible
 - front-load parameter and file checking
 - error out early
- undergoing beta testing; intended to be ready for CIAO 4.14 release

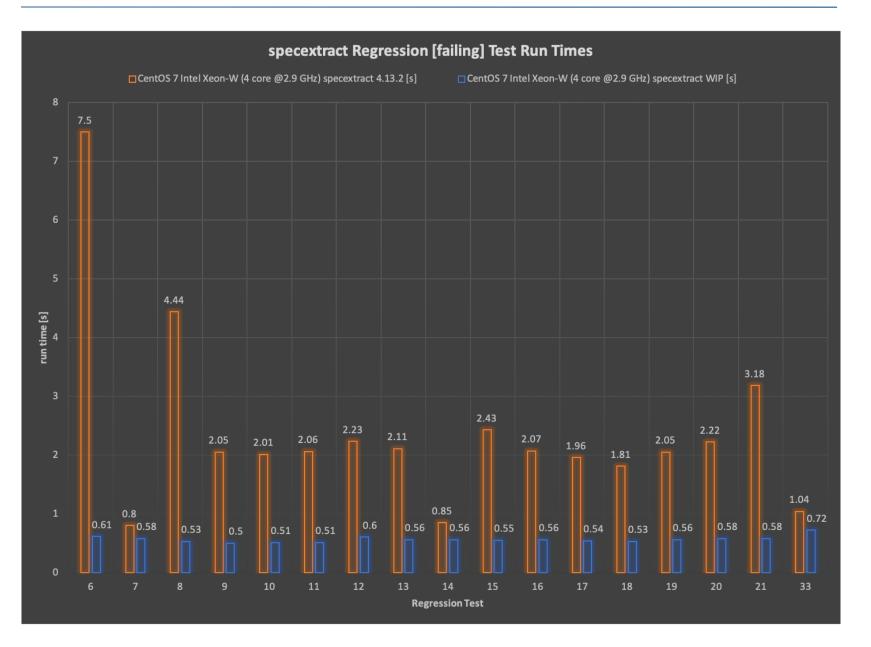
















Common use case: extract source and background spectrum, generate weighted responses for both







Catalog Support





SDS supports the catalog at 2.0 FTE level (Primini, Burke, Lee, with some additional relevant work by others)

Details of progress in I Evans presentation

Primini:

- Statistical Characterization of the catalog
- Aperture Photometry
- Quality assurance and reviews on data, detect list and data products

Burke

- convex hull support,
- quality assurance/reviews of source properties pipeline
- visualization (catalog display in WWT)

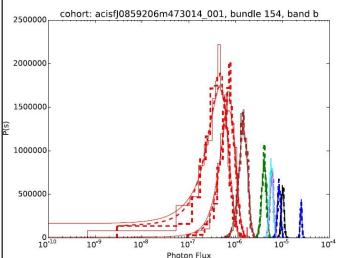
Siemiginowska

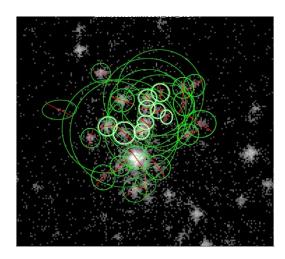
- fitting

Nynka

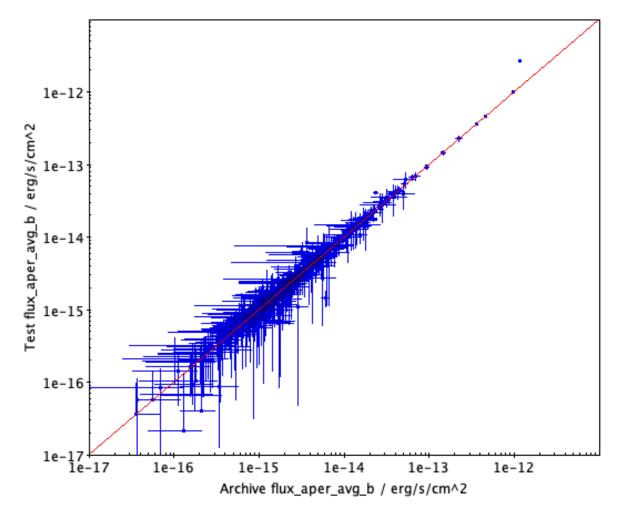
- simulations

Burke, McDowell, Lee, Primini, Siemiginowska - Documentation.





Testing CSC 2.1 Aperture Photometry



A test CSC 2.1 catalog was cross-matched with CSC 2.0, and fluxes of matching master sources were compared (for clarity, every 10th point from a sample of ~6600 matches are plotted). The results are in good agreement. In this sample, the use of pymc3 when CSC 2.0 photometry algorithms failed resulted in an increase of ~7% in the number of valid master source fluxes.





Sherpa





- Sherpa releases during the last year:
 - 4.12.2 on October 27, 2020.
 - CIAO version was released on Dec.15, 2020; It contained the code introduced in the Sherpa standalone during 2020 development year and released in 4.12.2.
 - 4.13.0 was released on Jan.8, 2021 to align the CIAO and Sherpa version numbers. This release contained a few documentation updates and infrastructure changes to migrate Travis-CI to GitHub Actions for integration testing.
 - 4.13.1 May 18, 2021
 - 4.14.0 October 7, 2021 this is the version which will be included in CIAO release in December 2021.

Jupyter notebooks for the Chandra Data Science Workshop:

https://github.com/cxcsds/sherpa-notebooks -chandra-data-science-august-2021

Sherpa Development 2021



- Release Highlights:
 - enhancements:
 - filtering and grouping area for binned spectral data
 - updates to allow users to change the hard limits of XSPEC model parameters
 - the sample_flux routine now returns correct information for the clip column
 - documentation changes:
 - improved PHA simulation documentation
 - improved Filtering and grouping of PHA data documentation
 - added sherpa.image module documentation
 - updates documentation for TableModel, Notice2D, cache support for evaluating models, and low level optimization code
 - Jupyter notebook updates
 - Infrastructure changes:
 - the github master branch has been migrated from master to main
 - updates to support Apple ARM
 - update to support Xspec version 12.12
 - update fftw from version 3.3.8 to 3.3.9
 - changes to support gcc 9.3.0 in conda defaults
 - updates to support python 3.9 including readline 8.1 upgrade, numpy minimum
 1.19 (numpy 1.18 minimum for python 3.7/8)
 - test infrastructure clean up and updates

Sherpa Development 2021



- bug fixes:
 - updates to fix several 'unable to parse region string: None' errors
 - fix issue where save_all() of a loaded image with no region filter would fail on reload
 - fixed issue with plot_model() being called before notice or ignore could lead to filters not getting applied
 - fix to error out instead of crash when grouping data using an unsupported method

Full Release Notes on https://github.com/sherpa/sherpa/releases



Sherpa Documentation



CIAO

Standalone Python Package

CIAO ahelp files autogenerated from python docstrings

sherpa.ui.set_source(id, model=None)

Set the source model expression for a data set.

The function is available as both set_model and set_source. The model fit to the data can be further modified by instrument responses which can be set explicitly - e.g. by set_psf - or be defined automatically by the type of data being used (e.g. the ARF and RMF of a PHA data set). The set_full_model command can be used to explicitly include the instrument response if necessary.

Parameters

- id (int or str, optional) The data set containing the source expression. If not given then
 the default identifier is used, as returned by get_default_id.
- model (str or sherpa.models.Model object) This defines the model used to fit the data. It
 can be a Python expression or a string version of it.

See also

delete_model

Delete the model expression from a data set.

Fit one or more data sets.

freeze

fit

Fix model parameters so they are not changed by a fit.

get_source

Return the source model expression for a data set.

integrate1d

Integrate 1D source expressions.

set_syserror
set_xlinear
set_xlog
set_ylinear
set_ylog
show_all
show_conf
show_covar
show_data
show_filter
show_fit
show_kernel
show_method
show_model
show_proj
show_psf
show_source
show_stat
simulfit
t_sample
thaw
uniform_sample
unlink
unpack_arrays
unpack_data
The sherpa.astro.ui module
The sherpa.ui.utils module
The sherpa.astro.ui.utils m
EBOOKS
pa Quick Start
book support in Sherpa

O Edit on GitHub

set_staterror

sherpa.ui.set_staterror(id, val=None, fractional=False)

Set the statistical errors on the dependent axis of a data set.

These values over-ride the errors calculated by any statistic, such as chi2gehrels or chi2datavar.

Parameters

- id (int or str, optional) The identifier for the data set to use. If not given then the default identifier is used, as returned by get_default_id.
- val (array or scalar) The systematic error.
- fractional (bool, optional) If False (the default value), then the val parameter is the absolute value, otherwise the val parameter represents the fractional error, so the absolute value is calculated as get_dep() * val (and val must be a scalar).

See als

load_staterror Set the statistical errors on the dependent axis of a data set. Load_syserror Set the systematic errors on the dependent axis of a data set. set_syserror

Set the systematic errors on the dependent axis of a data set.

get_error

Return the errors on the dependent axis of a data set.

Note

https://sherpa.readthedocs.io/en/latest/NotebookSupport.html#

CIAO 4.12

Edit	View Insert Cell Kernel Widgets Help	Not Trusted	Python 3 C					
F BC C	B B + + H Ban B C + Matchan							
	Notebook support in Sherpa							
	A rumber of objects have been updated to support HTML output when displayed in a Jupyter notebook. Let's take a quick tour!							
	Data1D, Data1Dint, and Data2D							
	First we have the data objects:							
In [1]:	import numpy as mp							
	from sherpa.data import Data1D, Data1DInt, Data2D							
	x = np.arange(100, 200, 20)							
	y = np.asarray([120, 240, 30, 95, 130])							
	<pre>dl = DataID('oned', x, y) dli = DataIDInt('onedint', x(z-1), x((z), y(z-1))</pre>							
	x0 = np.amarray([150, 250, 100]) x1 = np.amarray([250, 200, 200])							
	y2 = np.asarray[[50, 40, 70]] d2 = Data20['twod', s0, s], y2]							
	Each can be displayed with print, which shows a textual representation of attribute and values:							
Te (2).	print(d1)							
TH (x).	tate = oted							
	x = Int64(5)							
	y = Int64[5] staterror = None							
	syserror = None							
	Or they can be displayed as-is which, in a Jupyler notebook, will display either a plot or a HTML table. The Data1D and Data1DInt classes will dislay a							
	for any car of outplayed area which, in a degree interest, we capely term a part of a risk, done. The database and database plot (if the pylab plotting backend is selected), and the Database class a table.	the oddere h	/4					
In (3):	d1							

Models and parameters

Models and parameters can also be displayed directly as HTML tables, mirroring their print output.



We can compare the model output (this also works with a single component, such as mgauss and mconst):

In [22]: print(mdl) (gauss2d +

Param	Type	Value	Min	Max	Units
	-75-				
gauss2d.fw	hm thawed	10	1.17549e-38	3.40282e+38	
gauss2d.xp	os thawed	3150	-3.40282e+38	3.40282e+38	
gauss2d.yp	os thawed	4520	-3.40282e+38	3.40282e+38	
gauss2d.el	lip frozen		0 0	0.999	
gauss2d.th	eta frozen		-6.28319	6.28319	radians
gauss2d.am	pl thawed	1	-3.40282e+38	3.40282e+38	
const2d.c0	thawed	1	-3.40282e+38	3.40282e+38	

In [23]: mdl

Out[23]: <BinaryOpModel model instance '(gauss2d + const2d)'>

. . .

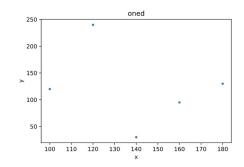
In [24]: print(mgauss.theta)

val	=	0.0
min	=	-6.283185307179586
max	=	6.283185307179586
units	=	radians
frozen	=	True
link	=	None
default_val	=	0.0
default_min	=	-6.283185307179586
default_max	=	6.283185307179586

CIAO 4.13

In [3]: d1

Out[3]: ▼ Data1D Plot



In [15]: mdl

Out[15]: ▼ Model

Expression: gauss2d + const2d

Component Parameter Thawed Value Min Max Units

	fwhm	\checkmark	10.0	TINY	MAX	
	xpos		3150.0	-MAX	MAX	
gauss2d	ypos		4520.0	-MAX	MAX	
yausszu	ellip		0.0	0.0	0.999	
	theta		0.0	-2π	2π	radians
	ampl	1	1.0	-MAX	MAX	
const2d	c0	V	1.0	-MAX	MAX	

In [17]: mgauss.theta

Out[17]: ▼ Parameter

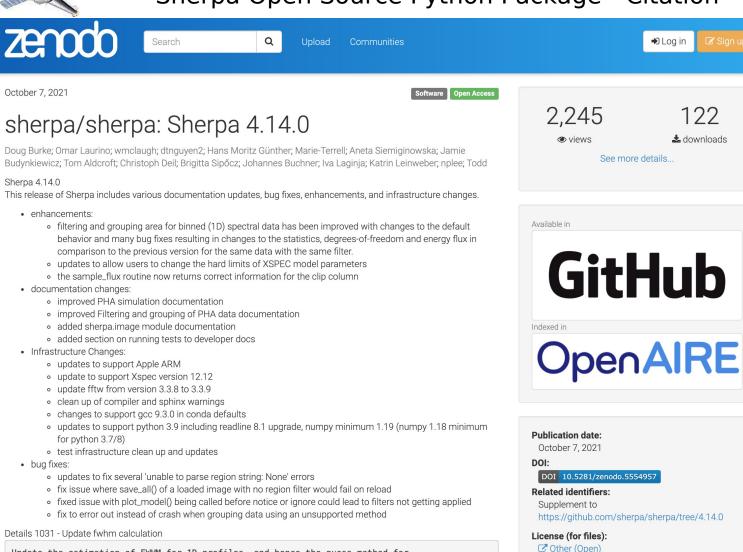
Component	Parameter	Thawed	Value	Min	Max	Units
gauss2d	theta		0.0	-2π	2π	radians

In [25]: mgauss.theta



Sherpa Open Source Python Package - Citation





Versions

Version 4.14.0 10.5281/zenodo.5554957

Oct 7, 2021

Details 1031 - Update fwhm calculation

Update the estimation of FWHM for 1D profiles, and hence the guess method for Gauss1D and related routines. The 2D models use the same routine so see these changes.

1073 - Allow fake_pha to be called with an identifier of None

The fake_pha command now treats id=None as the default id. This addresses #1064.

1106 - The sample_flux routine now returns the correct information for the clip column



~1401 publications in ApJ, AJ, MNRAS and A&A use Sherpa (since 2001 and including astro-ph abstracts)

342 citations to Freeman et al 2001 SPIE paper

91 research papers published in 2020

7 PhD theses listed in ADS that used Sherpa

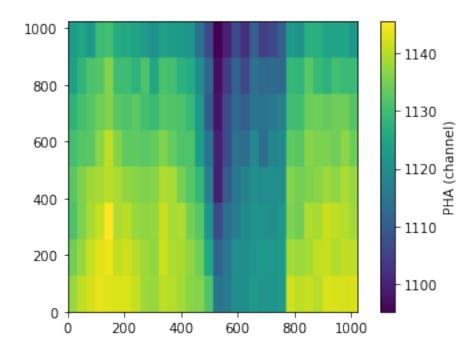




Instruments and Gratings

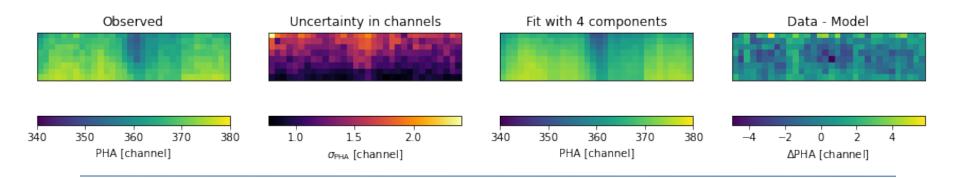






New approach to gain calibration using principal component analysis (PCA)

Gain evolves with time but there are consistent spatial patterns 3 to 5 spatial components found – requires less cal observational data than 256 independent tiles



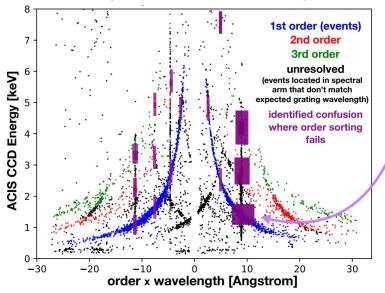




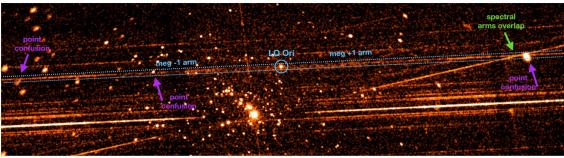
Chandra HETG Confusion Tool: Progress Update

HETG observation of the Orion Nebula Cluster

- HETG observations of fields with multiple X-ray bright sources drastically complicates the analysis of high resolution X-ray spectra.
- Chance alignment of sources in crowded fields cause *confusion* (events in the HETG spectrum of an extracted source belonging to a different object). This tool identifies the location of these *confusing events* using the position of X-ray sources in the observation.
- When fitting spectra, users can then remove (ignore) the small wavelength range associated with confusion instead of removing the entire arm.



LQ Ori Spectral Extraction Order Sorting: MEG



Example Confusion Table for LQ Ori to be used when fitting spectra

obsID	confusing_source_#	Oth_order_confused_cnts	grating_type	order	confusion_type	confusion_wave	wave_low	wave_high
3	697	280.	MEG	1	spectral	2.613	2.51584	2.71869
3	1391	1762.	MEG	1	spectral	8.988	7.17296	11.29768
3	188	1741.	MEG	1	point	8.905	8.748	9.062
3	266	25.	MEG	1	point	4.928	4.891	4.965
3	455	14.	MEG	-1	point	1.218	1.173	1.263
3	516	17.	MEG	-1	point	2.578	2.556	2.6
3	653	243.	MEG	-1	point	4.409	4.377	4.441
3	892	56.	MEG	-1	point	7.571	7.507	7.635
3	1114	564.	MEG	-1	point	11.266	11.18	11.344
3	188	1741.	MEG	1	arm	1.	1.	32.
3	266	25.	MEG	1	arm	1.	1.	32.
3	455	14.	MEG	-1	arm	1.	1.	32.
3	516	17.	MEG	-1	arm	1.	1.	32.
3	653	243.	MEG	-1	arm	1.	1.	32.
3	892	56.	MEG	-1	arm	1.	1.	32.
3	1114	564.	MEG	-1	arm	1.	1.	32.

This HETG confusion tool is designed to identify all primary sources of confusion: (1) a point source landing on a grating arm, (2) when two different source's spectral arms intersect, and (3) when a point source lands on an arm and is bright enough to potentially confuse the entire arm with its dispersed spectrum.

The tool can be run on any number of epochs and produces an easy to read fits table (above) for all objects (or a subset) in the field-of-view identifying where confusion occurs, the source responsible for confusion, and a wavelength range to ignore for spectral fitting. Supplementary products can include an order sorting 'banana plot' (left) overlaid with the affected wavelength range for identified confusion (purple). This tool has been tested using MARX simulations and continues to be tested on the 70 HETG observations of the star-forming Orion Nebula Cluster (individual epoch above).





Other instruments and gratings work:

Repro5 V&V complete

HRC-S time dependent gain map updated

Addressed some minor issues with pipeline grating responses (handling of edge cases such as zero response).

TGCAT grating catalog software port work continuing

Work on gratings with HRC-I, ACIS-I support continues

Order sorting improvements in work





PSF





ChaRT and Marx

- ChaRT pdated to use CIAO 4.13
- ChaRT Routine OS and system upgrades
- Marx routine upgrades





ChaRT

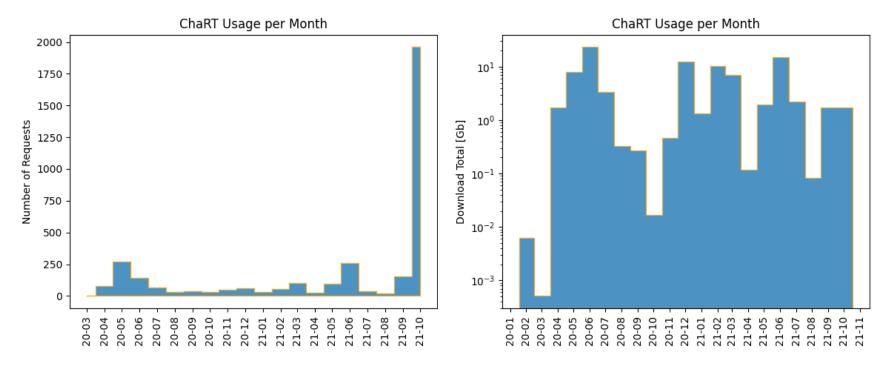


ChaRT server upgraded to CIAO-4.13.

CUC Nov 2021





Visualization





SAOImageDS9

• Releases

- Version 8.2 released in Dec 2020 with CIAO 4.13
- Version 8.2.1 released in Feb 2021
- Version 8.3b1 released in May 2021
- Version 8.3rc released in Oct 2021

• New Features 8.3

- Multiple color bars and graphs, one per frame.
- New notes support, users may add notes to presentation to be saved with backup.
- Added support for user defined Analysis button bars.
- MacOS ARM64 binary support, for both X11 and Aqua ports.
- Imported MatPlotLib colorbars.
- Assorted improvements and enhancements.





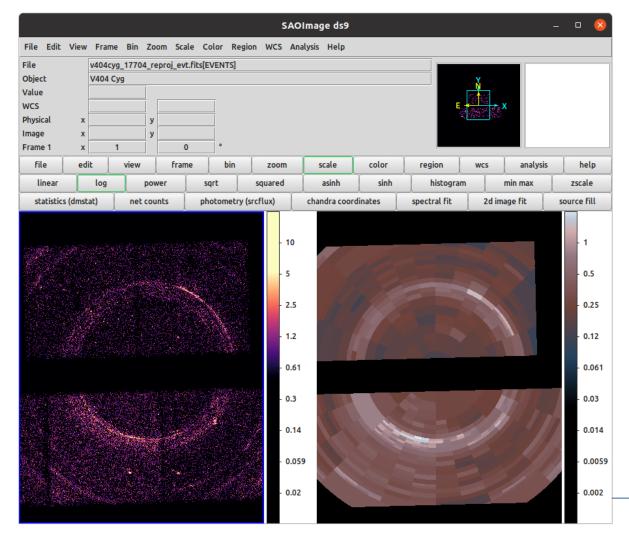
SAOImageDS9 - May 2021 to Oct 2021

- GitHub Activity
 - 213 Commits
 - 25 Release Note Entries
- Help Desk
 - 54 CXC HelpDesk
- Downloads
 - 17847 unique IP





ds9 8.3 + dax 4.14



- New button bar
- Multiple color bars
- Matplotlib colors
- Added dmradar

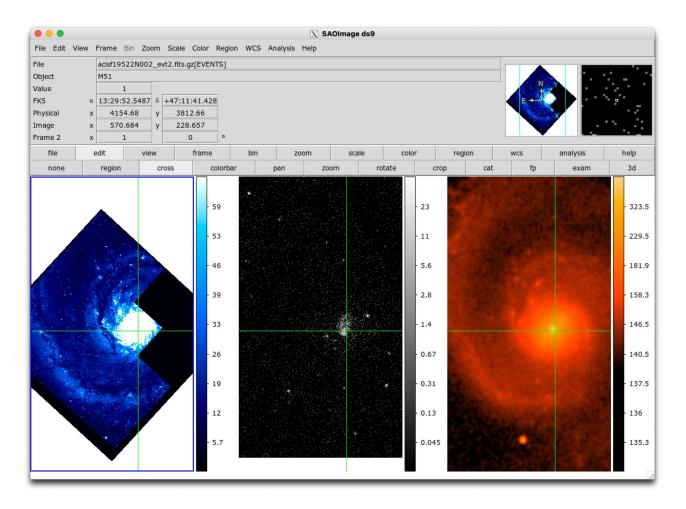
Left: broad band counts image OBSID17704. Right: adaptively binned to w/ snr=5 (25 counts/bin).

CUC Nov 2021





SAOImageDS9 - Multiple Colorbars



CUC Nov 2021 | SDS