



Science Data Systems

Jonathan McDowell



COVID-19: SDS SAO staff hybrid working (typically 2 days/wk remote)

Team: McDowell, Fruscione (1/2), Siemiginowska, Burke

(SAO scientists: CIAO, Sherpa, Docs, HRC, Catalog, User support)

Glotfelty, Lee, Joye

(SAO computer specialists: Docs, User support, scripts, DS9)

Huenemoerder, Guenther, Principe, Nynka

(MIT scientists: Gratings, ACIS, PSF, V&V, Catalog, Sherpa, User support)

Overview:

Ensure the science community can turn data products into science papers:

Define, test and support CIAO - the Chandra user data analysis package

User support for data analysis

Maintain and improve science algorithms, data products

Simplify and codify evolving best practices for analysis (scripts, threads)

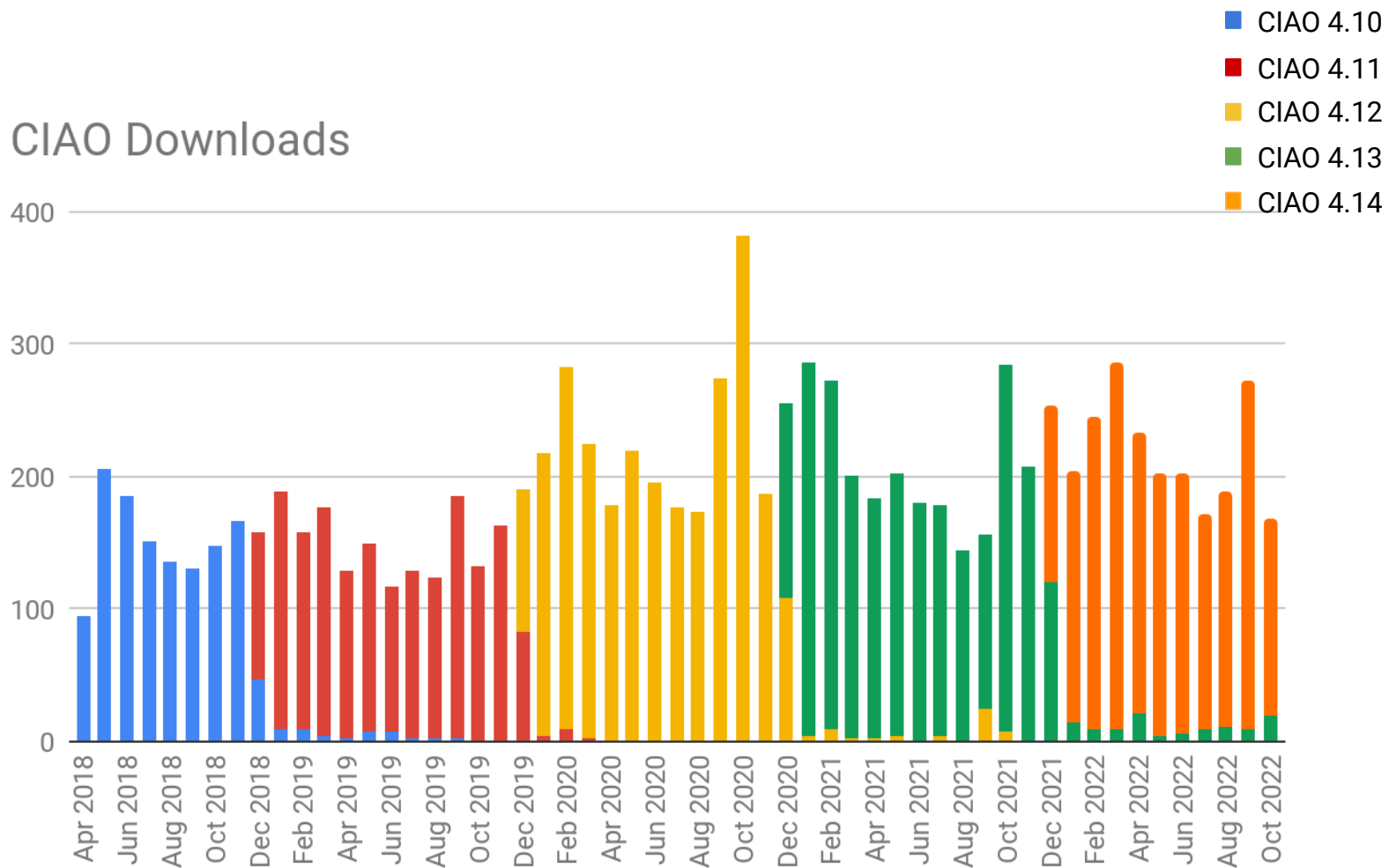


Community Support:

Downloads,
Documentation,
Helpdesk

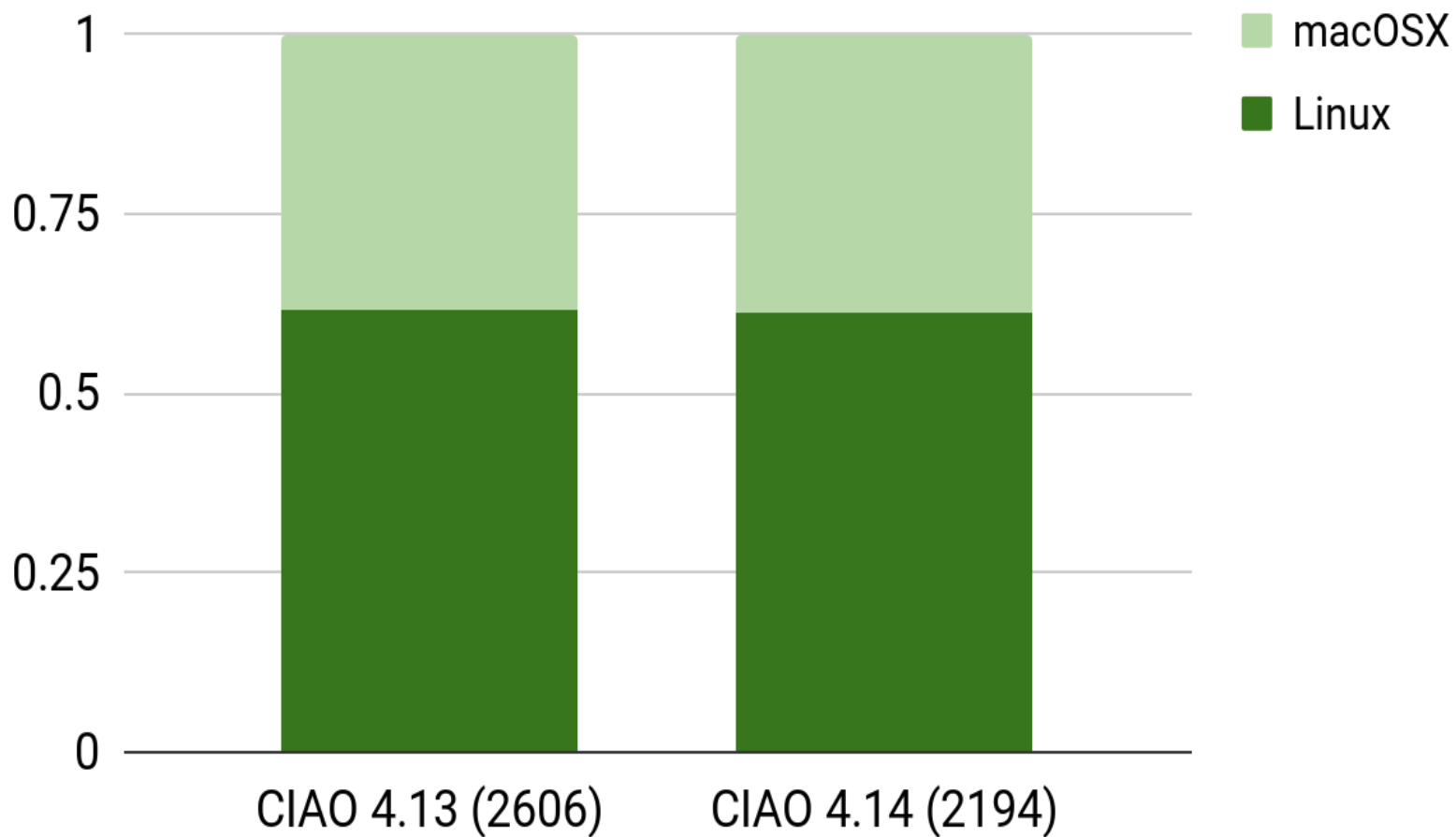


CIAO Downloads



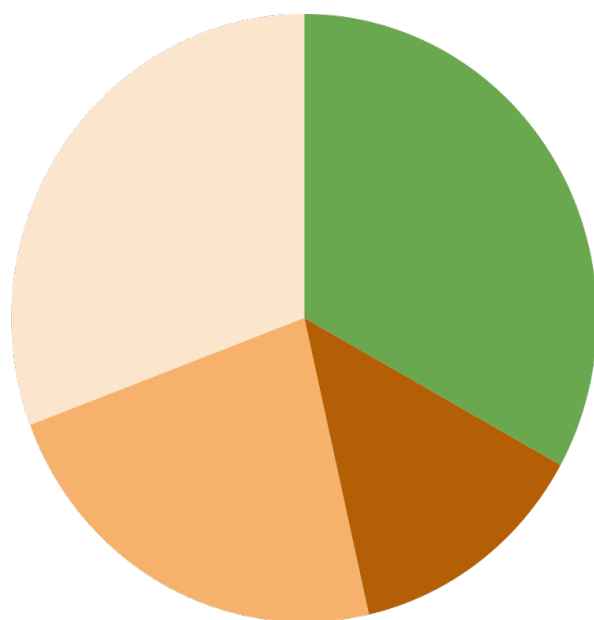


Download by OS

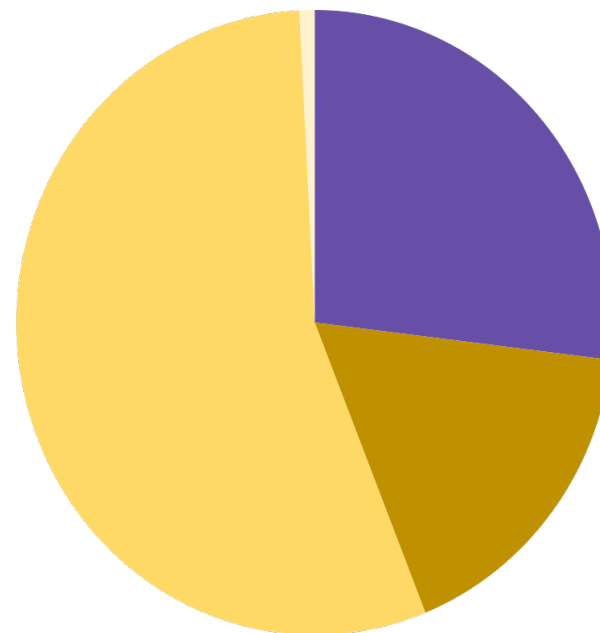
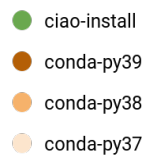




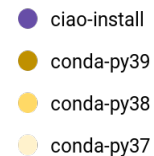
CIAO 4.14 Installation Options – Most recent 6 months



Linux



Mac



The default python version for the conda installations has changed. When released in December the default was 3.8. Currently it is 3.7 on Linux.



Downloads (lifetime)

OS	CIAO 4.13	CIAO 4.14
Linux	1605	1341
macOSX	1001	853
	2606	2194
Source	42	34
Total	2648	2228



Documentation



Documentation

- CIAO 4.14 routine roll-out and updates
 - Update doc threads, help files, release notes
 - Completed review of threads for Repro-5 changes
 - Still a few with early ObsIDs that have not been through Repro-5 yet.
 - Conda installation
 - Several updates related to changing to use **conda-forge** channel to get needed libraries for XSpec.
 - Simplified the download conda page and migrated details and installation options into a new thread.
 - Added information about how to use conda installed MARX
-

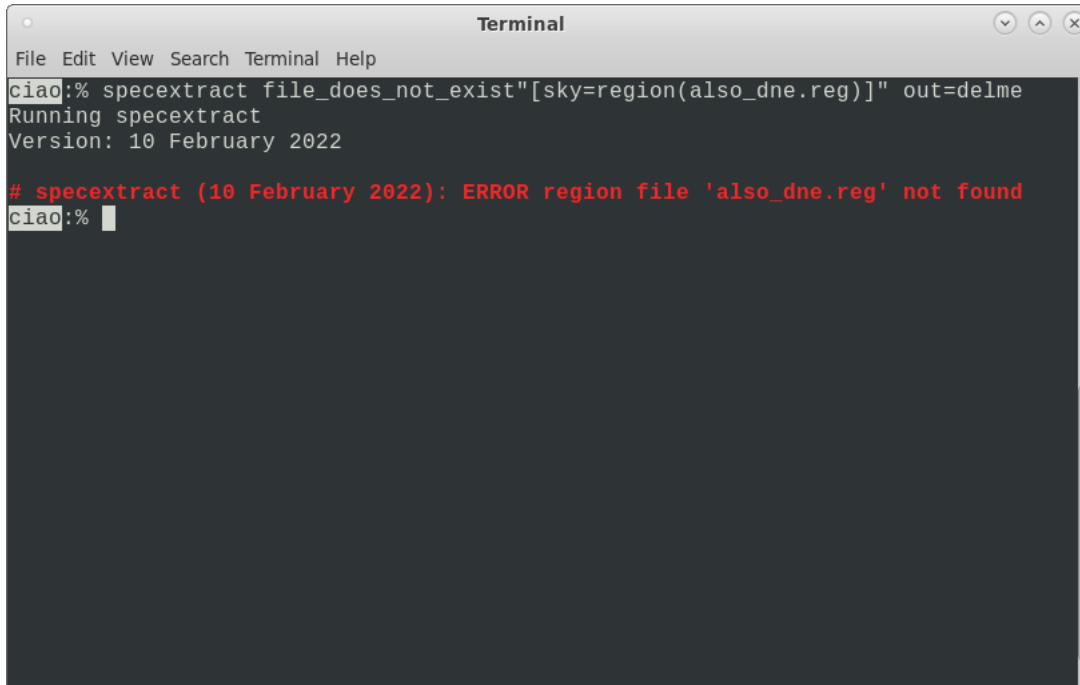


Documentation

- Revising user documentation for merging data to provide updated guidance and more coherent overview
 - New Analysis Guide: Using srcflux for acis_extract users
 - ACIS warm focal plane
 - Adding warning to specextract thread for specific conditions that they may want to filter out, if possible, times when ACIS FP_TEMP is warm.
 - Updating ACIS FP_TEMP filtering page: what was an anomaly will now sometimes become SOP.
 - Updating ACIS FP_TEMP why pages
 - Reviewing doc updates for changes to default dither parameters
-



Improving CIAO User Experience



```
Terminal
File Edit View Search Terminal Help
ciao:% specextract file_does_not_exist"[sky=region(also_dne.reg)]" out=delme
Running specextract
Version: 10 February 2022

# specextract (10 February 2022): ERROR region file 'also_dne.reg' not found
ciao:%
```

- In ciao-contrib-4.14.0 we added "escape sequences" to make error messages **RED**.
 - Users can disable by setting NO_COLOR environment variable.
- This only applies to the contributed scripts.



ahelp – possible enhancement

```
Terminal
File Edit View Search Terminal Help
dmlist tools
SYNOPSIS
List contents or structure of a file.
SYNTAX
dmlist infile opt [outfile] [rows] [cells] [verbose]
DESCRIPTION
dmlist dumps the contents or header of a file or block (a block is a
subfile or FITS extension) to ASCII in an organized way. It corresponds
to the FTOOLS fdump and fstruct programs, but interprets the input file
at a higher level. All CXC data model formats are supported (see "ahelp
dm" [https://cxc.harvard.edu/ciao/ahelp/dm.html] for more
information on the CXC Data Model).
dmlist uses a comma-delimited list of options to select which
information is displayed. Options indicate both what sections of a file
are to be displayed (blocks, keys, comments, cols, subspace, data,
header, struct, full) and in what format (array, all, clean, raw). See
```

sections



syntax

URLs

Also improved lists



docs: ahelp context changes

The [ahelp pages](#) provide the most complete list of all CIAO components: tools, scripts, modules, concepts).

However, since all tools are grouped in a single category, it is difficult to locate related tools. For example the "destreak" tool is not grouped with the other ACIS tools.

CIAO 4.15 will introduce more specific context to better organize the tool list.

dmclist example: the "tools" have been split into 3 categories.

```
Terminal
File Edit View Search Terminal Help
SEE ALSO
-----
CONTEXT | SUBJECTS
-----
concept | subspace
dm      | dm, dmascii, dmfiltering, dmopt
tools::coordinates | dmcoords
tools::core | dmcopy, dmextract
tools::statistics | dmstat
-----
LAST MODIFIED
December 2022
(END)
```



Merging Central (WIP)

Merging Central
Introduction to Merging Chandra Data
Data Preparation & Merging
Final Merging Analysis
Reviewing Astronomy for Imaging & Spatial Analysis
Combining Datasets for Imaging & Spatial Analysis
Combining Datasets for Spectral Analysis
Timing Analysis Across Multiple Observations

CXC Links
CXC
Science
CXC

Merging Central

Introduction to Merging Chandra Data

Spatial/Imaging Analysis **Spectral Analysis** **Timing Analysis**

Chandra observations can be split in segments or can cover different time spans for two main reasons:

- Scientific reasons:** the same target or patch of the sky has been observed many times during the course of the mission
- Engineering reasons:** because of spacecraft thermal restrictions, long observations are broken into shorter exposures. As of late 2021, observations longer than about 60ks are usually broken up into 30ks segments for planning purposes, and observations of these segments may be separated by significant periods of time Proposers' Observatory Guide [63.3.3](#)

Whatever the reasons, observers may want to merge the various observations for example to:

- detect faint sources
- study variability across time
- cover a large area
- recover the unsplit exposure time

The direction in which to go when merging observations is strongly dependent on several factors. The main ones are:

- the scientific goal
- the instrument used
- the time separation between observations
- the spatial separation between observations

It is also crucial to understand what is possible or is not possible to do with the merged output once observations are combined and the limitations therein.

The aim of this website is to provide instructions and guidance in all aspects of merging observations.

Note
'Merging', 'combining', and 'co-adding' are all synonymous in the context of this website.

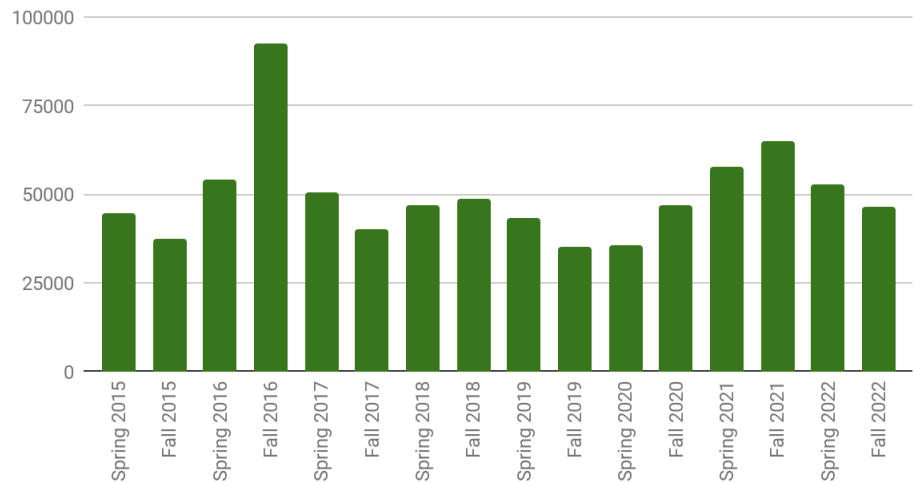
Merging Overview: Analysis-Dependent Paths

- Merging should only be done when necessary
 - broad overview and reprocessing guide
 - repository of existing documentation
- Provide real-world examples
 - drawn from user-inquiries via helpdesk and workshops
 - organized on generalized analysis theme
- Demonstrate limitations of combined data sets
- Upshot: combining datasets occurs at the end of processing, regardless of science interest

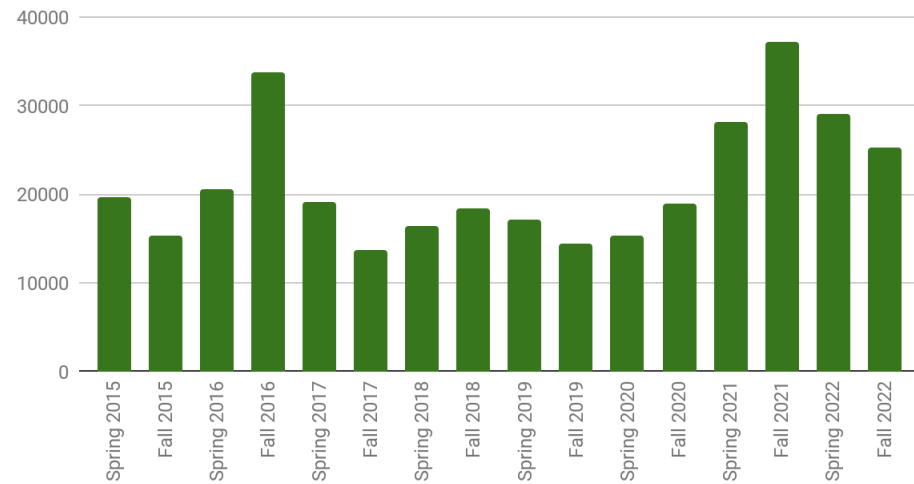


Long Term Website Trends

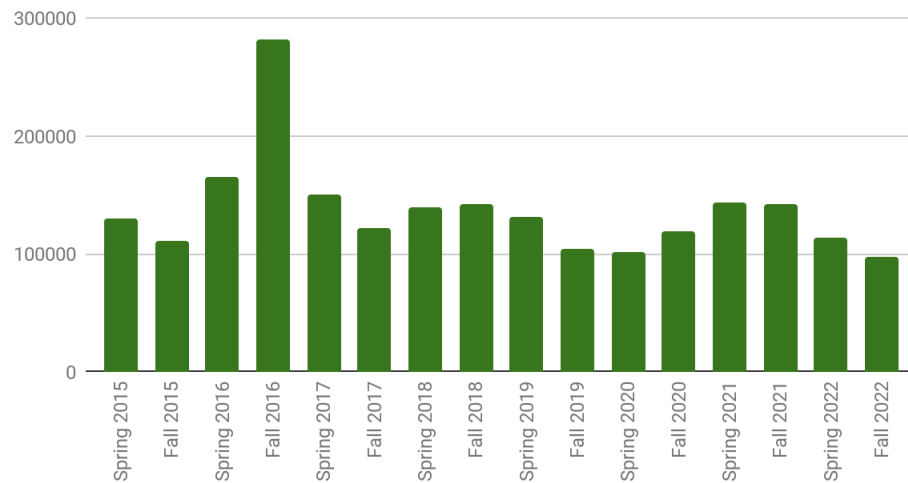
Sessions



Users



Page Views





Special Topic:
ACIS_EXTRACT and *srcflux*



Background

“ACIS_EXTRACT” is a suite of programs developed by PSU members of the ACIS IPI team, optimized for automated analysis of fields with many point sources.

It includes a set of (non-free-software) IDL scripts which run CIAO tools for most of the actual analysis.

There is concern from the community about long term support for ACIS_EXTRACT given the planned retirement of core developers.

What should those users do now?



CUC request re acis_extract

"The CUC requests that the CXC review the possibilities of creating a tool or thread recommending procedures equivalent to those performed by [acis extract]." - Chandra Users Committee Report, May 12, 2022

SDS has reviewed the **basic** functionality of acis_extract and finds that the currently available [srcflux](#) script already performs most of the same reductions and analysis in a comparable amount of time.

Reference threads:

- [Calculate source count rates and fluxes](#)
- [Calculate source count rates and fluxes for combined datasets](#)



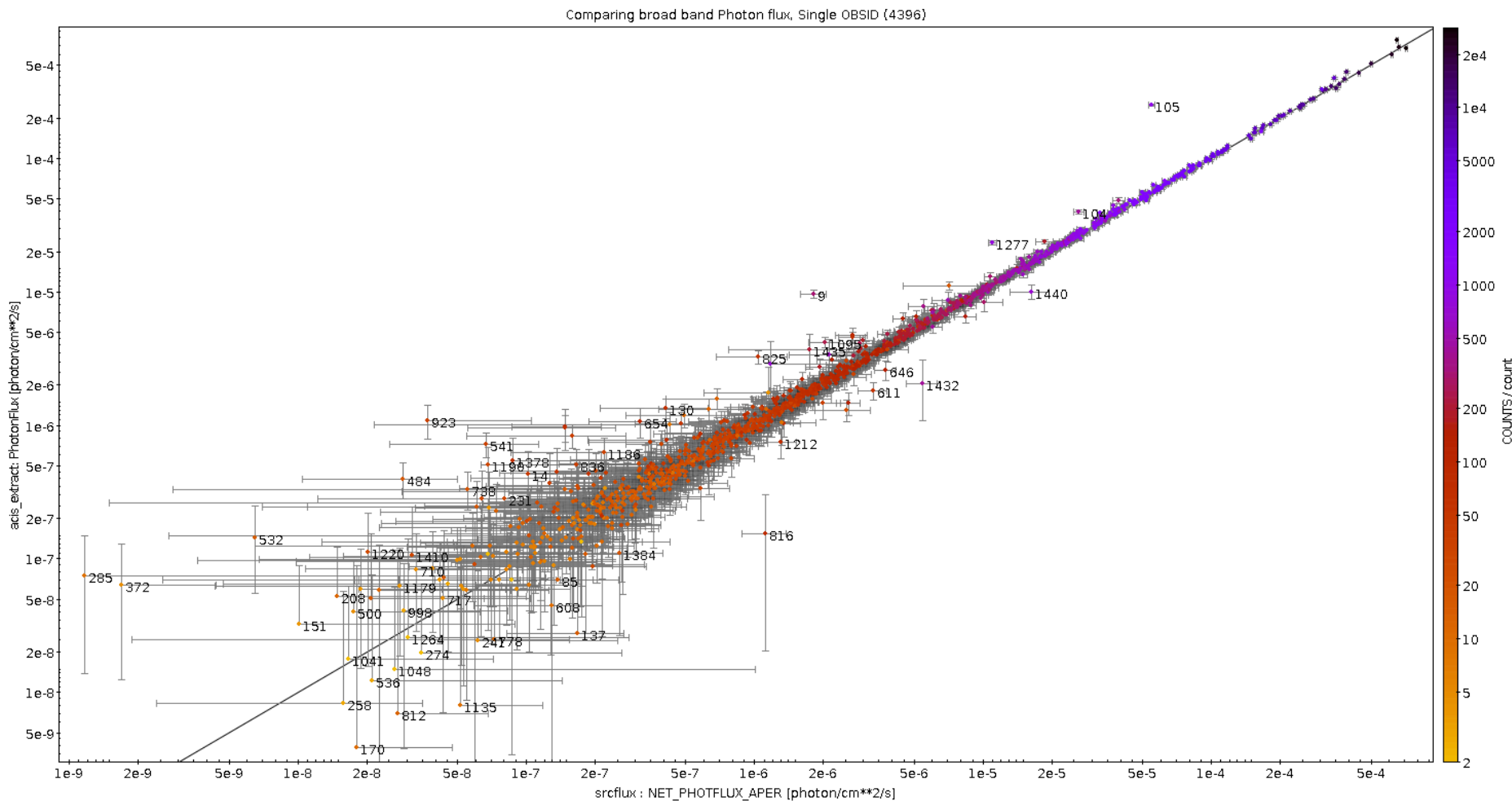
Actions to respond to CUC

- ✓ Review `acis_extract` basic usage and map to `srcflux`.
- ✓ Identify gaps.
 - automatic spectral fits
 - refine source position
 - "optimal" extraction regions
- ✓ Compare `acis_extract` outputs against `srcflux` to confirm consistent results.
- ✓ Provide guide for `acis_extract` users
 - cxc.cfa.harvard.edu/ciao/guides/srcflux_for_ae_users.html

- Fill in the gaps

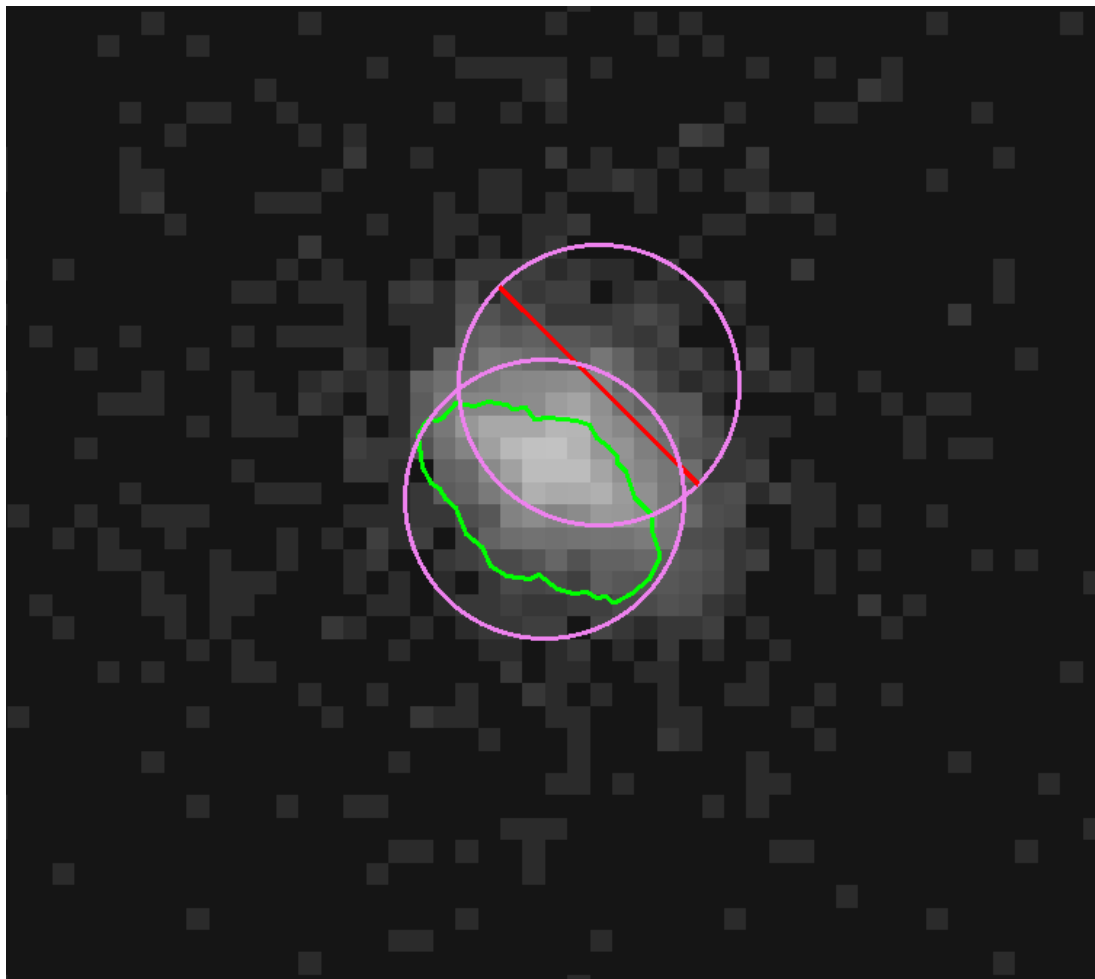


Comparison of photfluxes for single OBS_ID: 4396





Outlier: 105

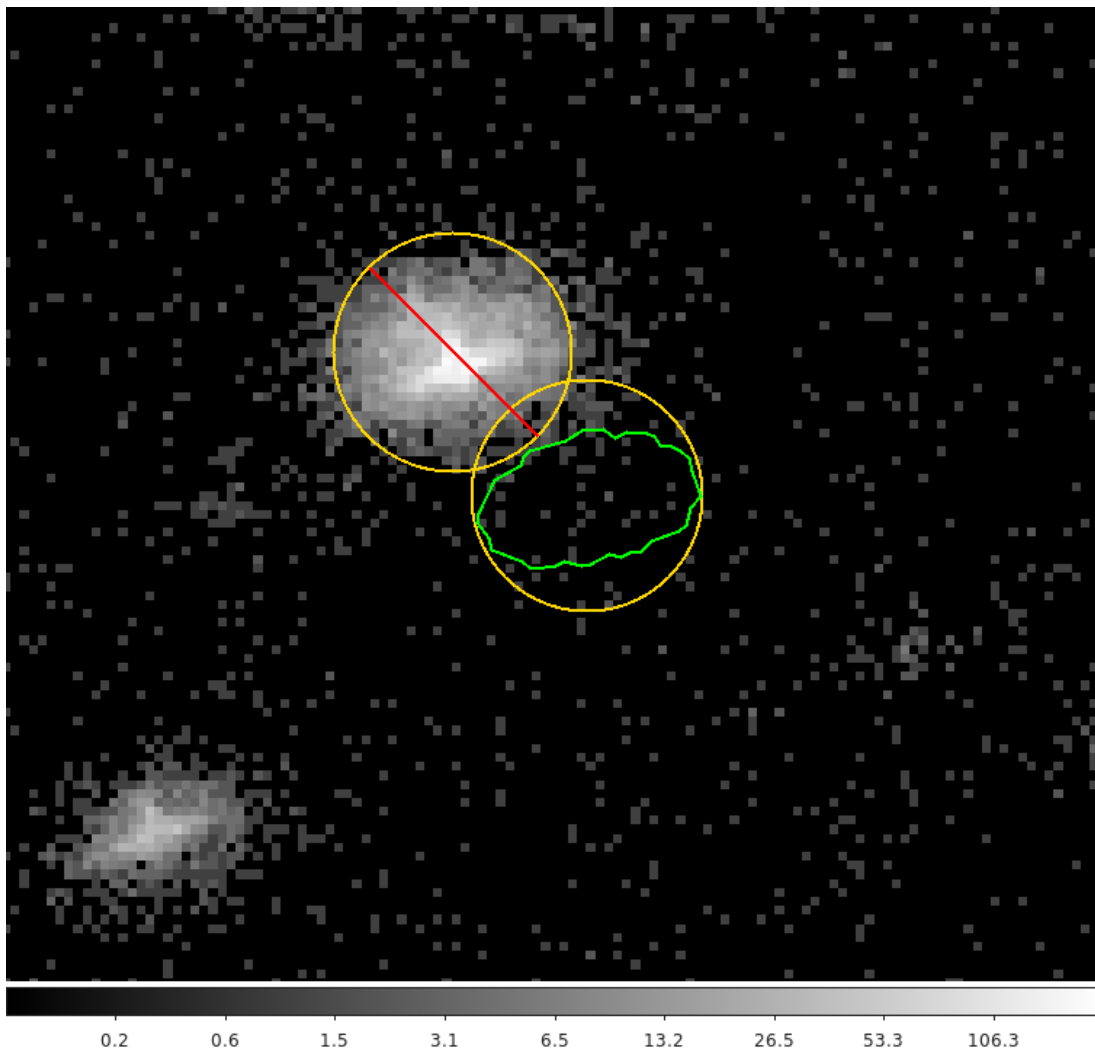


We used 2CXO source positions which reports two nearby sources. The srcflux regions suffer from large overlap affecting the flux estimate.

1 2 5 10 21 42 85 172 342



Outlier: 816

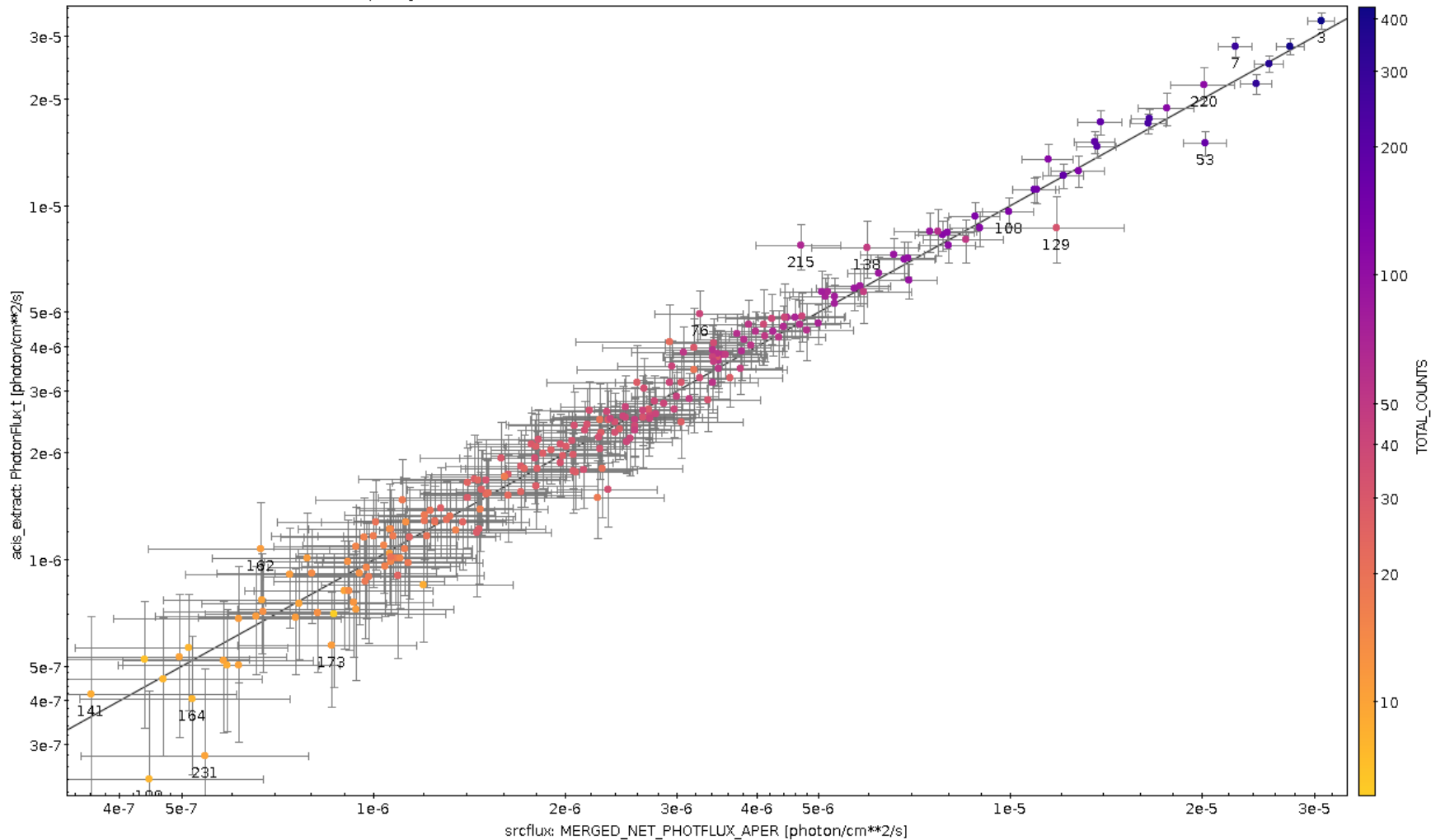


A very faint source (non-detection), next to a very bright source. The srcflux region contains a fraction of the overlapping bright source's PSF which biases the results.



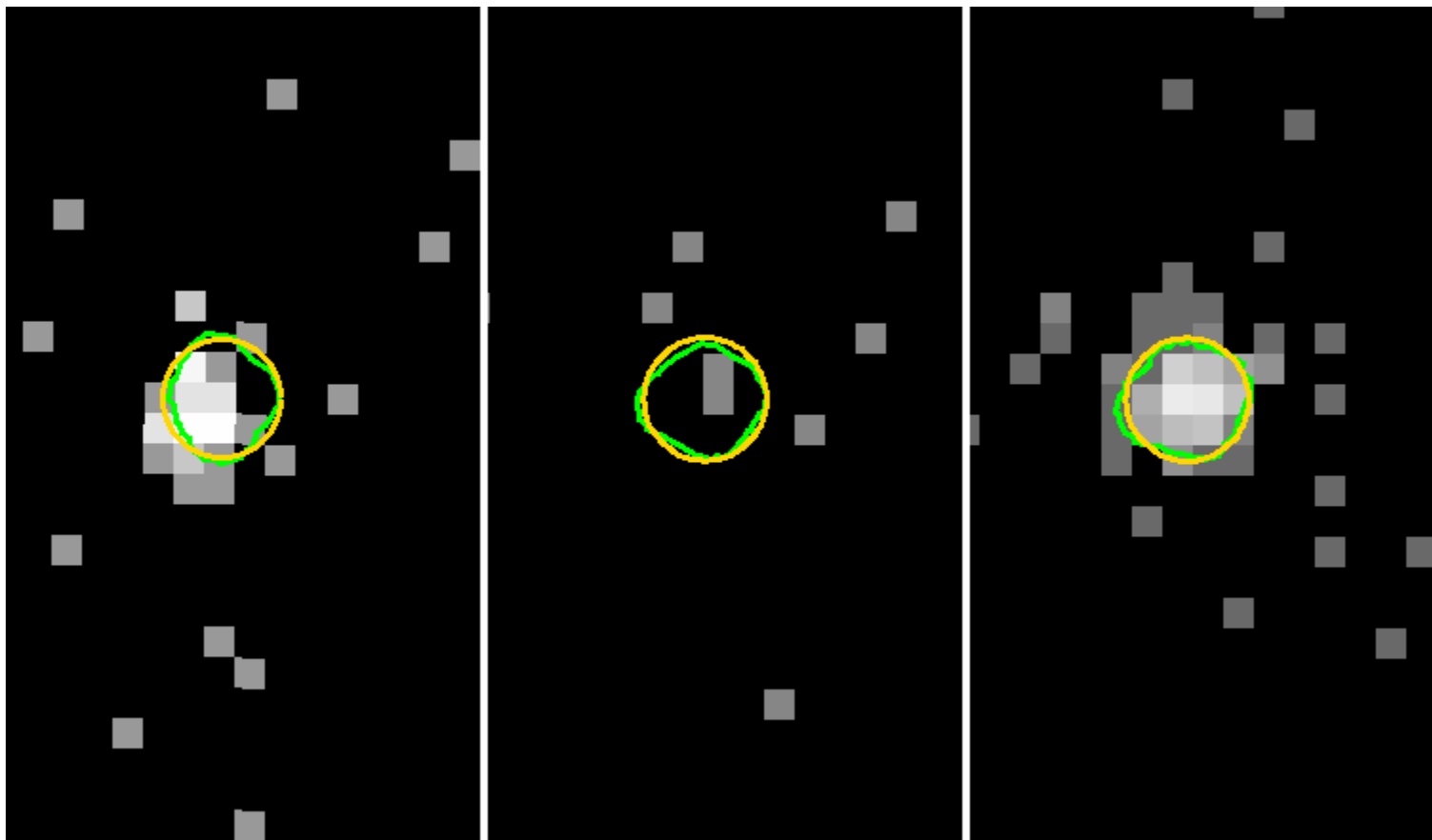
... for combined dataset: Colinder197 (3 OBS_IDs)

Comparing Broad Band Photon Fluxes, Combined Dataset: Colinder 197 (obsids 22689,24757,25099)





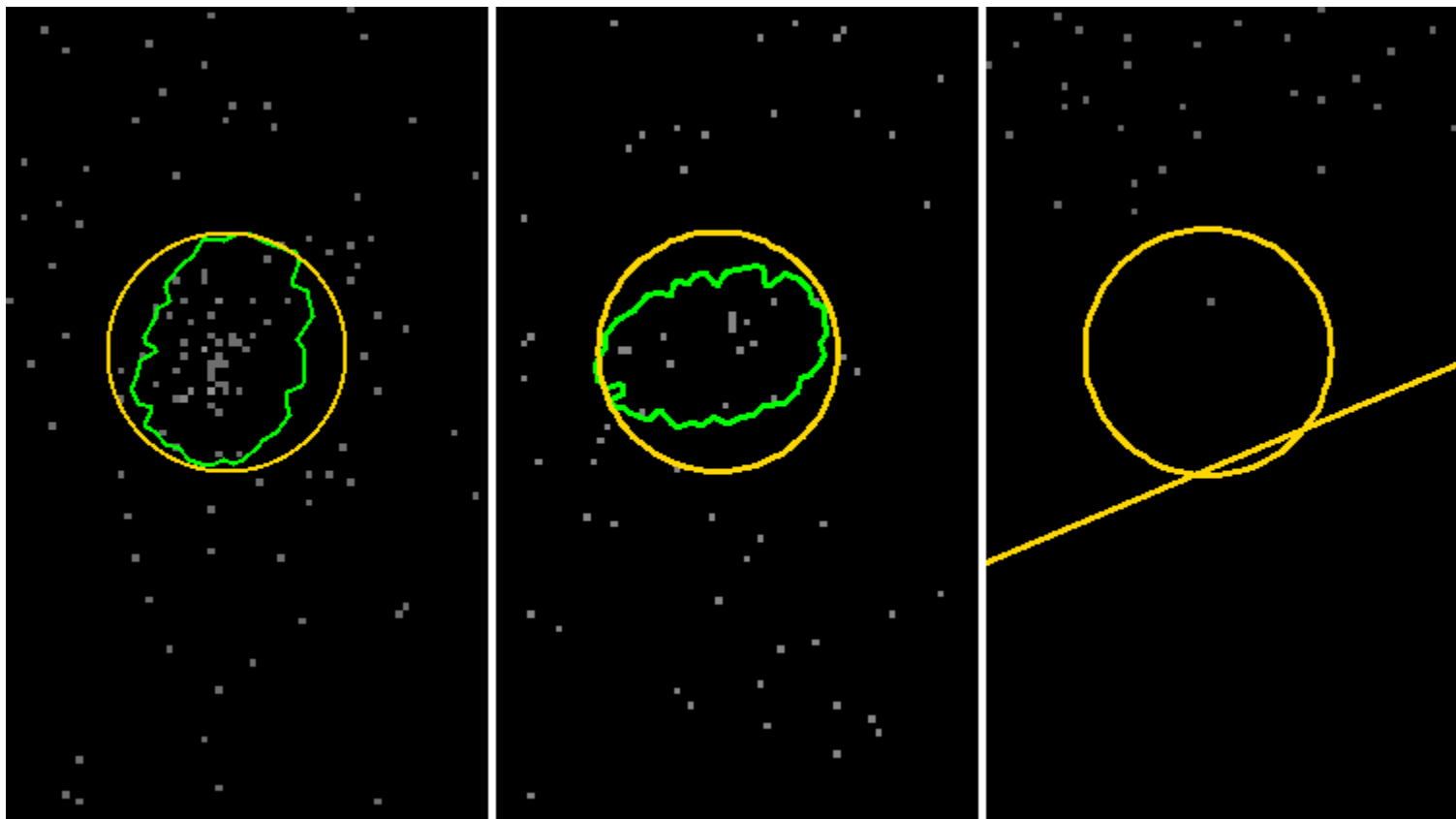
Outlier: 53



A variable source. Since we didn't apply any fine astrometric correction the region in one observation is offset from the source.



Outlier: 215



An edge source. srcflux uses the 3rd observation, acis_extract omits it.



How to fill in the gaps?

- srcflux will be updated to allow users to run their own custom analysis tasks and have the output included with the existing products.
 - Based on a template and several examples users will be able to write a **plugin** to do things like
 - automatically fit spectra
 - recompute source position
 - compute hardness ratios
 - perform deconvolution
 - KS test
- We are working to provide users with new scripts to create source and background regions akin to those used by `acis_extract`
 - source region based on simulated PSF
 - background region enclosing min number of counts.



PSF 90 regions from CSC2

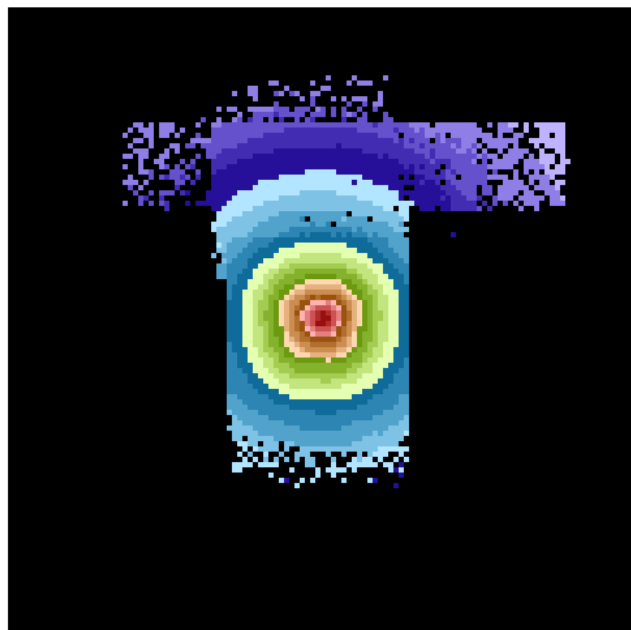
The `acis_extract` polygon regions *can* provide better flux estimates when dealing with overlapping sources compared to the circular regions used by `srcflux`.

As a spin-off on the study of the apertures used by `acis_extract`, we looked at using the PSF parameters from CSC2 to create elliptical regions without needing to actually simulate the PSF.

- CSC2 contains ~1.5M detections where the PSF is simulated.
- 90% ECF ellipse parameters, in multiple energy bands, are already computed:
`mjr_axis_aper90`, `mnr_axis_aper90`, `angle_aper90`
- Randomly distributed across the detectors with more detections closer to the aimpoint.



CSC2 PSF Parameters, ACIS-I



1.0 1.2 1.7 2.5 4.2 7.7 14.5 28.2 55.4

Above: Median `mjr_axis_aper90_m` for ACIS-I observations in pseudo-detector coordinates. The `mnr_axis_aper90_m` map looks similar.

Below: Median `angle_aper90_m` for ACIS-I observations in pseudo-detector coordinates.

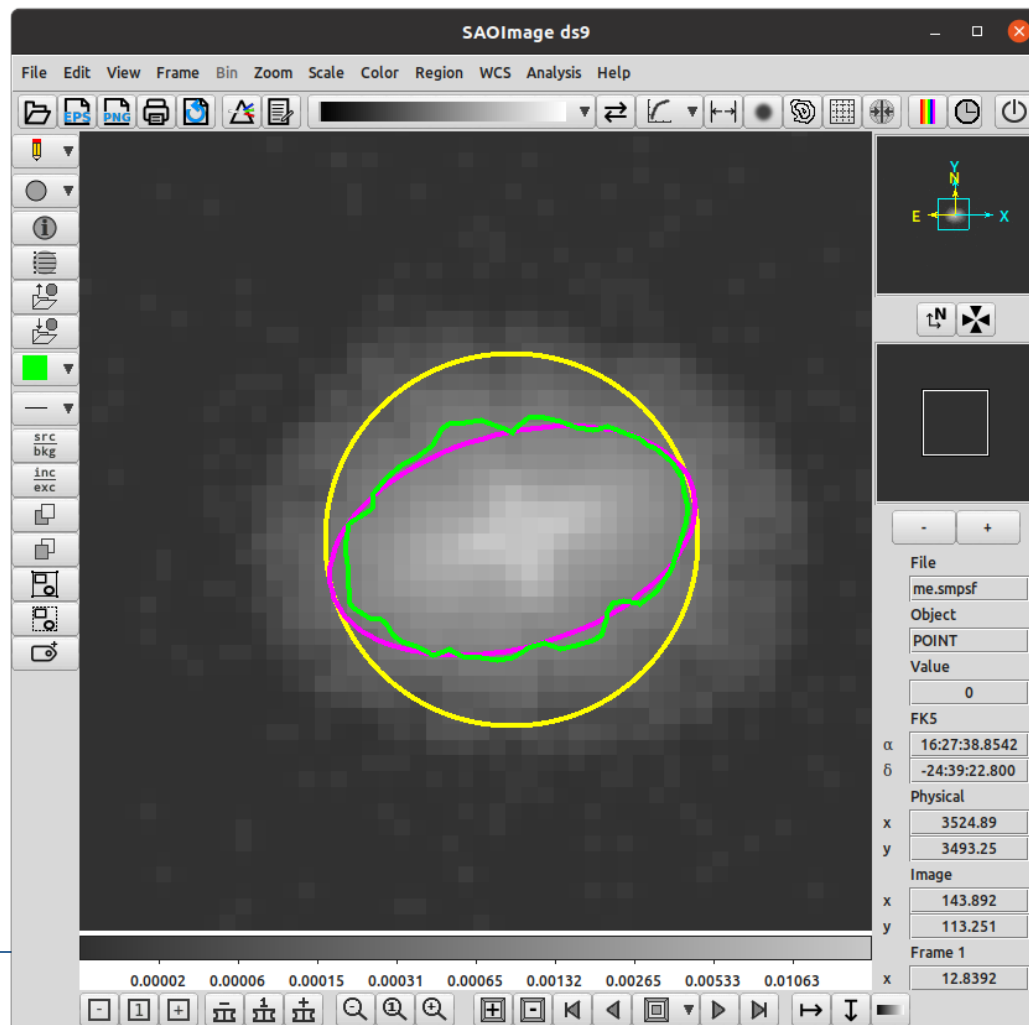


20 38 55 73 91 108 126 144 161



Example

The **yellow** circle is the default regions used by srcflux. The **green** polygon is approximately the same as the region used by acis_extract. The **magenta** ellipse is the region obtained from the CSC2 PSF parameterization maps.





Helpdesk



Helpdesk Stats

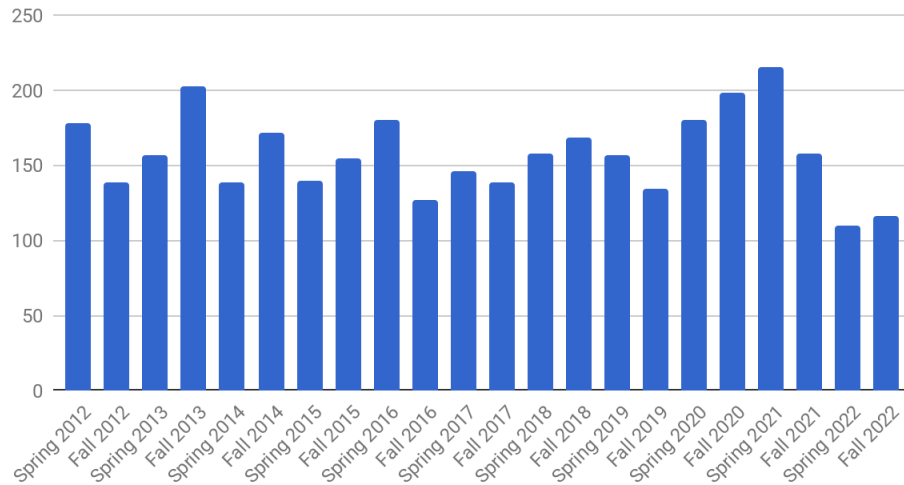
	2021-11-01 to 2022-04-30	2022-05-01 to 2022-10-15
Time period [months]	6	5.5
Number of Tickets	110	116
Median time to 1st contact [hrs]	1.68	0.94
Median time to close [hrs]	17.66	7.93
Maximum time to close [hrs]	1728.2	433.2

Longest ticket was about variability values in the catalog – user's expectations did not match what was in CSC.

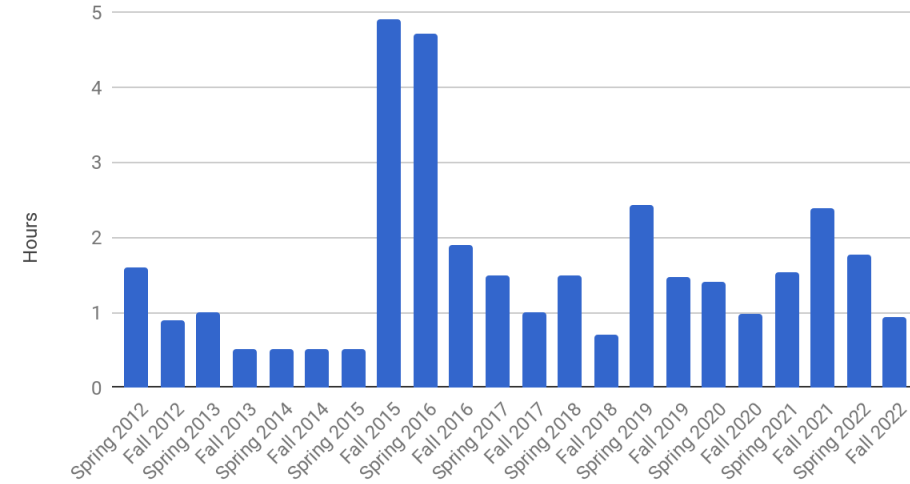


Long Term Helpdesk Trends

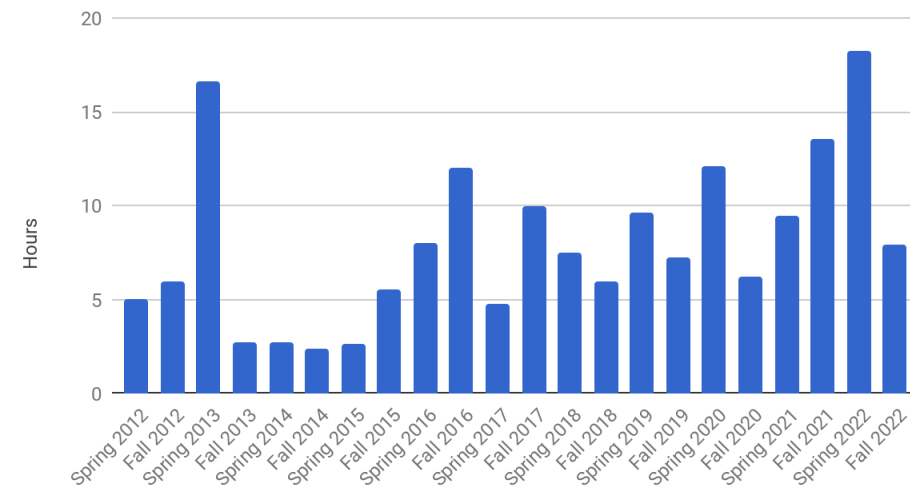
Number of Tickets



Time To Answer

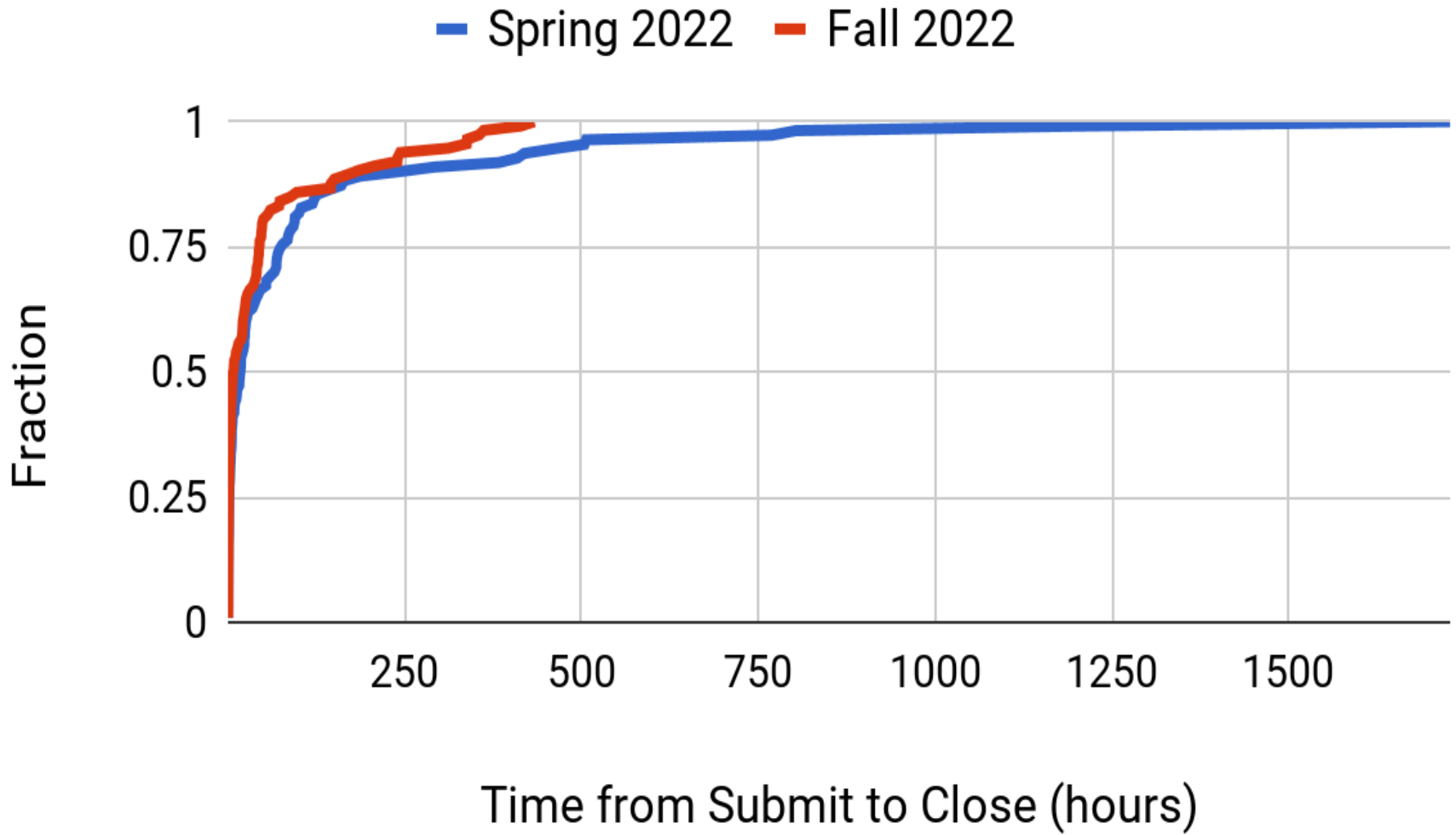


Time To Close



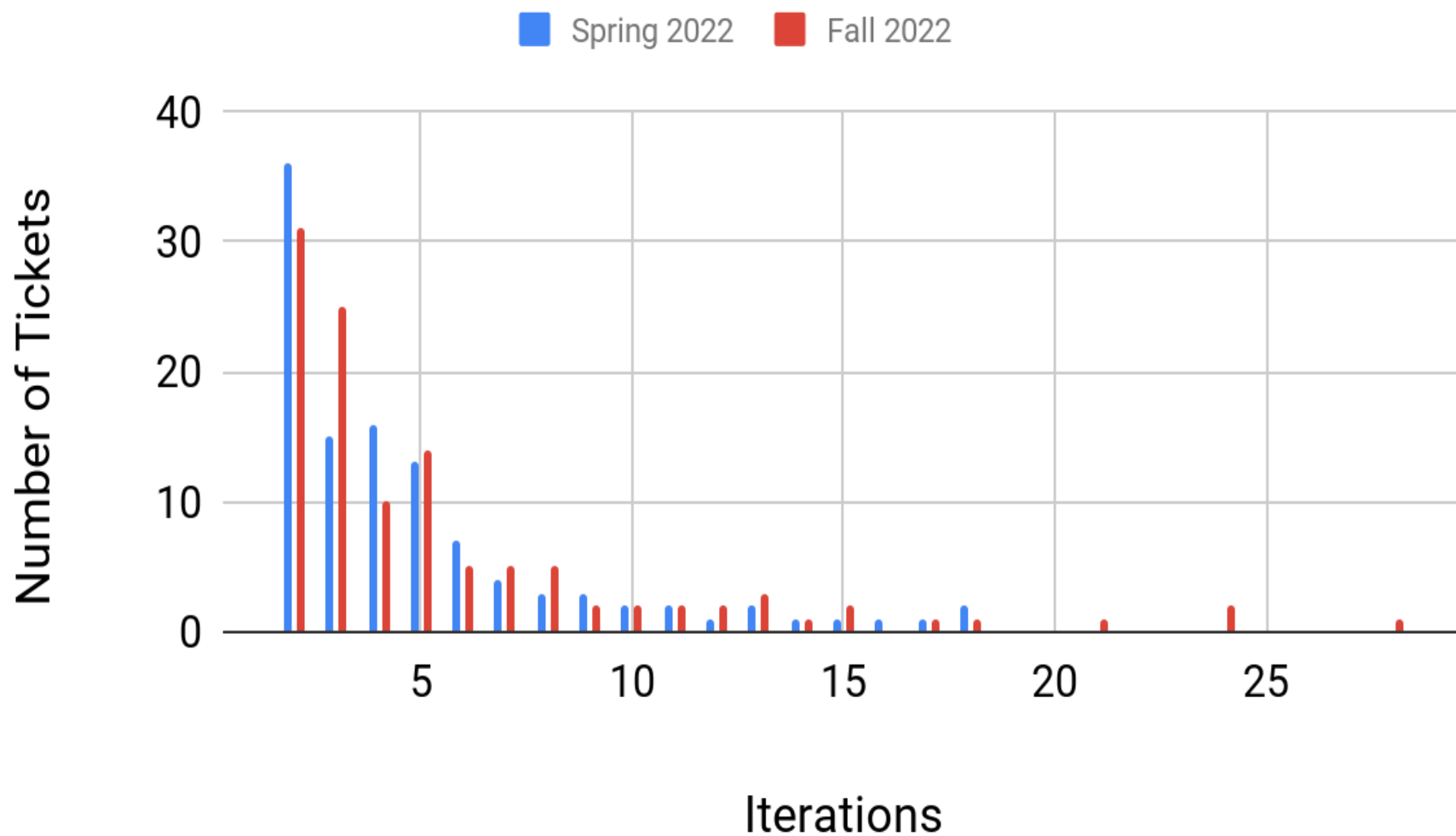


Ticket Aging





Iterations per Ticket





Helpdesk

- Examples of bugs
 - wcs_match when 0 matching rows crashes
 - specextract random failures when running in parallel
 - issue with mismatch pixel size when allowing scale adjustments via wcs_update
 - more filtering problems when using "@" filters
- Example of documentation updates
 - Based on several helpdesk tickets, added Windows Subsystem for Linux (WSL) being unsupported warning message on additional pages.
 - Also based on helpdesk feedback, updated the text on the download buttons to clarify that they are indeed buttons.
Added more links to alternative download instructions for conda users – many users have problem downloading CALDB and sherpa tar files from CXC (timeouts)

Added new Watch Out topic about special characters (especially spaces) in file and directory/folder names. There has been an increased number of users affected by this.



Community



Community Outreach

- AAS 239 - January 2022
 - Scheduled workshop was cancelled due to lack of signups; entire conference then cancelled
 - Arranged two one-on-one zoom sessions with participants to address individual data reduction questions
- American Physical Society (APS)
 - Supported first time presence at Chandra booth (dual role w/ CDO)





Community Outreach

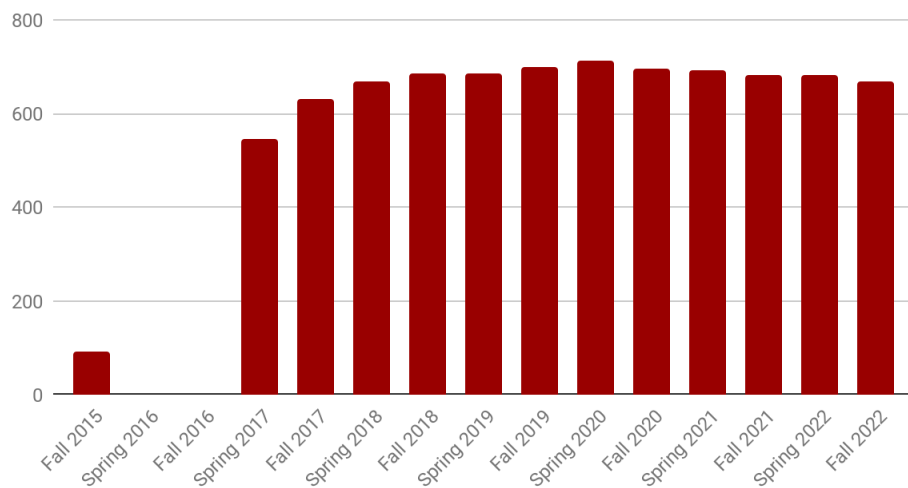
- Support for IAU-OAD project AstroSprint: The Next Generation (3-5 June 2022)
- Scheduled in person CIAO workshop at AAS 241, Seattle.
 - Registration has been light. Go/No-Go TBD.
- Plans for one-on-one or one-on-few support for video calls on request to offer support between a mutli-day workshop commitment and a simple helpdesk request.



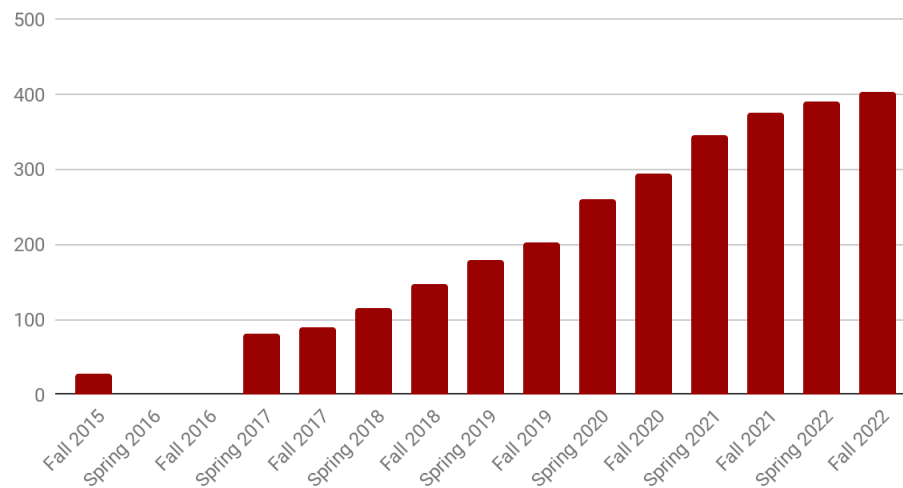


Social Media

Facebook



Twitter



Facebook	ChandraCIAO	670 followers (-11)
Twitter	@chandraCIAO	402 followers (+13)



CIAO 4.15 and Scripts Overview



CIAO 4.14 Highlights

- The ardlib library has added support for the Frame Store Shadow bits. This fixes the problem where the Effective Area is erroneously higher at the bottom edges of the CCDs (so the work-arounds added in the contributed scripts are no-longer needed).
 - The azimuth angle, phi pair used to determine the vignetting factor to the effective area has been corrected.
 - A bug with dmextract and multi-chip observation times has been addressed (only significant if the per-chip GTIs are significantly different).
 - The dmdiff comparison tool has been overhauled, improving support for array data and WCS metadata.
 - The dmradar tool has been added (adaptive binning in polar coordinates) to complement dmnautilus.
 - Update Sherpa to support XSPEC 12.12.0 model library (new models and abundance tables).
 - The conda installation method now uses the conda-forge channel.
-



CIAO 4.15 Highlights

- `tg_resolve_events`: fix for HRC-I+HETG overlap area
- `mkgrmf`: support off-axis PSF for extraction of off-axis grating sources in all grating/detector combinations
- `dmimgcalc`: operation can now be a stack allowing much longer expressions.
- `wavdetect`: double precision FFTs fixes instability of detecting low significance sources, fix to remove duplicate sources.
- `dmmakepar`: write units so that they are recognized as such in FITS header
- `wavdetect`, `axbary`, `reproject_aspect` and other scripts now record proper HISTORY records.
- several tools updated to propagate WCS from input image to output table.
- many tools updated to support NULL values (as opposed to NaN values)
- `sherpa` updated to XSPEC 12.12.1c, focus on fixing data access issues (e.g. setting `stat/sys` error values to match data), UX improvements for filtering data



Contributed Scripts

Highlights

1. 4.14.0 (December 2021)

- a. dax upgrades: take advantage of new button bar support in DS9 and better integration with DS9's table display functionality ("prism")
- b. specextract now takes advantage of multiple cores and has improved diagnostic messages
- c. Added FOV output to the image and merging scripts (fluximage, merge_obs)

2. 4.14.1 (January 2022)

- a. specextract regressions have been addressed (blank-sky and solar-system observations)

3. 4.14.2 (April 2022)

- a. chandra_repro now reminds users to look at the V&V report and has seen improvements for grating and solar-system observations
 - b. srcflux now reports on intra-observation variability using glvary (similar to the Chandra Source Catalog approach)
 - c. specextract has improved support for analysing data from the ACIS readout streak
-



Contributed Scripts

4.14.3 (Sep 2022)

updated search_csc scripts to be able to access the "current", CSC 2.1 in-production catalog.

In progress, planned for 4.15 release

apllimits tool to calculate upper limits following Bayesian method of Kashyap et al. (2010)



Sherpa



Sherpa Development 2022



- Sherpa releases:
 - CIAO version 4.14 was released on Dec.14, 2021; It contained the code introduced in the Sherpa standalone during 2021 development year and released in 4.14.0
 - 4.14.1 May 20, 2022
 - 4.15.0 October 11, 2022
 - CIAO release in December 2022 will include the changes from both standalone releases.
- 4.14.1 Release highlights:
 - Enhancements:
 - Various plotting backend improvements to support future planned updates
 - Data I/O backend improvements and data object class improvements
 - Beta support for py 3.10 and XSPEC 12.12.1
 - Support for Mac Clang 12.0 compiler
 - 15 bug fixes and 52 other changes



Sherpa Development 2022

- 4.15.0 Release highlights:
 - Enhancements:
 - Improved validation of arguments when creating Data objects
 - Filter setting with notice/ignore are reported to the screen for the users of the UI layer.
 - Increased test coverage for plotting
 - Infrastructure changes:
 - Drop support for Python 3.7
 - Updates to start creating Python 3.10 Conda packages.
 - Various improvements to the GitHub Actions and GitLab workflows
 - Bug fixes:
 - Ensure *chi2specvar* errors match XSPEC when 0 counts are present during background subtraction
 - Remove model instances from the global symbol table when clean is called
 - Addresses new warnings in the tests for Matplotlib 3.6.0 and AstroPy 5.1
 - Minor issues in *fake pha*
- 35 Sherpa Pull Requests (PR) including 6 bug fixes.

Full Release Notes:

- <https://github.com/sherpa/sherpa/releases/tag/4.15.0>
- <https://zenodo.org/record/7186379>



Sherpa Ongoing Development



- Upgrade to plotting backends to allow dynamic plotting (2023)
- Datastacks support for grating analysis (2023)
- Models - ongoing efforts to improve python build of xspec model library
- Infrastructure changes to support build with new compilers, python version and updated platforms
- Documentation and bug fixes

Details of the development, including issues and pull requests for the new code are available on GitHub:

<https://github.com/sherpa/sherpa>



Using Sherpa in Astronomy Research



1473 publications in ApJ, AJ, MNRAS, A&A and others use Sherpa (since 2001 and including astro-ph abstracts)

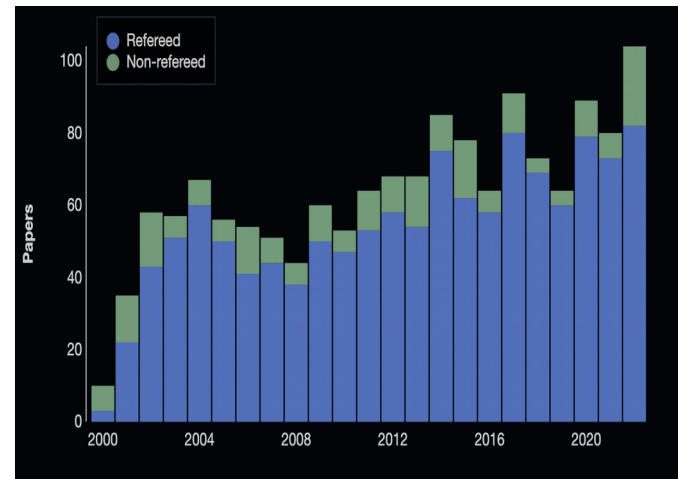
<https://ui.adsabs.harvard.edu/public-libraries/X6orMXwpRtSPy8x1uiiRMg>

389 citations to Freeman et al 2001 SPIE paper

20 citations to zenodo releases: DOI: [10.5281/zenodo.593753](https://doi.org/10.5281/zenodo.593753)

104 research papers published in 2022

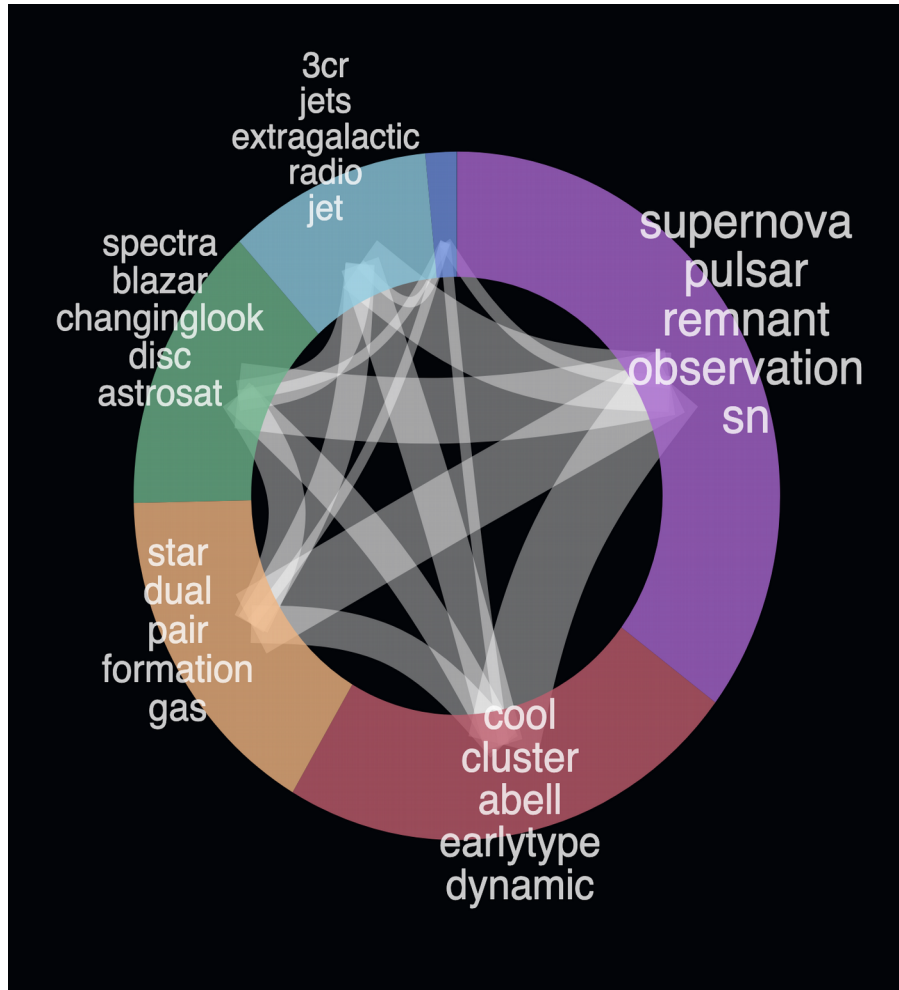
7 PhD theses listed in ADS that used Sherpa



Statistics from ADS stats



Using Sherpa in Astronomy Research



Paper Network in ADS

for 500 papers published in 2016-2022



Instruments/Grating S





- Added support for TG with HRC-I (HETG/HRC-I, LETG-HRC-I): updates to CIAO tools and CALDB.
 - Bugs found in grating region handling for higher orders, small effects to HRC-S/LETG effective areas due to errors in applying the CALDB enclosed energy fraction at a given diffraction angle; details reported to DS and ticket opened.
 - Update to order-sorting tables: new tool specified for CALDB file ("OSIP") construction, to keep up-to-date with changes in the ACIS response calibration, especially given higher temperature operations. Algorithms specified, delivered to DS for implementation.
 - V&V: Provided support for V&V of all pipeline products for grating observations. Handled some special cases requiring custom processing for un-filtered transient bad pixels or occasional bad zeroth order positions. Performed V&V for all public data ingested into TGCat.
 - TGCat: normal processing, and a few special processing requests.
 - Continued support for ISIS package via isis-users list
-



Responding to changing temperatures



-
Included warning message for users running specextract on 'warm' observations where the focal plane temperature is warmer than -109 C.

In coordination with the ACIS calibration team, SDS is planning contributed scripts for reducing and creating responses for observations with large ACIS focal plane temperature variations.



PSF

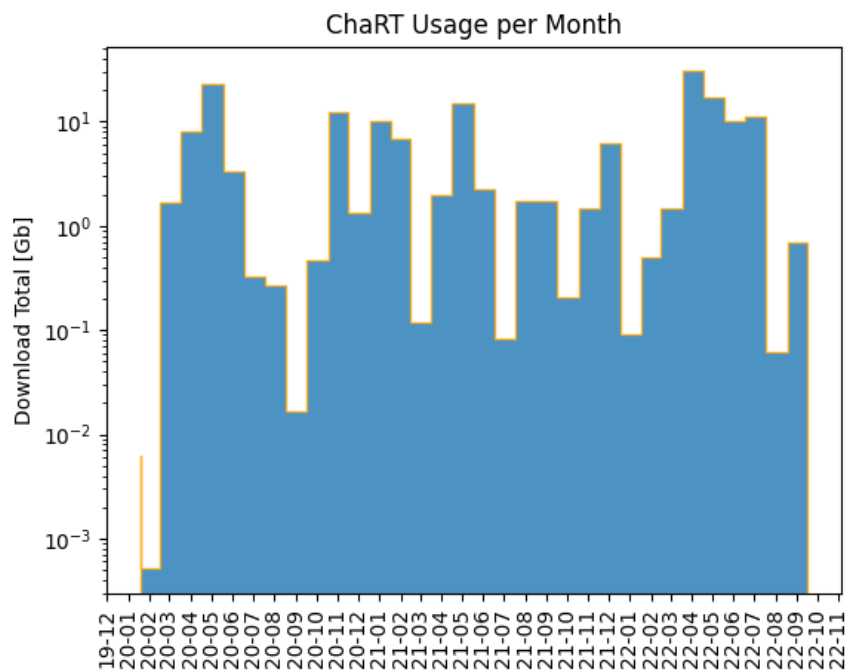
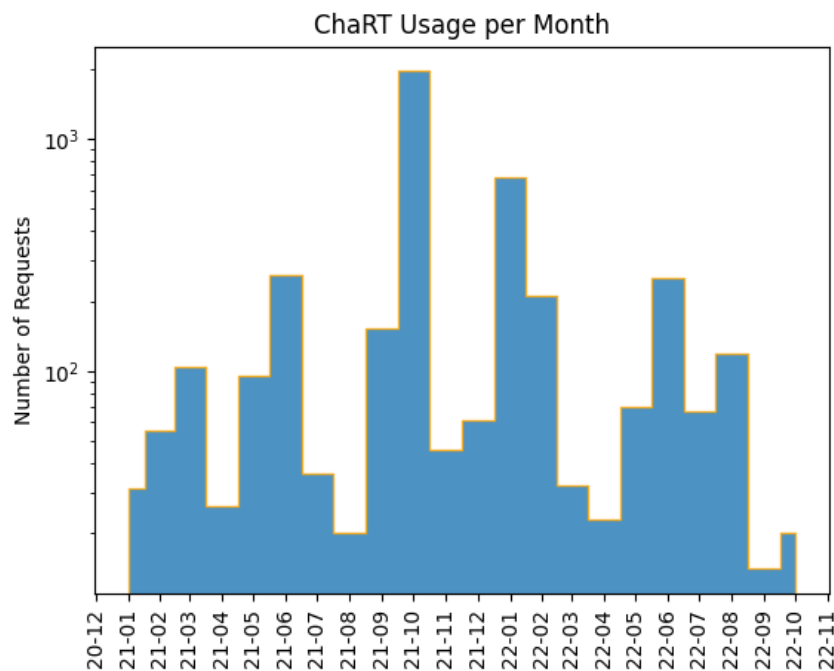


ChaRT and Marx

- ChaRT – monitored usage, routine OS and system upgrades
- Marx – no activity



ChaRT



Over the summer we had the first instance where a user supplied an input spectrum which generated over 50Gb of data; exceeding current capacity. Contacted via helpdesk and resolved issue (user failed to read threads/instructions to provide proper units).



Visualization





SAOImageDS9

- Releases

- Version 8.3 - released in Dec 2021 with CIAO 4.14
- Version 8.4b1 - released in Apr 2022
- Version 8.4b2 - released in Aug 2022
- Version 8.4rc - released in Oct 2022
- Version 8.4 - planned release Dec 2022

- New Features 8.4

- Multiple Layout options: horizontal, vertical, basic, advanced
- New Illustration mode, users may add graphics for presentations.
- Fade between images, save as GIF movie.
- Improvements in access to regions analysis results via XPA, SAMP
- Plots can display Sum, Average, and now Median of data values.
- MacOS ARM64 binary support, for both X11 and Aqua ports, Monterey and Ventura.

- Other improvements and enhancements.



dax: Interactive Grating Coordinates Vector

The screenshot shows the SAOImage ds9 software interface. The main window displays a Chandra X-ray image with a yellow vector line. The 'Coordinates' panel shows the tool's settings, and the 'tgcoords' panel shows the resulting grating coordinates and energy.

Parameter	Value
Grating Arm	HEG
Order	1
TG_R [deg]	0.054735
TG_D [deg]	-0.00059019
Energy [keV]	6.487

Adds a special vector (line) to the display. One end is 0th order location (taken from [REGION] block). Move the arrow end around to get updated grating coordinates and energy.

Uses new pyxlib routines made available in CIAO 4.14.

Uses ds9 region "callback" functionality.

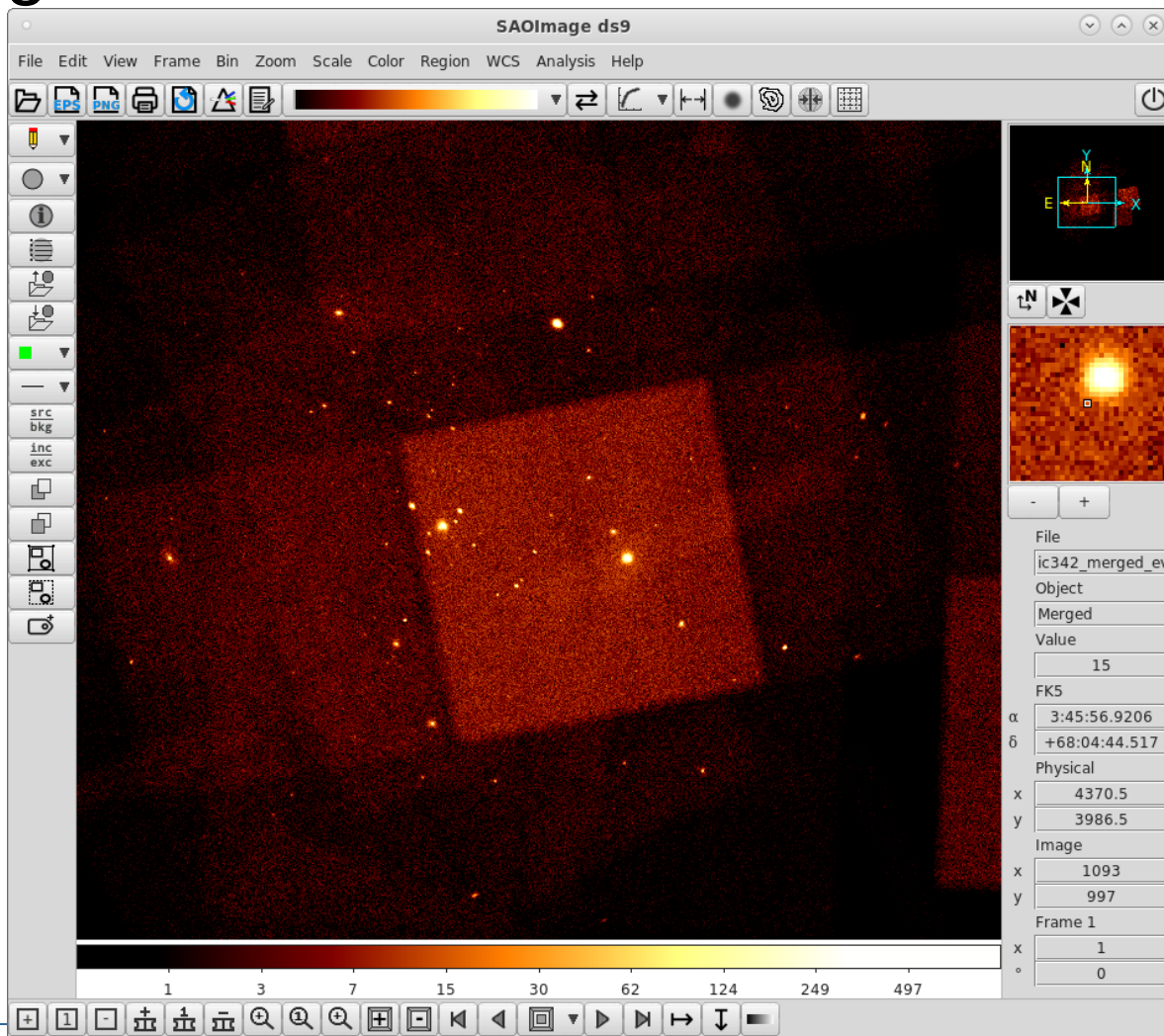


SAOImageDS9 Advanced View

I/O

region controls

panner and magnifier controls



File
ic342_merged_ev
Object
Merged
Value
15
FK5
 α 3:45:56.9206
 δ +68:04:44.517
Physical
x 4370.5
y 3986.5
Image
x 1093
y 997
Frame 1
x 1
y 0

all buttons have tooltips

bin, block, zoom, frame controls



SAOImageDS9 - Nov 2021 to Oct 2022

- **GitHub Activity**
 - 341 Commits
 - 58 Release Note Entries
- **Help Desk**
 - 110 CXC HelpDesk Requests
- **Downloads**
 - 48015 unique IP addresses