



Science Data Systems

Jonathan McDowell



CIAO Update – Jonathan McDowell



- Helpdesk, documentation updates and survey results
- Regression tests revamped
- CIAO 4.1.2 patch released in April
- CIAO 4.2 release in test for Dec release
- PSF study underway
- Catalog characterization on continuing basis
- Catalog HRC R&D work mostly done, preparing for production (I Evans presentation)



Community support

- Responded to CUC
 - carried out major user survey
 - Added new ChIPS and Sherpa examples
 - Investigating end-of-life plan for CIAO
- No staff changes since last CUC
- Helpdesk: 189 new tickets (Mar 18 – Oct 19), 11 still open
 - Median time to first answer 1 hour (longest 2 weeks, science staff unavailable)
 - Median time to final answer 10 hours, but some stay open for several months (need feature in next release, or continuing support of difficult analysis)
 - Answers generated 24 new docs, 3 new bugs, 14 RFEs
 - 60% did not require scientist or DS support
- Gave catalog GUI (CSCview) and CIAO demonstrations at Pasadena (summer) AAS meeting – updated CIAO demo
- Sherpa demo, X-ray astronomy lecture, fitting statistics lecture at Paris Multiwavelength Astro School (July)
- Supported demos and solicited feedback from community at 10 Years conference in September.



Documentation



- New threads since last quarterly:
- Simulating ACIS-S LETG spectra with Sherpa
- Search for variability in a source
- Measuring source extent
- Catalog with DS9/TOPCAT thread
- Catalog coverage thread
- Running scripts in Sherpa
- Estimated Background Image
- Extract a Spectrum from the ACIS Readout Streak
- Extract a Spectrum of a Solar System Object
- Compute Upper Limit for Unresolved Source

- Updated existing threads to accommodate new functionality in scripts package
- Extensive updates to catalog documentation, e.g. docs for Google Earth interface
- Planning to update CIAO web to match new CXC front page design



Sherpa

[Sherpa 3.4 Homepage](#)
[Sherpa 4.1 Homepage](#)

[About Sherpa](#)
[Choosing Python or S-Lang](#)
[Converting to the New Syntax](#)

Analysis Threads

[Sherpa](#)
[Science \(CIAO\)](#)
[ChIPS](#)
[ChaRT](#)

Help Pages (AHELP)

[Alphabetical: Python](#) | [S-Lang](#)
[By context](#)
[Using ahelp](#)

Gallery of Examples

[Python](#)
[S-Lang](#)

Models, Statistics, and Methods

[Models](#)
[Statistics](#)
[Optimization Methods](#)

Documentation

[Latest Updates](#)
[Known Issues and Limitations](#)
[FAQ](#)

[Sherpa 3.4 to 4.1 Conversion](#)
[References](#)
[Publications](#)

Download Software

[Contributed Sherpa Packages](#)
[Contributed Sherpa Scripts](#)

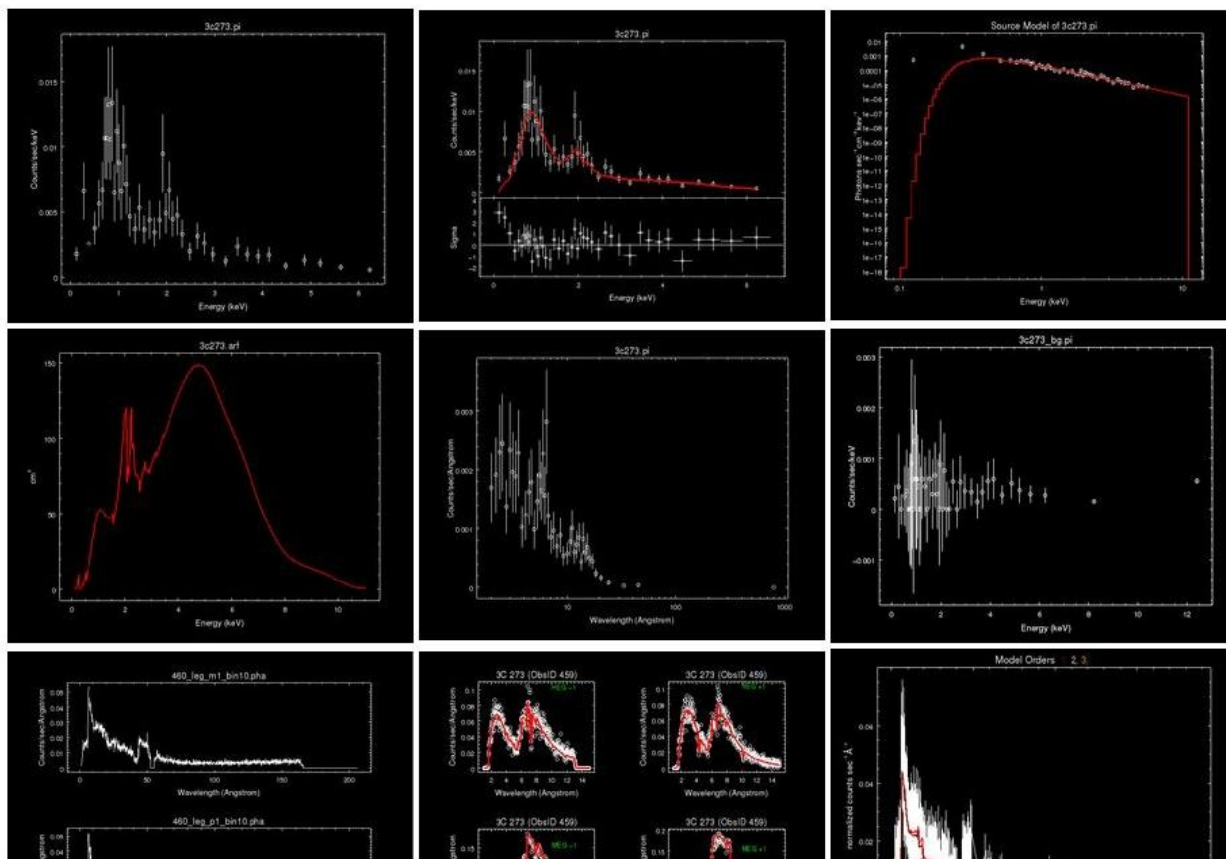
CXC Links

[CIAO \(Data Analysis\)](#)

Thumbnails of Sherpa Examples (Python)

The examples below are intended to illustrate the various plotting and fitting capabilities of Sherpa, as well as help you transition from the Sherpa 3.4 to 4.1 syntax. Click on a thumbnail plot to view and compare the Sherpa 3.4 and Sherpa 4.1 (Python) scripts used to create it.

See also the [List View](#), in which the examples below are split into Sherpa 1-D and 2-D "Plotting" and "Fitting" categories, with descriptive titles.





Scripts

- New deflare script combines old `lc_clean` and `lc_sigma_clip` with new ChIPS
- `check_ciao_caldb`: we were getting more helpdesk tickets arising from broken CALDB installations and provided this diagnostic script.
- `fullgarf` modified to use CALDB bad pixel file by default
- assorted python/S-Lang modules to simplify user threads – e.g. `save_instmap_weights` routine used in 'Calculating Spectral Weights for `mkinstmap`' thread.
- ensuring compatibility of scripts package with `ciao_install`
- Planning script for easier spectral weighting in `mkinstmap`, `ChART`, `mkwarf`
- Planning rewrite of `merge_all`



CIAO packaging improvements

- CIAO download process was getting too complicated – even more so with split into separate packages
- new 'ciao_install' script helps you get CIAO - worked design with DS
- Testing and feedback on prototype
- Preparing to integrate into CIAO web site
- Updating documentation
- Our view is that ciao_install is a huge improvement and makes it much easier to download CIAO



CIAO 4.1.2

CIAO 4.1.2 released in April, has been downloaded extensively

Downloads of CIAO 4.x to 466 IP addresses
(some addresses download more than one platform:)

Linux	300
OS X Intel:	224
Mac PPC:	26
Solaris:	7
Source build	51 (mostly for Linux platforms)

13 downloads of CIAO 3.4

Conclusion: Linux and Mac-Intel remain our users' dominant platforms
Mac PPC and Solaris continue to decline – very few users
User survey provides independent dataset which is in good agreement with these data.



SDS internal regression testing is done on multiple platforms. Major reorganization and update of regression tests and test data completed over the summer. New baselines run and analysed on all platforms.

We update our test operating systems to keep up with the user community
Platforms for 2009:

- Fedora Core 8 (may upgrade to more recent FC)
- Mac Intel OS 10.5 (plan to do some testing on 10.6 just released)
- Mac PPC 10.4 (legacy OS, candidate for end of support)
- Solaris 10 (used in data system and by CXC ops team; few external users)

SAO-HEAD/CXC moving internally to CentOS machines (more stable OS)
DS is shifting to CentOS for the Linux build of CIAO
Not seeing CentOS as a community trend yet – SDS will continue regression testing on recent Fedora OS releases, using the CentOS CIAO binaries.

Now seeing many more users with Ubuntu Linux than before
Did spot and download testing on Debian and Ubuntu to check no drastic problems
Also did download testing on Fedora Core 11 released this summer
Will download-test Fedora 12 when it is released in November



CIAO User Survey – 1: General



Last survey was in 2003

We wanted to understand how CIAO is being used, and CUC encouraged us to carry out a new survey. What do users like? What do they need? What platforms are they using?

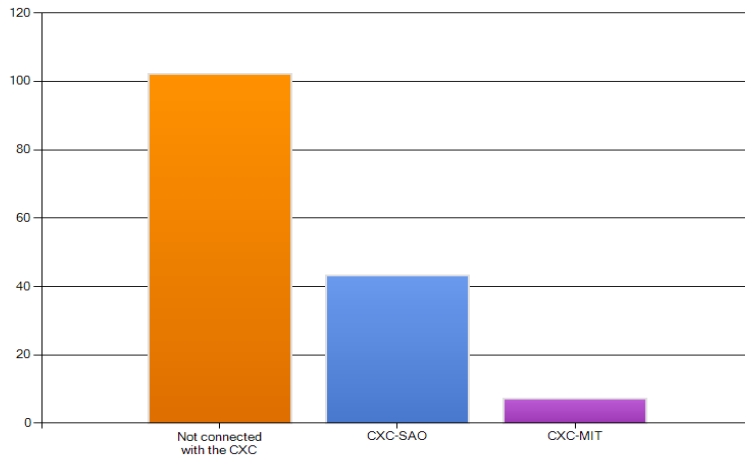
We had 151 respondents (101 from outside CXC). No major differences between CXC and non-CXC responses. In general the responses are consistent with the SDS team's qualitative impressions from HD requests and from interactions with the community at AAS and elsewhere.

Responses on internal web page:

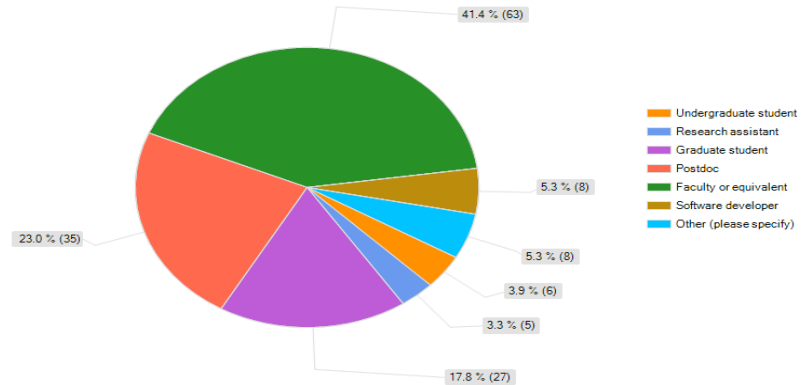
http://icxc.harvard.edu/sds/survey_2009/index.html

Will make summary of responses available to community.

What is your connection with the Chandra X-Ray Center (CXC), (if any)?



Your experience:





CIAO User Survey – 2: Tools and infrastructure



Some results:

Tools -

ubiquitously used, and comments are positive (DM tools are widely seen as a strong feature). 31 percent of users use CIAO at least sometimes to look at data from other (non X-ray) wavelengths, but they are nevertheless self-identified (131 out of 151) as mainly X-ray astronomers.

Scripting -

Most (60%) never use scripting language capabilities
The rest are split 50/50 between S-lang and python

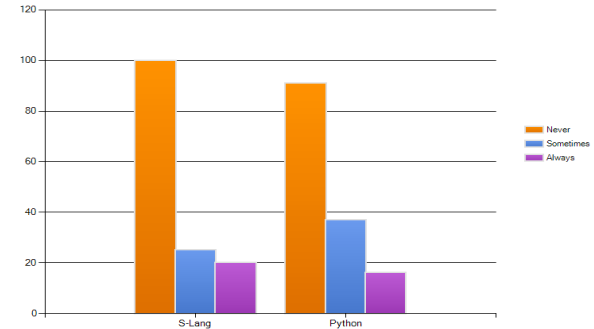
70% of users plan to learn python vs 47% planning to learn S-lang

We asked about importance of python, S-lang and old-style CLI for Sherpa/ChIPS interfaces. 30% said the python interface was 'very important', 40% said a CLI was very important, 10% S-lang.

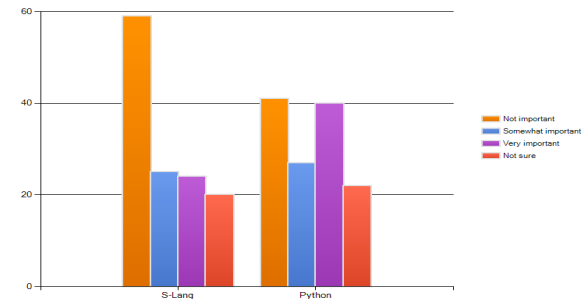
Documentation -

Generally satisfactory, with all methods (ahelp, threads, helpdesk) being widely used.

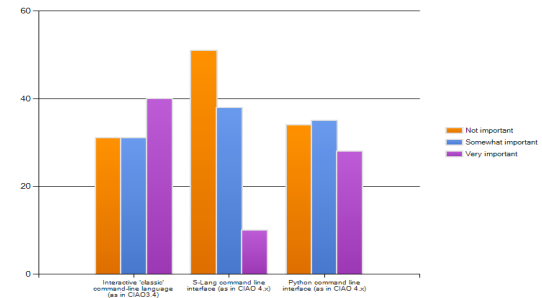
CIAO currently supports two scripting languages, S-Lang and Python. Do you write scripts in any of these languages? (even if you do not currently use them in CIAO)



In general, how important is it to your work that CIAO support these languages?



How important are the following interfaces in Sherpa and ChIPS?





CIAO User Survey – 3 : Sherpa/ChIPS

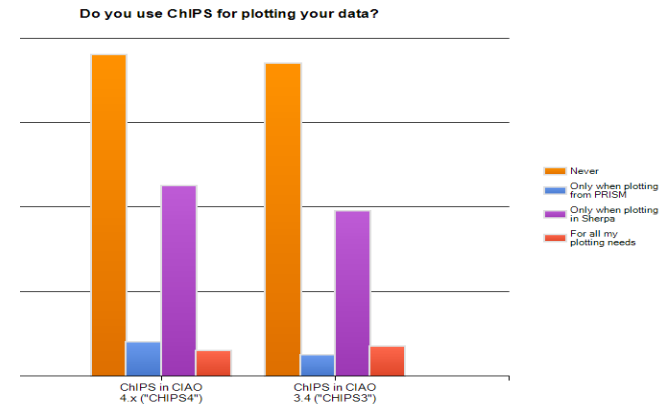
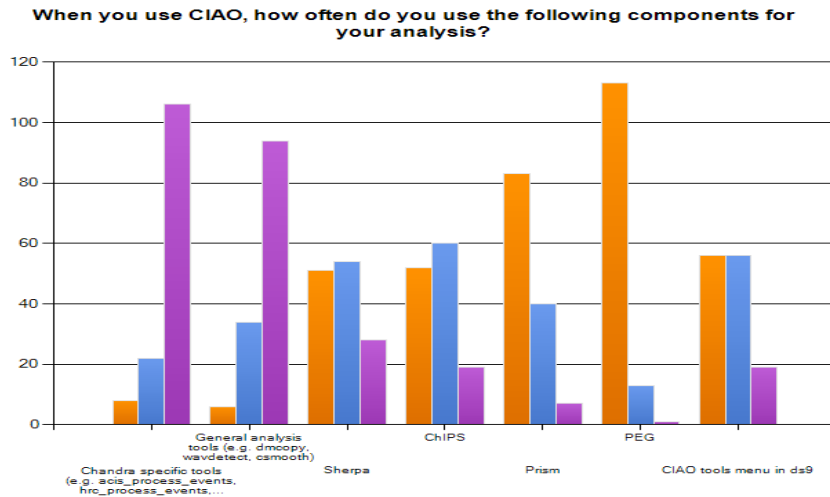


Sherpa and ChIPS:

There is a significant Sherpa/ChIPS user community. 60 percent of CIAO users sometimes use Sherpa/chips (q9: 82/133 respondents use Sherpa sometimes or always, 79/131 for ChIPS), 35 percent sometimes use Prism (use of the peg parameter editor is rare, 10%).

For comparison, 20% report that they sometimes use ISIS, 85% sometimes use XSpec (split evenly between XSpec 11 and 12).

You can use Sherpa and ChIPS by entering 'python' and importing them as modules, but users don't make use of this: they mostly just type 'sherpa'. Most ChIPS use is via Sherpa – users currently use IDL, pgplot and sm for general plotting tasks; only 7 percent of users use ChIPS outside of Sherpa.

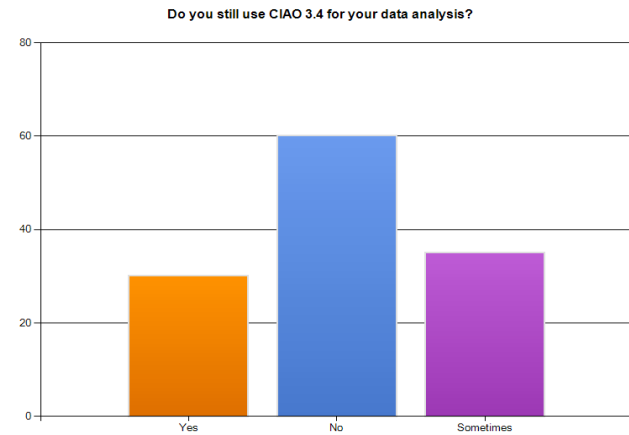




Sherpa and ChIPS: user comments

Comments on changes to Sherpa interface were largely negative – many users preferred the old style CLI, and feel that documentation for the new Sherpa is still insufficient. We are reviewing how to support users who prefer a CLI interface and making improvements to the python interface and documentation. However, there were a number of positive comments on python support, and we are aware via helpdesk and personal interactions of more users transitioning in recent months. Presentation of the new capabilities to younger astronomers at the multiwavelength school was very positive.

Use of CIAO3.4 is still substantial – Sherpa use is currently about 50/50 between CIAO 3.4 and CIAO 4.x. A minority of users (17%) have a large number of old CIAO3.4 scripts; some users report they don't have motivation or time to learn a new interface.





CIAO User Survey – 4: Platforms

Platforms:

Linux: 64 bit now slightly more common than 32 bit

Fedora still dominant distro (FC2 to FC11 in use) (45/74)

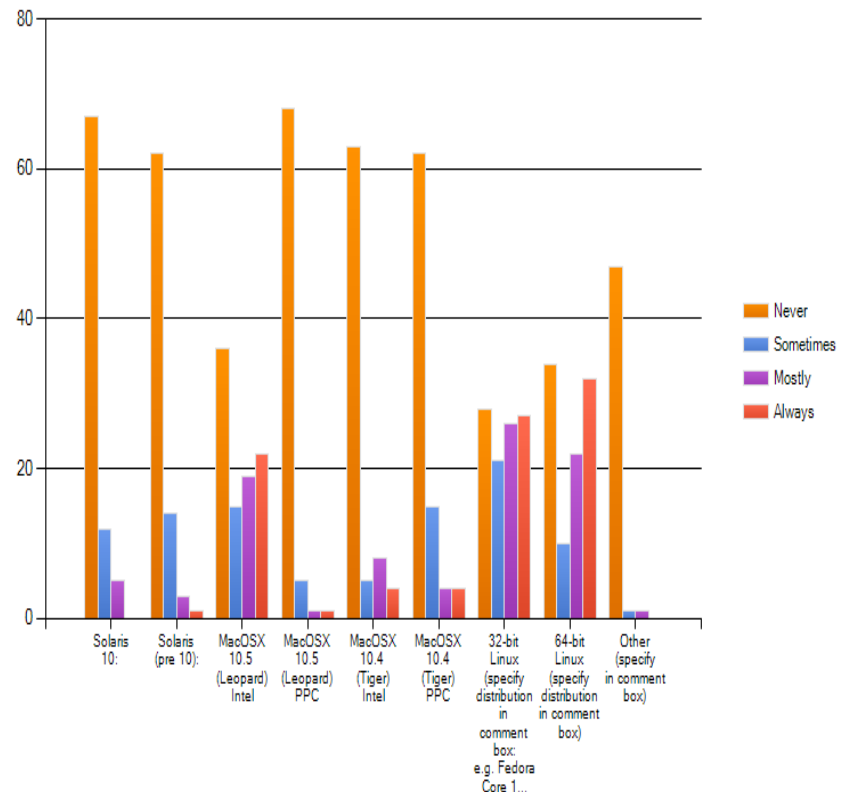
Debian, Ubuntu, SciLinux, RHEL, SUSE, CentOS

Mac: Of those minority of users who are planning a switch of their current OS in the near future, they are all intending to go to Mac Leopard.

Solaris use very small outside CfA

Of 2 institutions recording intent to install Solaris for multiple users, we know that at one (GSFC) users **actually** are using their Linux desktops. Also largely true at CfA, outside of CXC data aides and software group: Solaris CIAO is available but scientists are using Linux or Macs.

Which types of computer platform do you use for data analysis?
(including e.g. work, home, laptop machines)





CIAO 4.1.2

CIAO 4.1.2 was released in April as planned
It contained a set of Level 3 tools, plus bug fixes.

- Sherpa
 - CIAO 4.1.2 Sherpa updated to the latest XSPEC12.5 models
- Science Tools
 - **evalpos** Simple script to get image pixel value at a celestial position (e.g. for sensitivity file values)
 - **glvary** - Variability tool developed for catalog
 - **modelflux** – New tool based on catalog script, to convert between flux and rate for arbitrary Sherpa model
 - **srcextent** – Modified from catalog tool for better user interaction; measures effective size of source and PSF and estimates intrinsic size.
 - **streakmap** - Modified from catalog tool to accept more general user inputs; calculates streak background map due to CCD readout 'out-of-time' events.
 - **pileup_map** - Simple script to estimate pileup based on frame count rates.



CIAO 4.2

CIAO 4.2 (Dec 2009) internal drop 1 unit and regression testing completed

Internal drop 2 testing nearly complete

Sherpa improvements plus tools and bug fixes

- Catalog tools
 - background map tool, combining streak and Poisson mean algorithms
- Pipeline and instruments
 - `acis_build_badpix` Tool to make bad pixel files for your observation, adding user-requested improved abilities to customize
 - `acis_process_events` support for graded CTI [request from Cal team]
 - `hrc_process_events` support for new gain files



CIAO 4.2 - Sherpa

Sherpa improvements for CIAO 4.2:

- general improvements to user interface, error handling
- statistical functions for advanced fitting
- it works as a python package outside CIAO - Brian Refsdal presentation at SciPy meeting in Pasadena (Aug)

- 'Chi primini' statistics
 - Uses iterative weighting: variance derived from model values in previous iteration
- Support case where ARF is at higher resolution than RMF
- Routine which calculates simulations to generate flux errors
 - The 'sample_energy_flux' function takes the current best fit parameters and covariance matrix, and runs simulations with a multivariate Gaussian distribution on the parameters, then calculates the flux for each realization. This gives you the probability distribution of the flux. We then provide the ability to calculate the mean and standard deviation, and even arbitrary quantiles of the flux distribution.
 - This python script is already available to users for download, and documented at http://cxc.harvard.edu/sherpa/threads/flux_dist/index.py.html



CIAO 4.2 – Sherpa (2)

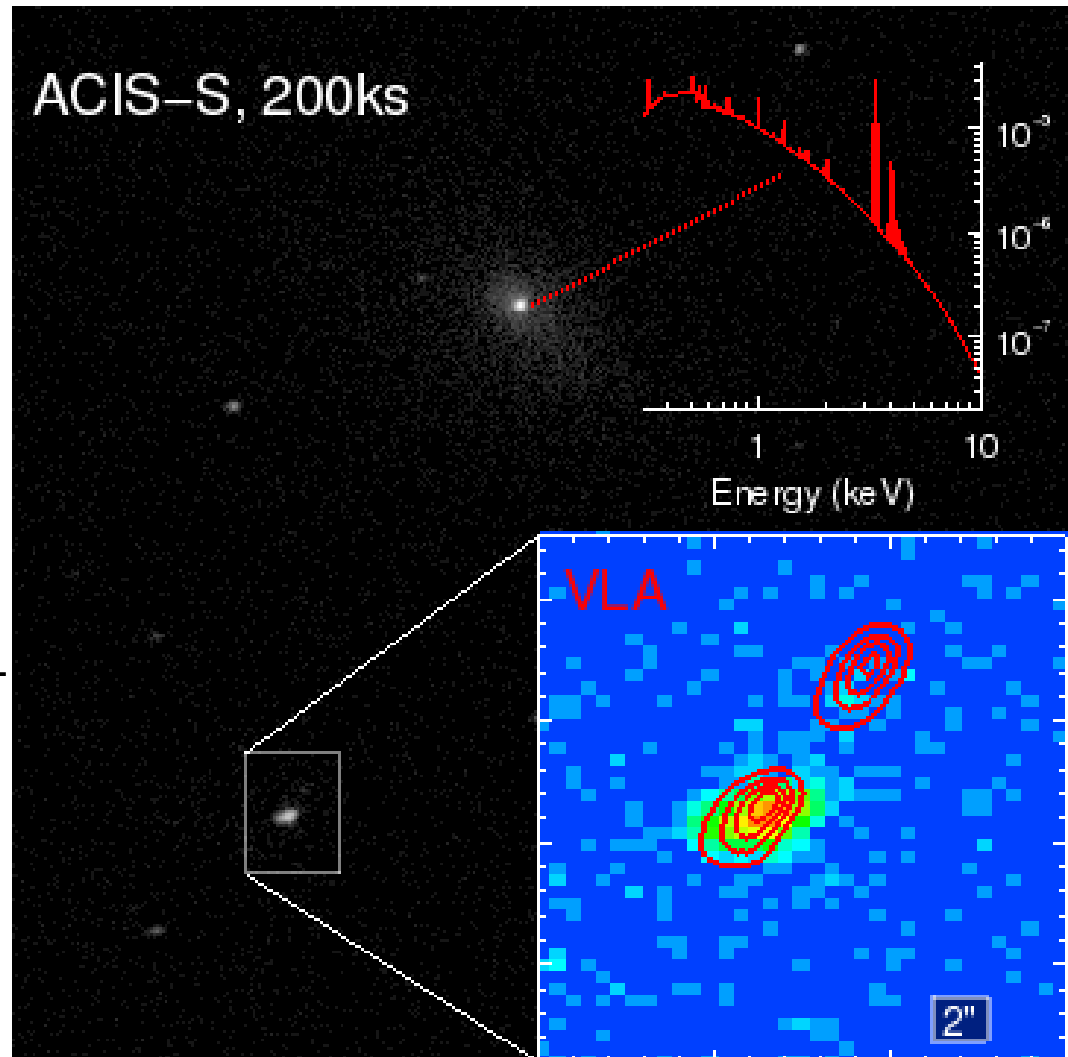
- Support for user statistics
 - `load_user_stat("priors", statfunc, errfunc, paramvals)`
 - Write your own statistic function for the minimization
 - Used to implement modern statistical minimization with Metropolis-Hastings (Bayesian algorithm)
 - Used to calibrate the flight operations ACIS thermal model
- General convolution with `load_conv()` to define your own convolution kernel
- New 'confidence' function for parameter uncertainties
 - More efficient than 'projection' and can handle cases which projection finds difficult
- Prototype support for calibration uncertainties
 - Randomize realizations of the ARF given cal errors
 - Derive combined statistical and calibration errors using simulations
- Support for parallelized computation of projection and confidence
- Added functions to make PSF fitting easier to use (`plot_kernel`, `show_psf`, `contour_kernel`, etc)



CIAO 4.2 - Plotting

CIAO 4.2 ChIPS

- SDS focus for CIAO4.2 is usability improvements
- Bug fixes and improvements to existing plot functionality
- Improved documentation
- Improved error messages
- Simplified common user operations
- Also testing and documenting the publication-quality plot support added to ChIPS in this release, e.g. allowing contours overlaid on image data



924 analysis ready observations

312 distinct objects:

631 HETG/ACIS,

92 LETG/ACIS,

201 LETG/HRC

139 GB archive

173 product downloads in last 6 months

Verification & Validation (V&V):

All products manually examined for zeroth order position, extraction region, & confusing sources.

Recently reprocessed to apply CALDB 4.1
Registered with VO; VO-table, VO cone-search support.

TGCat

RESULTS: Found 4 matching extractions

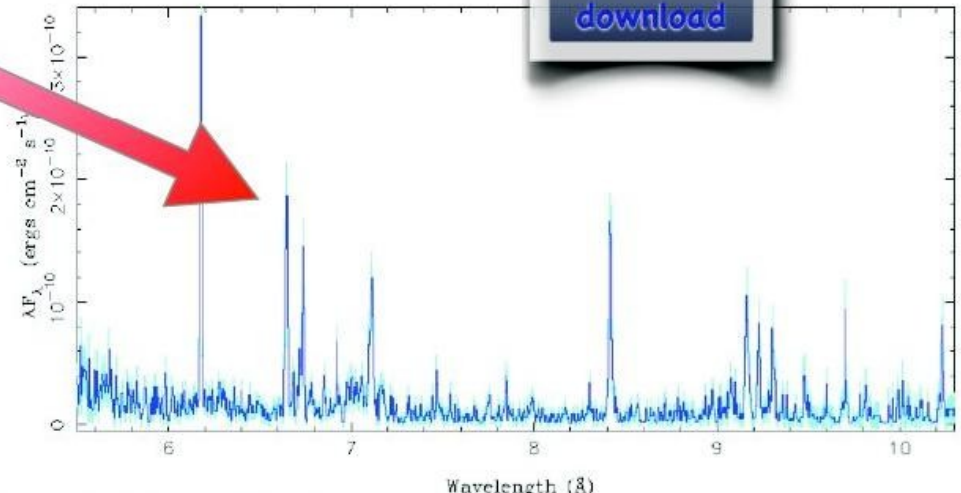
Links	YobsidA	YobjectA	YinstrumentA	YgratingA	Yra(h:m:s)A	Ydecl(d:m:s)A	Ydate_obs(y-m-d t)A	Yexposure(s)A
<input type="checkbox"/> 0-p-v-s	1926	Vela X-1	ACIS	HETG	09:02:6.838	-40:33:16.920	2001-02-11 21:20:17	84286.8
<input type="checkbox"/> 0-p-v-s	102	Vela X-1	ACIS	HETG	09:02:6.847	-40:33:16.776	2000-04-13 09:57:52	28018.8
<input type="checkbox"/> 0-p-v-s	1927	Vela X-1	ACIS	HETG	09:02:6.845		ObsID 1927 :: Vela X-1 :: 30 ks :: 2001-02-07 09:57:17	
<input type="checkbox"/> 0-p-v-s	1928	Vela X-						

Flux Spectrum (Click to preview all)

press "go" to operate on selections: limit download

Go Change Columns

Search,
sort, select,
preview,
plot,
download



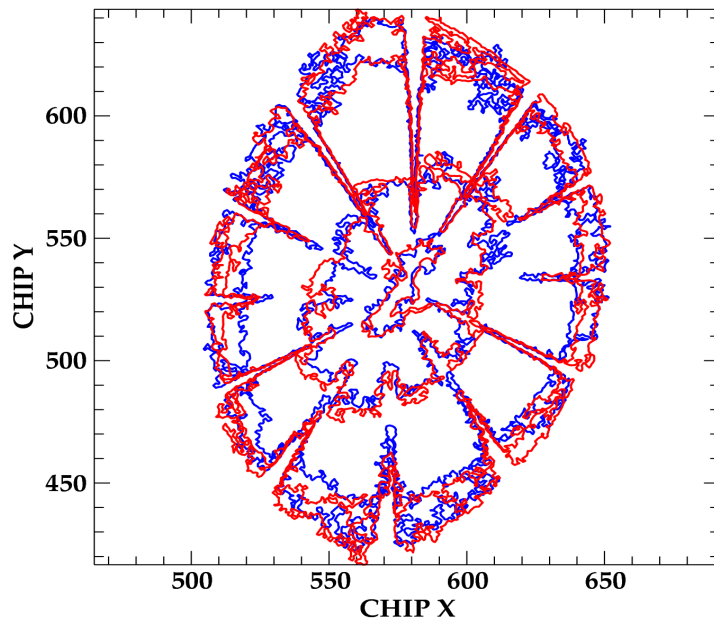


R&D: PSF

MARX and SAOTrace comparison

We have two PSF modelling codes: the SAOTrace detailed ray trace and the MARX approximate ray generator. MARX was recently improved to match SAOTrace better (e.g. include mirror support structures).
Next step: compare both with real data.

ACIS-S: $E = 1 \text{ keV}$, $\theta = 20'$ $\phi = 0^\circ$



Encircled energy contours at 1 keV

Far off axis (20 arcmin)

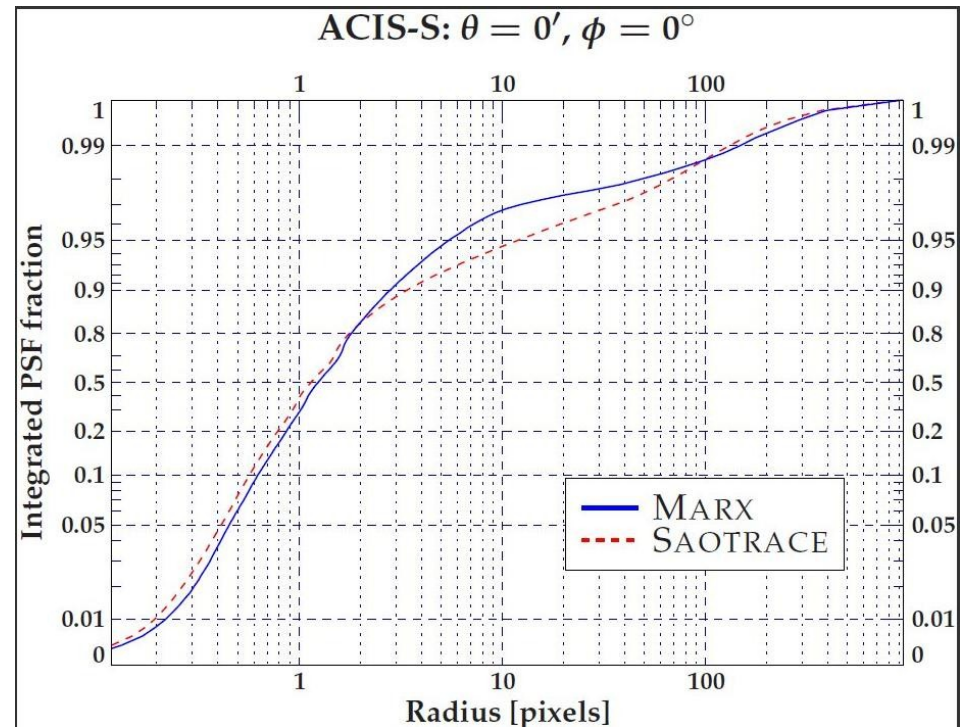
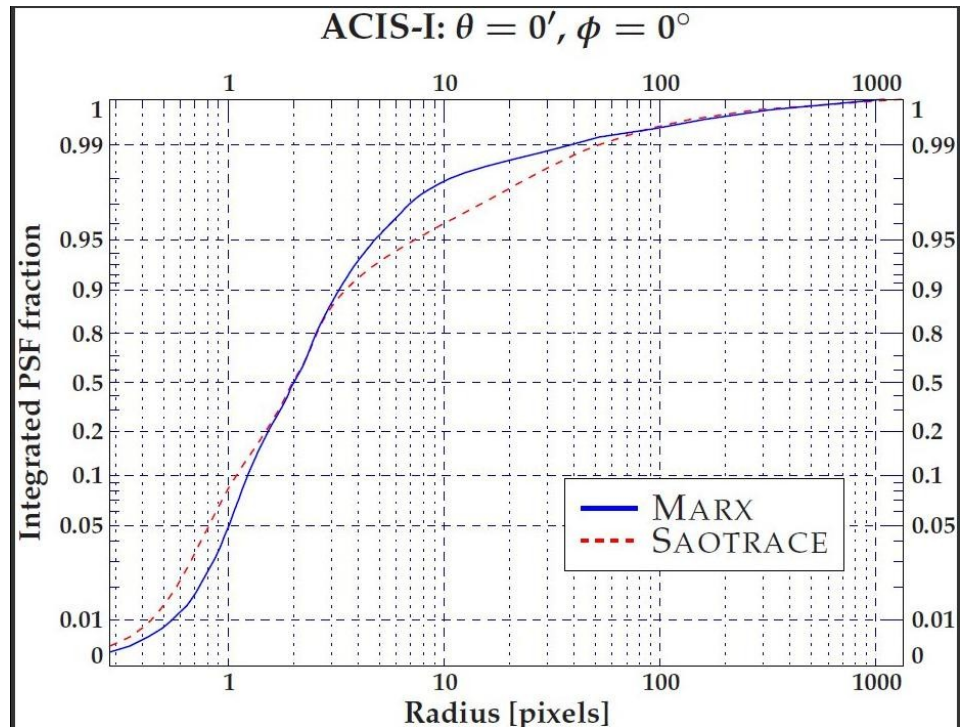
SAOTrace: red
MARX: blue



Aperture correction versus radius

- Good agreement in the 1 arcsecond region
- Problems for MARX in the subarcsecond regime but only small differences at larger radii
 - MARX PSF has more power at 3-5 arcsec:
 - SAOTrace has more at 5-20 arcsec, thanks to a better scattering treatment.

Vertical scale has been exaggerated in region of deviation





Modern PCs come with multiple cores, a capability we don't fully exploit right now.

Spectral fitting models for high resolution grating data are compute intensive.

We have been exploring approaches using parallel processing for both multiple-core machines and for heterogeneous networks.

A Python multi-core module approach is being adopted by Sherpa for CIAO 4.2 for more rapid projection and confidence interval calculations. We also see improvements on multiple cores from initial testing of this code.

MIT team produced a version of the XSTAR code using PVM (parallel virtual machine), so can efficiently generate XSTAR model tables for import into spectral fitting codes. See Miller et al 2009 (arXiv:0907.3114) for recent use.

Using ISIS as a testbed for scientist algorithm experiments, the MIT team implemented parallelized Levenberg-Marquardt optimization by forking of processes to n processors. Do see close to factor- n improvement in some demanding cases. Research continuing and will generate report on lessons learned.



Following user survey and internal discussions, **we have decided to phase out support for the S-Lang interface** in Sherpa and ChIPS – only the Python interface will be supported.

We are also migrating our CIAO scripts to python; of course, user scripts running CIAO tools in an slsh shell will continue to work, unless they use specific CIAO slang modules (which have not been widely used as far as we can tell.)



After CIAO 4.2 we do not expect to support Mac PPC

Linux and Mac-Intel are now our main user platforms. The 64-bit Linux build is coming soon.

We have improved the SDS testing process, but for CIAO 4.2 we expect to test Linux, Mac PPC, Mac Intel, Solaris and perhaps Linux64 – this is too many given the resources we have.

What do we do about Solaris? For now, will continue to support since heavily used by our ops team. But over the next few years the user base will continue to shrink – can we provide some kind of reduced support (less testing, etc) ?

The biggest burden for each supported platform is the maintenance of external software dependencies – e.g. cfitsio, GTK, fftw libraries. We build these as binaries for each port, and they must be kept up to date and in sync with the system. Building them from source is not trivial, which is why we do it for you. Eventually we may stop active support for some platforms and require you to build them and their external libraries from source.