Progress Report

Ian Evans
On behalf of the Chandra Source Catalog Project Team

Chandra Users' Committee Meeting
October 13, 2011

Executive Summary

Summary

- Current catalog version: 1.1; Released: 2010 Aug 10
  - 106,586 master source
    - Includes 104,628 ACIS-only, 1,034 HRC-only, 924 both ACIS and HRC
  - 158,071 source detections
    - Includes 152,296 ACIS, 5,775 HRC
  - 5,110 observations with at least one detected source
- Statistical characterization of rel. 1.1 is complete
- Team is working on algorithm development for rel. 2.0
  - Science/software team has developed a prototype source detection pipeline to validate new approach
  - Science team is working algorithm refinements for source properties analysis
  - "High-risk" items have largely been addressed or a path forward is believed to be understood

Science Highlights Since Last CUC Meeting

- Completed statistical characterization of rel. 1.1
  - Statistical properties of release 1 are available on the catalog web site
  - Paper describing the statistical properties of rel. 1.1 (Primini, F. A. et al.) was published in ApJS (see below)
- Updated public web site with latest user documentation and threads
  - **http://cxc.cfa.harvard.edu/csc/**
- IPAC updated the NASA Extragalactic Database (NED) to include basic data for CSC release 1.1 sources
- Working algorithm development and prototypes for release 2.0
- Publications
  - CSC absolute astrometric error paper
    - *Rots, A. H. & Budavári, T. 2011, ApJS, 192, 8*
  - Catalog statistical characterization paper
    - *Primini, F. A., et al. 2011, ApJS, 194, 37*
  - Aperture photometry algorithm paper
    - *Kashyap, V. & Primini, F. A. 2011, in preparation*

Software Highlights Since Last CUC Meeting

- Released several archive and user interface updates to support rel. 1.1
- Continued developing prototype rel. 2 source detection pipeline including Sherpa Maximum Likelihood Estimator (MLE)
- Evaluated rel. 2 software infrastructure and began infrastructure work
- Managed pipeline operations in support of algorithm development

Catalog-Related Releases

CAT	3.2.6	22 Nov	CSCview add cross-match capability; CSC 1.1 web services updates (LimSens, CSC Sky, Xmatch)
CAT	3.2.6.1	4 Feb	LimSens security update
CSC Sky	1.2	15 Feb	DMZ-related updates
CAT	3.2.7	28 Jun	Sybase 15.5 updates

- Publications
  - The Sherpa Maximum Likelihood Estimator
    - *Nguyen, D., et al. 2011, ASP Conf. Ser., 442, 517*

Chandra Source Catalog Usage Statistics

	2010 Sep 01 – 2011 Feb 28		2011 Mar 01 – 2011 Aug 31	
	Number	**% Non-CfA**	**Number**	**% Non-CfA**
CSCview catalog browser initializations	314 /month	67%	381 /month	81%
VO cone searches*	1518 /month	80%	2727 /month	99%
CSC Sky in Google Earth	103K hits/ month	9252 visits/ month	128K hits/ month	2846 visits/ month

* Excludes 254K cone searches from a single non-CfA user in 2010 Dec and 4.6M cone searches from a single internal user in 2011 Jun

Usage Statistics 2010 Sep 01 – 2011 Aug 31

- Statistics for different interfaces are not directly comparable
- CSC Sky interface includes a significant mix of professional and non-professional usage; other interfaces are primarily professional usage
- Increasing number of CSC Sky page hits with an associated decreased number of visits suggests that the fraction of "real" usage is increasing relative to casual accesses
  - During the period 2011 Mar 01 – 2011 Aug 3, CSC Sky dropped off Google's "featured sites" list

Comparison of Actual CSC Rel. 1.1 and Expected Rel. 2.0 Performance

	Release 1.1	*Release 2.0*
Number of Sources	106,586	~ 280,000
Limiting Net Source Counts (point source, on-axis)	~ 10	~ 5
Instrument Data Included	ACIS and HRC-I Imaging	Same
Number of Observations	5,110	~ 6,500
Source Detection Runs On	Single observation	Stacks of observations with the same instrument and pointings within 60 arcsec
Source Sizes Included	$\lesssim$ 30 arcsec	$\lesssim$ 60 arcsec

Enhancements Compared to Release 1.1

- Include latest calibrations
  - Updated photometric calibration with "new" QE, gains, contamination
  - New afterglow correction
  - ACIS sub-pixel resolution
- Significantly improve source detection process
  - Stack observations with same detector configuration and similar pointings (within 60″) prior to source detection
    - Observations are corrected to common astrometric frame before stacking
  - Run `wavdetect` with "permissive" parameters followed by Sherpa MLE grading
    - Allows detection to faint limit (~5 counts on axis with 35 ks background) with possibly significant false source rate (later resolved by MLE)
  - Run `wavdetect` on single blocking factor data with √2 scale factor sequence
    - Generates better source regions and eliminates issues associated with merging multiple blocking factor runs, at cost of significantly increased memory usage (need ~16 GB to run)
  - Use Sherpa MLE to evaluate source candidate likelihood
    - Likelihood value for source determines whether source is definite, questionable, or should be rejected as false
    - Requires substantial "up-front" simulation work (unlike rel. 1 which required mostly post-facto characterization simulations)

Enhancements Compared to Release 1.1 (cont.)

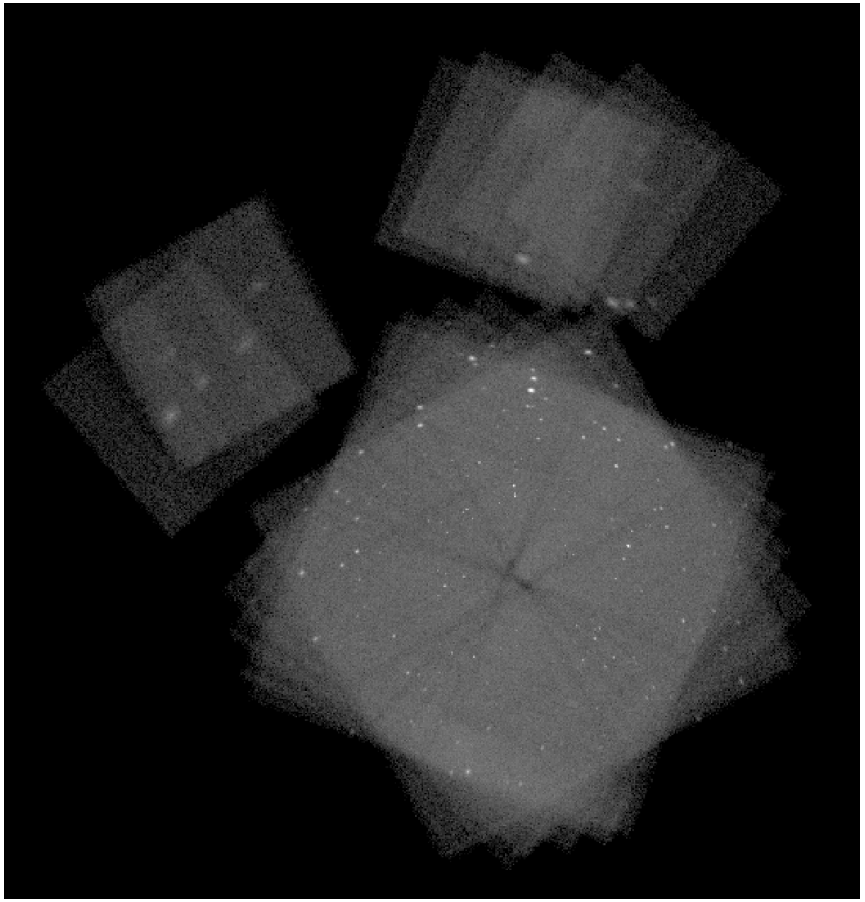
- Improve source detection background maps
  - Exclude exterior detector edges and adjacent ACIS BI/FI chip edges
    - Many rel. 1 false sources "detected" at these boundaries
  - Investigating new Voronoi tessellation/Delaunay triangulation background algorithm
    - Promising, but still evaluating performance in crowded fields
- Enhance aperture photometry algorithms
  - Use spectrally weighted ARF and apply energy-dependent ECF corrections
  - Simultaneously determine intensities of nearby/overlapping sources and background using a Bayesian formalism
  - Improve method for merging multiple intensity values for a source by using informative priors
  - Revise limiting sensitivity algorithm for consistency with rel. 2 approach
    - Approach based on MLE fit statistic similar to 2XMM
- Include upper limits from non-detections in temporal variability statistics
- Enhance software infrastructure
  - Redesign processing pipelines to support observation stacks
  - Update support infrastructure (e.g., automated processing, archive, database, and UI's) to support access to new properties and data products derived from observation stacks

CSC Release 2.0 Status

Current Status

- Science team: working on release 2.0 tasks
  - Enhancing the source detection thread and Sherpa MLE algorithm
  - Enhancing rel. 1 aperture photometry and limiting sensitivity algorithms
  - Developing Voronoi tessellation/Delaunay triangulation background
  - Investigating algorithms for handling upper limits in time variability
  - Developed code based on MARX to generate multi-ObsId simulations
    - Inject point and elliptical Gaussian extended sources
    - Tag source photons for downstream analysis
- Software team: working on release 2.0 tasks
  - Developed prototype source detection pipeline including updated `wavdetect` and Sherpa MLE for functional testing and parameter tuning
    - Prototype works on simulations and observations included in CSC rel. 1
  - Defining pipeline architecture and requirements for processing system
  - Working modifications to automated processing (AP) infrastructure to run release 2.0 more efficiently
  - Established a new software build to support rel. 2.0
- Release 2.0 hardware
  - Upgraded existing "morris" cluster memory to support testing
  - Currently evaluating new hardware requirements for rel. 2.0 production

Enhanced `wavdetect` and MLE

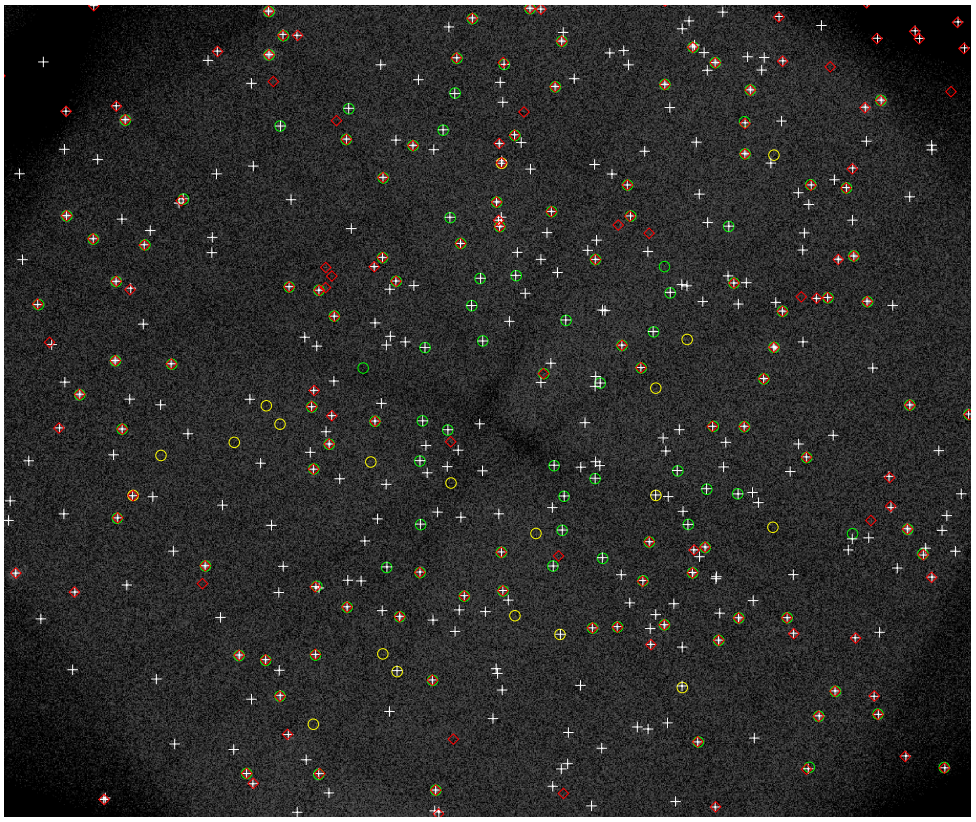


Stack of eight ACIS observations (ObsIds 02312, 02313, 02405, 02406, 08594, 08595, 08596, 09596) of the CDFS totaling 827 ks exposure

- Observation stacks must all have the same instrument configuration (i.e., cannot mix ACIS and HRC), and be co-pointed within 60″

CSC Release 2.0 Source Detection (cont.)

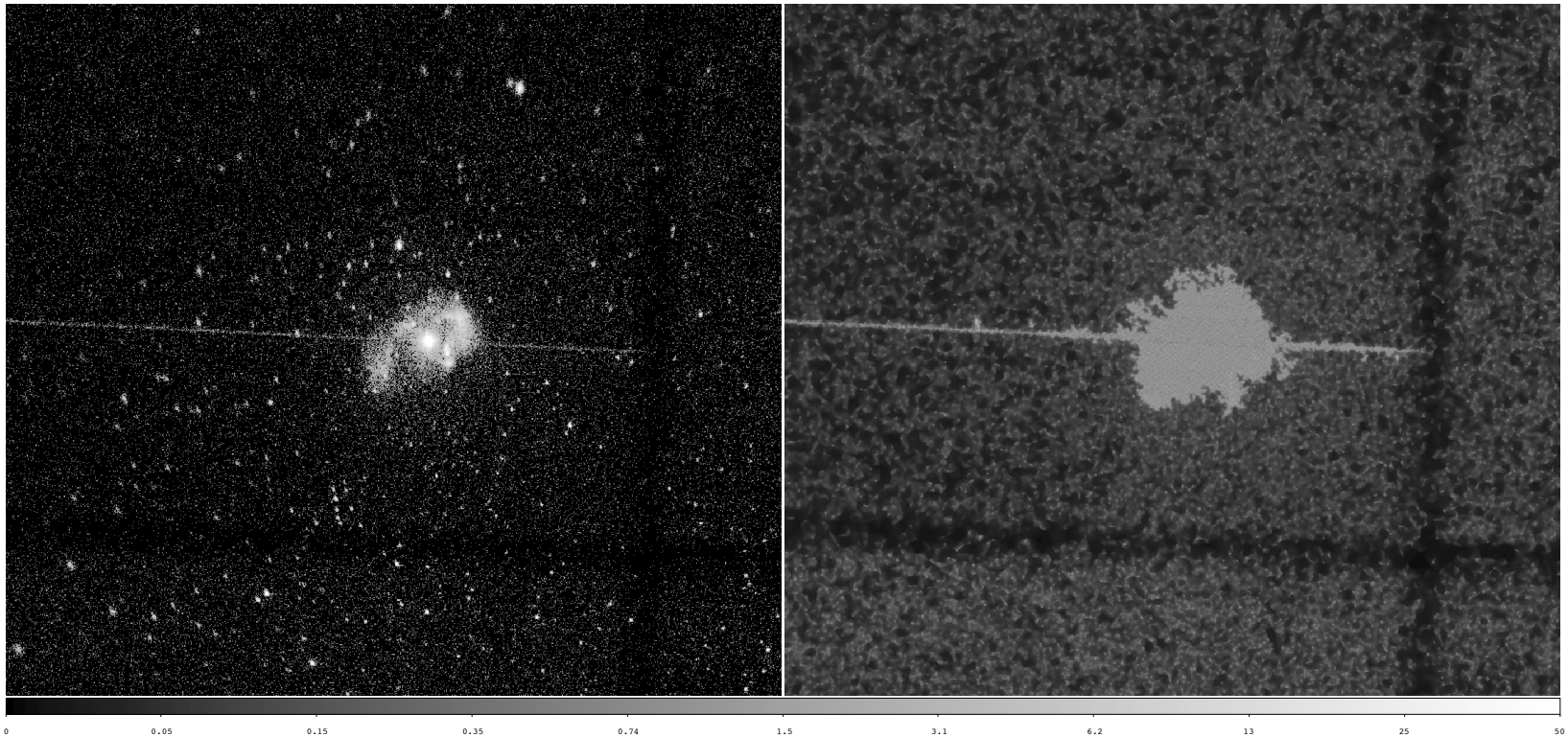
Enhanced `wavdetect` and MLE (cont.)



Sources in the central region of Chandra Deep Field South (stack of 8 ObsIds totaling 827 ks)

- White crosses indicate source positions from the published 2 Ms catalog (2008 ApJS, 179, 19)
- Source regions displayed as red diamonds are from CSC release 1.1
- Source regions displayed as green circles were detected in the broad, hard, or medium energy bands using the prototype CSC rel. 2.0 pipeline and have log likelihood ≥ 9 (~ questionable or better sources)
- Source regions displayed as yellow circles were detected in the soft energy band

Voronoi Tessellation/Delaunay Triangulation Background

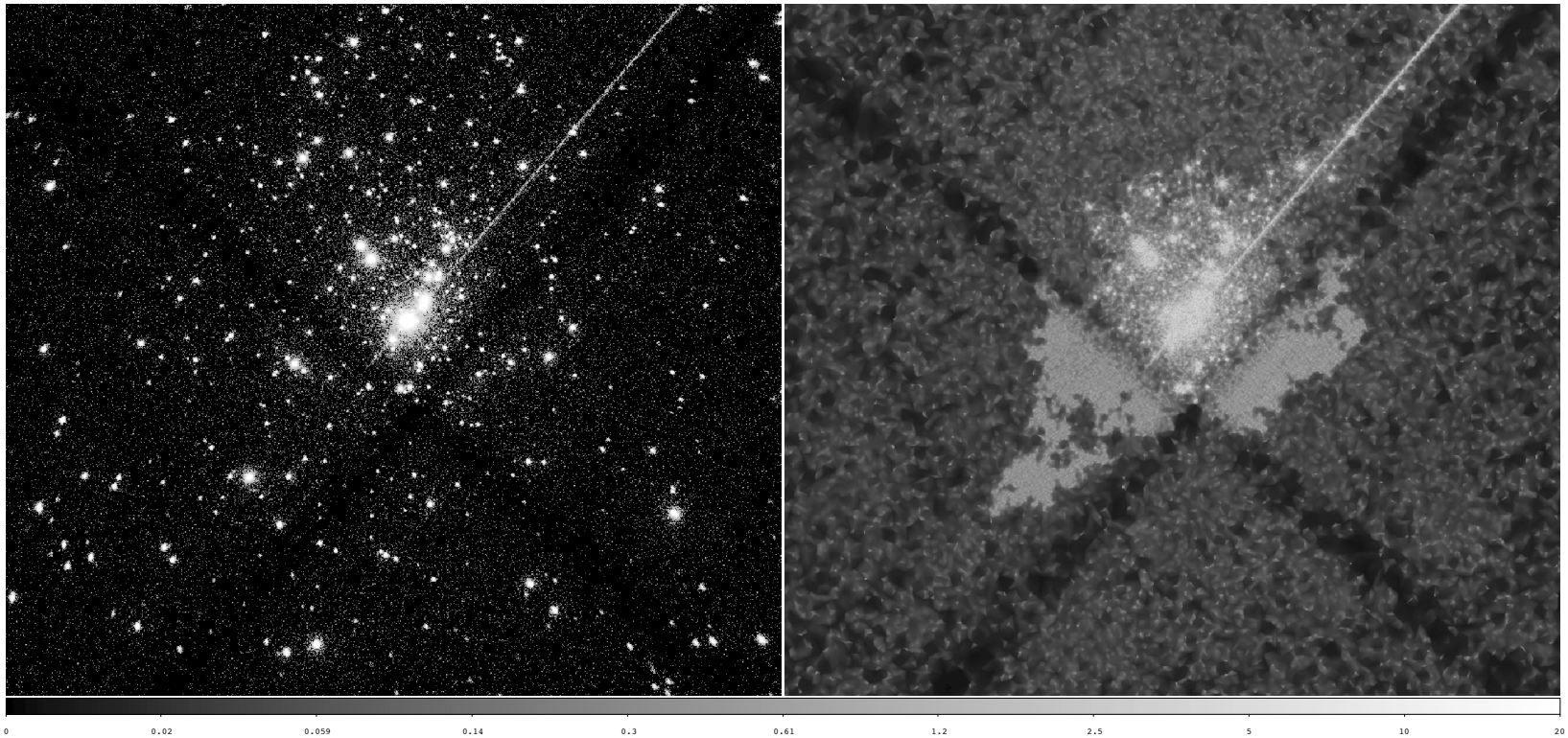


Left: Central region of Trumpler 16 (ACIS ObsId 06402); *Right:* Background map derived using VT/DT

- VT/DT background effectively identifies readout streak and enhanced emission background in the vicinity of Trumpler 16
  - The background can be used as input to `wavdetect` to improve the detection efficiency for point/compact sources overlying the extended emission
  - The boundary of the enhanced emission background *may* be usable as a region definition for identifying extended emission components that *could* be included separately in rel. 2.0

CSC Release 2.0 Backgrounds (cont.)

Voronoi Tessellation/Delaunay Triangulation Background (cont.)



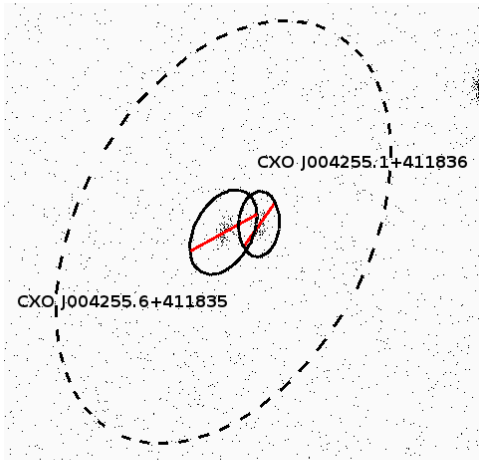
Left: Central region of the Orion Nebula Cluster (ACIS ObsId 04395); *Right:* Background map derived using VT/DT

- VT/DT background is effective in areas with low-medium source density, but currently evaluating performance in crowded fields
  - Some point/compact sources appear to be included in the background, which may reduce their detectability

Aperture Photometry Algorithm Enhancements

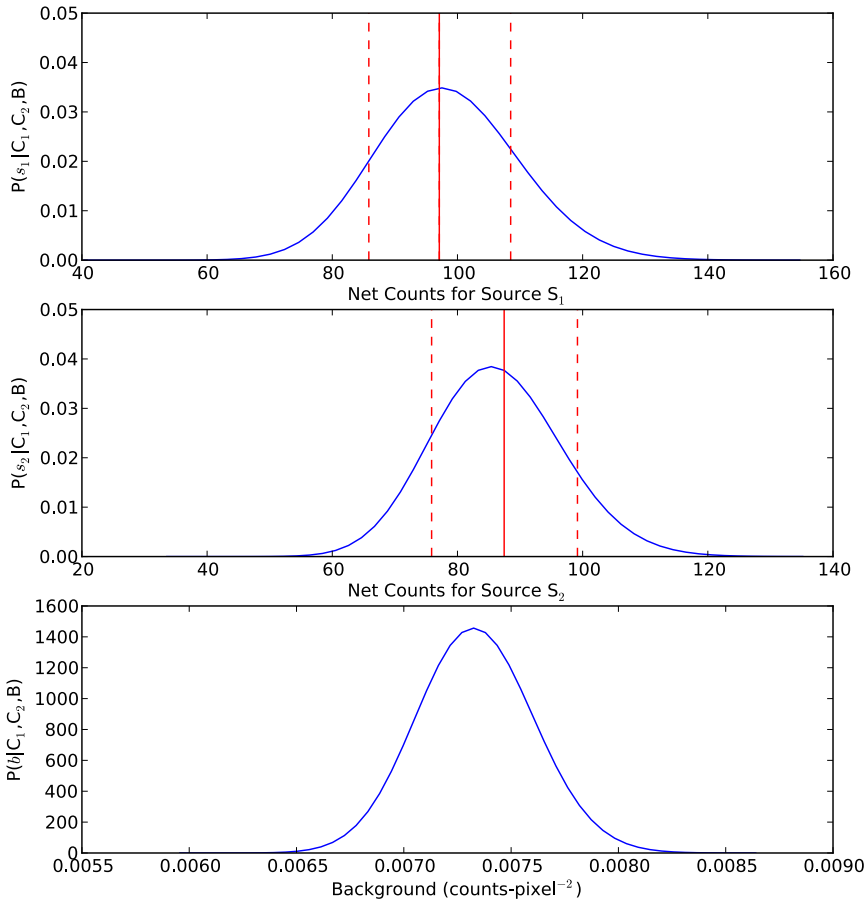
Below, Left: Two nearby sources with overlapping source regions and surrounding background region included in CSC rel. 1 (ObsId 1912)

- The fractions of each source in each source region and the background region must be evaluated
  - For point sources these can be determined from the PSF models, but for extended sources must be estimated independently
- A Bayesian formalism is used to solve for all of the source and background PDFs simultaneously
- The new algorithm is more robust and easier to use than the rel. 1 algorithm that is included in the CIAO tool `aprates`



Below, Right: Normalized PDF's for the source intensities and background level

- The rel. 1 source intensities and confidence limits are shown in red



CSC-Derived CIAO Tools

- Rel. 1 resulted in the release of 14 new CIAO tools
  - These tools were originally developed to perform tasks necessary to construct CSC rel. 1, but were later deployed for users in CIAO 4.1–4.2
  - Several of these tools provided CIAO users with the capability to easily extract source properties for the first time since launch
    - Example source properties include aperture photometry fluxes, source variability statistics, and source extent measures
  - These tools were developed for the CSC, but are not catalog-specific
    - They can be used with CIAO users' own data
- Rel. 2 will provide additional new capabilities that could be released to CIAO users
  - The Sherpa MLE tool and the associated source detection/likelihood grading thread should greatly simplify and significantly improve source detection, especially for multiple observations split because of thermal constraints
  - The Voronoi tessellation/Delaunay triangulation background algorithm should simplify construction of background maps
  - Existing aperture photometry, source variability, and source extent tools will be enhanced for wider applicability and improved accuracy
- Like rel. 1, we expect that such CIAO tools could be delivered with or soon after the completion of rel. 2 of the catalog

Future Directions

Short Term Plans

- Rel. 2.0 development
  - Continue working rel. 2.0 algorithm and software development
  - Aim to begin rel. 2.0 processing no earlier than fall 2012
- Public Interfaces
  - External Interfaces
    - Continue discussions with outside agencies (CDS, HEASARC) aimed at providing access to a subset of rel. 1.1 catalog master source properties through those interfaces

Long Term Plans

- Catalog releases
  - Future releases
    - Simultaneous source detection across overlapping observations with different detectors and pointings (and thus very different local PSFs)
    - Detection and classification of very extended sources