CHANDRA SOURCE CATALOG

# Progress Report

Ian Evans
On behalf of the Chandra Source Catalog Project Team

Chandra Users' Committee Meeting
October 22, 2013

# Current Catalog Release



## Summary

- Current catalog version: 1.1; Released: 2010 Aug 10
  - 106,586 master sources
    - 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

## Usage Statistics

| Release 1.1 | 2012 Oct 01 – 2013 Mar 31 | | 2013 Apr 01 – 2013 Sep 30 | |
|---|---|---|---|---|
| | Number | % Non-CfA | Number | % Non-CfA |
| CSCview catalog browser initializations | 379 /month | 85% | 180 /month | 90% |
| Command-line (CLI) searches | 728 /month | 4% | 864 /month* | 32% |
| VO cone searches | 9441 /month | 79% | 9763 /month | 77% |
| CSC Sky in Google Earth | 412 visits/ month | | 416 visits/ month | |

*Excludes ~75K (~100% non-CfA) searches in Sep. 2013

# Current Catalog Release Support

## CSCview

- Updated version of CSCview supports Java 7 (aka Java 1.7)
- Now available as a downloadable application for Mac
  - Avoids OS X limitations on Java applet execution in browser

## Command Line Tools

- New tools distributed as part of the contrib package support the most common queries using a CIAO syntax interface
  - `search_csc` performs position-based queries
  - `obsid_search_csc` performs ObsId-based queries

## Documentation

- Updated existing threads and added a new science thread "Investigating Colors of Variable Galactic Sources"

CHANDRA SOURCE CATALOG

# Release 2 Summary

CXC

| | |
|---|---|
| ***Expected Number of Distinct Sources on the Sky*** | ~ 280,000 |
| ***Expected Number of Source Detections*** | ~ 380,000 |
| ***Expected Number of Observations*** | ~ 7,000 |
| ***Expected Limiting Net Source Counts (point source, on-axis)*** | ~ 5 |
| ***Instrument Data Included*** | ACIS and HRC-I Imaging |
| ***Source Detection Runs On*** | Stacks of observations with the same instrument, and pointings within 60 (TBR) arcsec |
| ***Sources Types Included*** | Point and compact sources in regions without extended emission<br>Regions of extended emission delineated by convex hulls<br>Convex hull thresholds TBD<br>Point sources within extended emission convex hulls<br>Detection sensitivity may be reduced in these regions |

## Science Highlights Since Last CUC Meeting

- Developed prototype algorithms and specifications for release 2.0
  - Aperture photometry specifications for individual- and combined-observation analysis
  - Limiting sensitivity specification
  - Extended (convex hull) source likelihood specification
  - Scargle Bayesian Blocks specifications for partitioning stacks into groups of observations for combined analysis
- Updated `mkvtbkg` background and extended source detection tool
  - Include ACIS readout streaks in the background
  - Smooth the (non-streak) background with a Gaussian whose σ depends on θ
- Supported pipeline development and testing
  - Evaluated numerous pipeline test runs and provided feedback to software team

## Software Highlights Since Last CUC Meeting

- Continuing development of release 2.0 pipelines and tools
  - Updated `mle` to add estimation of confidence limits on all parameters
  - Integrated `mkvtbkg` backgrounds and added convex hull detection
  - Added limiting sensitivity pre-population
  - Added cratered source (ACIS pile-up) detection
  - Added quality assurance processing for all pipelines through "stacker"
  - Implemented data archiving for all pipelines through "stacker"
  - Added support for HRC-I observations
  - Working later-stage tools as specs become available
- Supported extensive science testing of initial implementation through "stacker" pipeline with actual and simulated data
  - Identified and resolved or worked around unexpected issues as necessary

# Pre-Production Testing



## Test Data

- Extensive pre-production testing has been primarily based on a set of ~20 observation stacks, selected to validate pipeline performance
  - Test stacks are chosen to exercise specific capabilities, ensure that pipeline processing is robust, and can handle all expected input data
  - Tests have identified many unexpected issues when processing some data

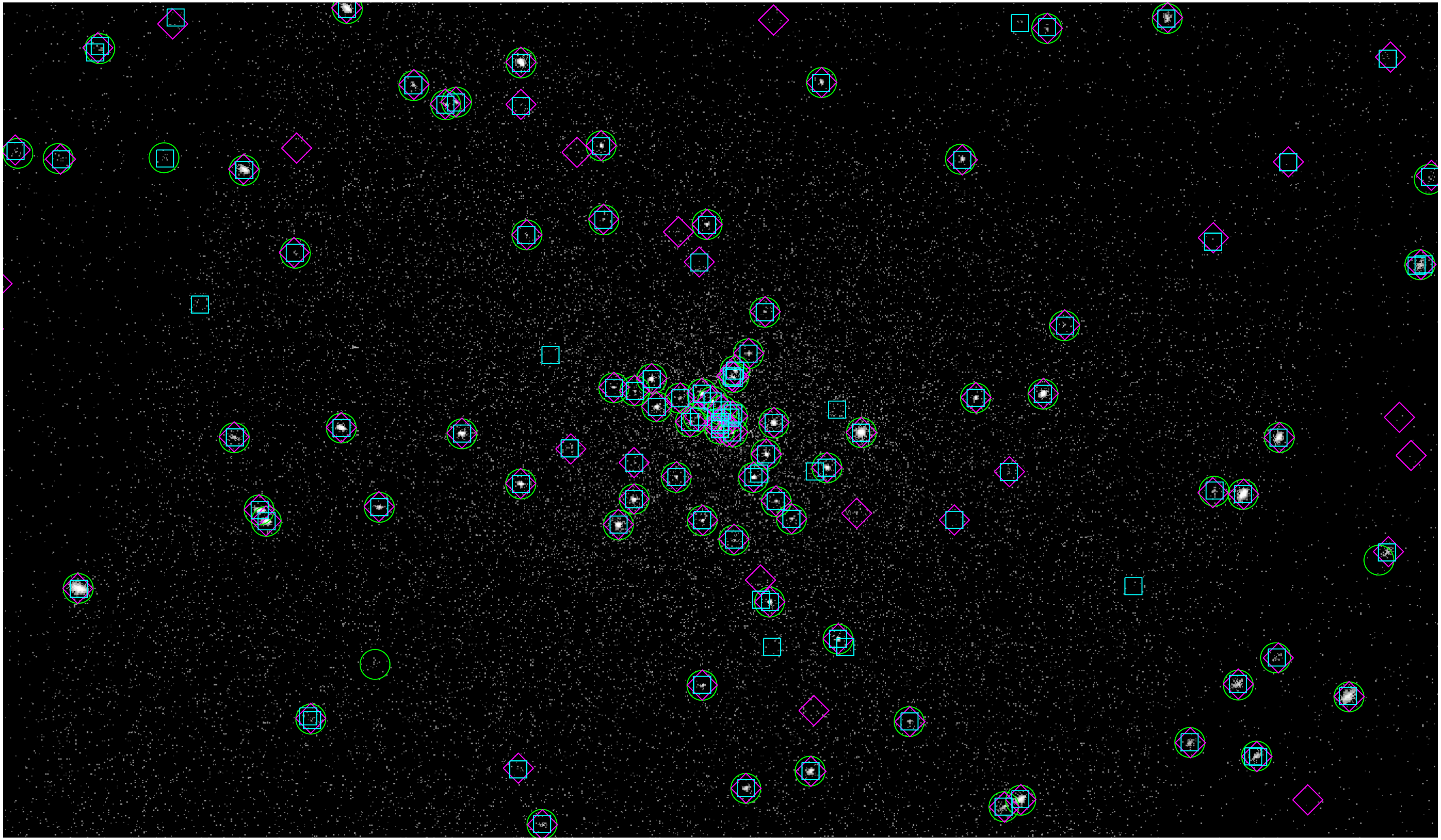
```
# PSR B1929+10
000160 acisfJ1932139p105929_001 06657_000,07230_000
# SN 1993J
000190 acisfJ0955229p690132_001 00735_001,09122_000,12301_000
# Orion (streak in same direction)
000512 acisfJ0535130m052259_002 03498_000,03744_000,04373_000,04374_000,04395_000,04396_000
# 0 degree RA overlap
000600 acisfJ0000082p135624_001 11490_000
# 360 degree RA overlap
000610 acisfJ2359567p004249_001 11591_000
# Polaris, high declination
000620 acisfJ0224387p891406_001 06431_001
# Subarray, zero sources?
000630 acisfJ0348574p125529_001 02158_000
# Extended source on various detector edges
000640 acisfJ0412549p102027_001 06929_000,07217_000,07218_000,07222_000,07234_000,07235_000
# Extended source
000650 acisfJ0602180m395930_001 03202_002,03450_002
# Faint, few sources
000660 acisfJ1120056p570223_001 06960_000
# Deep extended emission w/ point sources
000670 acisfJ1259510p275430_001 13993_000,14410_000
```

## Test Data (cont.)

```
# 1.0 keV low energy ACIS event threshold
000680 acisfJ1742590m293012_001 02283_000
# Sgr A* stress test, 02943, 02953, 05954 interrupted by SCS107
000690 acisfJ1745400m290026_001 02943_000,02951_000,02952_001,02953_001,02954_000,03392_001,
       03393_000,03549_000,03663_000,03665_000,04683_000,04684_000,05360_000,05950_000,05951_001,
       05952_001,05953_001,05954_000,06113_000,06363_000,06639_000,06640_000,06641_000,06642_000,
       06643_001,06644_000,06645_000,06646_001,07554_001,07555_000,07556_000,07557_000,07558_000,
       07559_000,09169_000,09170_000,09171_001,09172_001,09173_000,09174_000,10556_000,11843_000,
       13016_000,13017_000
# M17
000700 acisfJ1820296m161055_001 00972_002,06403_001,06420_000,06421_000,08460_000,08461_000
# M31
000710  hrcfJ0042403p405201_001 00260_000,00261_000,00263_000,00264_000,00265_000,00266_000
# M31
000720 acisfJ0042432p411635_001 00310_000,00312_000,04360_000
# NGC 1068
000730  hrcfJ0242412m000034_001 12705_000
# NGC 1068
000740 acisfJ0242410m000123_001 00344_000
# NGC 2903
000750 acisfJ0932079p212854_001 11260_000
# Large Y-offset (is the local PSF computed correctly?)
000760  hrcfJ1033161p534103_001 01400_000
# Few sources
000770  hrcfJ1744519m283909_001 06194_000,06195_000,08533_000,09032_001,09038_000
# Central source; 08547 contains an instrumental feature that is removed in L2 and should be
# removed in L3 also
000780  hrcfJ1520426m571003_001 08547_000,08556_000
```

# Rel . 2 Source Detection Example

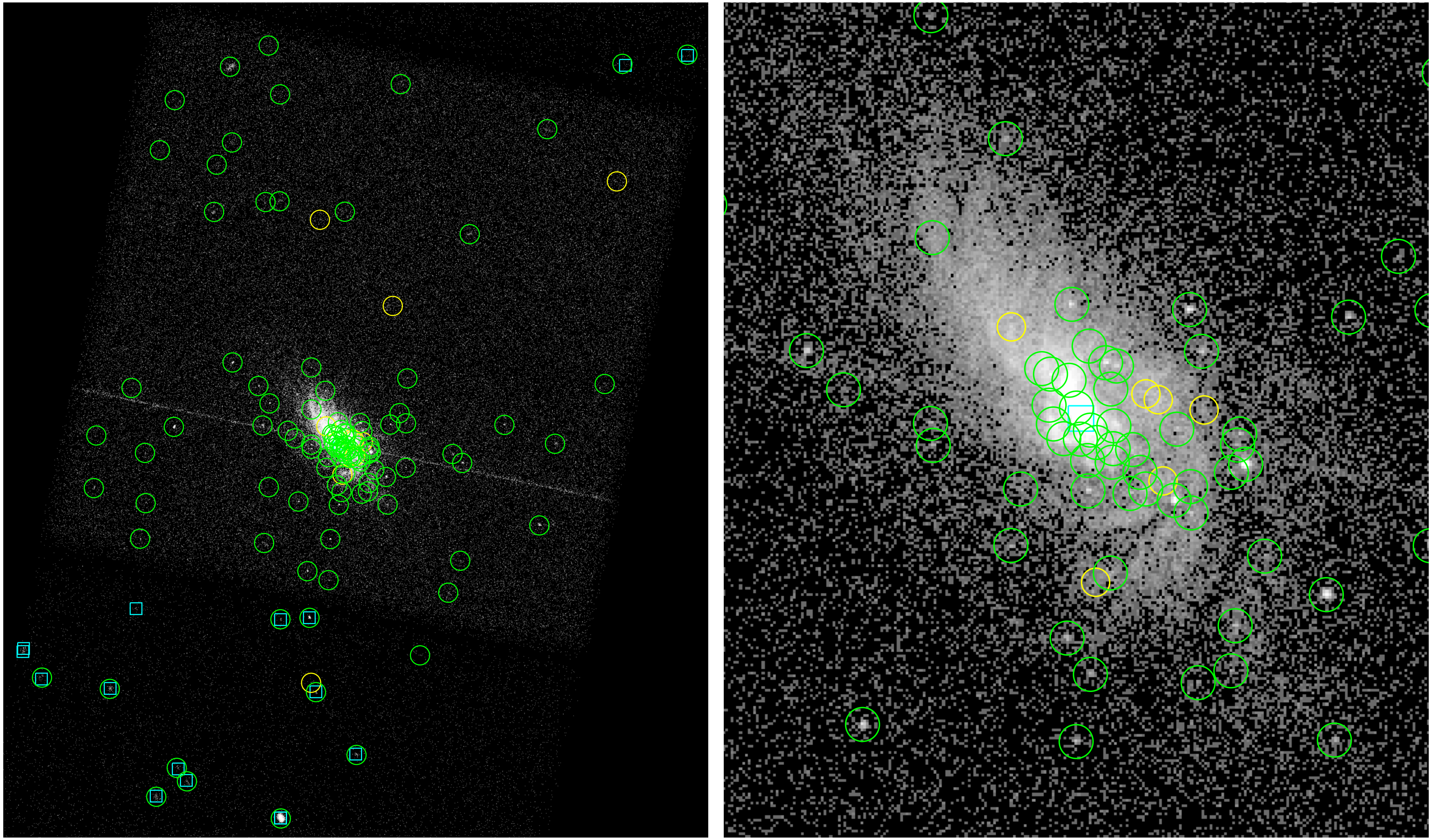




- Core of M31, StackId `acisfJ0042432p411635_001` (white light), Rel. 1 ‘`b`’ band sources shown with cyan squares, Rel. 2 sources shown with green circles, marginal sources shown with yellow circles, Kong et al. (2002) M31 *Chandra* point source catalog sources shown with magenta diamonds

# Rel. 2 Source Detection Example

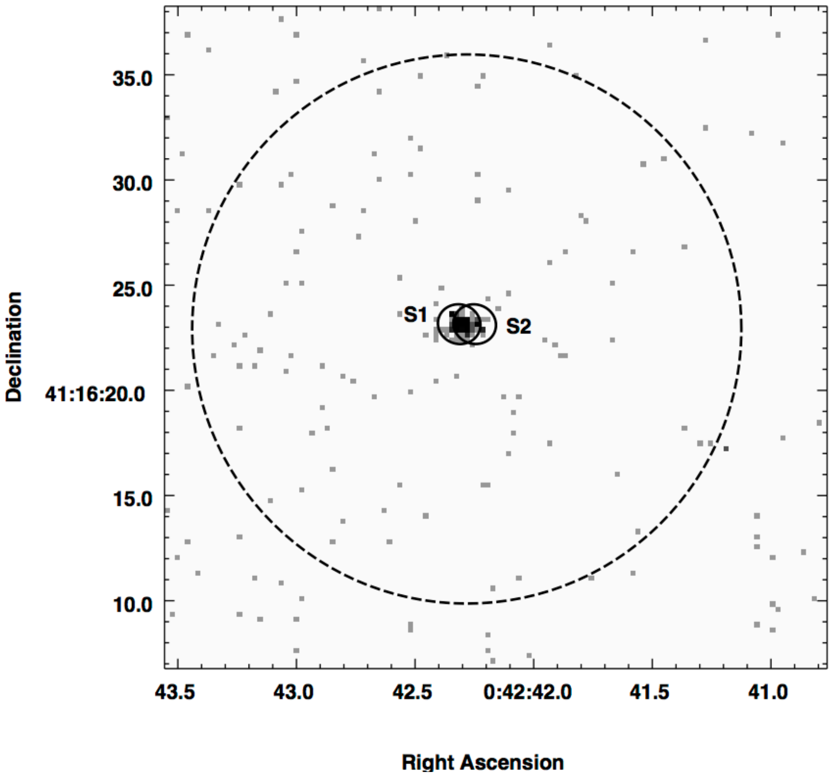


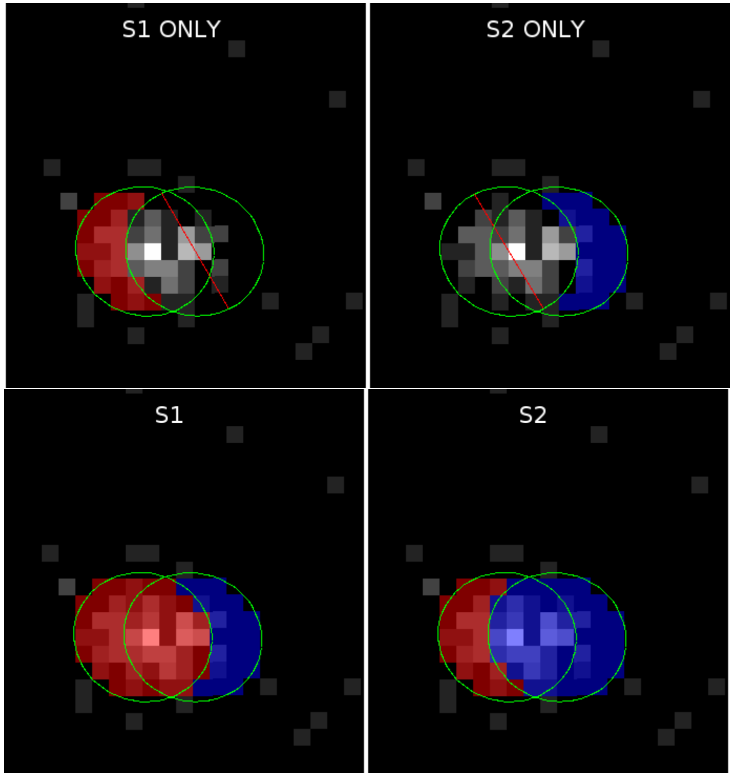


- NGC 1068, StackId `acisfJ0242410m000123_001` (white light), Rel. 1 ‘`b`’ band sources shown with cyan squares, Rel. 2 sources shown with green circles, marginal sources shown with yellow circles; note the dramatic increase in source detections on the CCD that contains the galaxy

# Rel. 2 Aperture Photometry



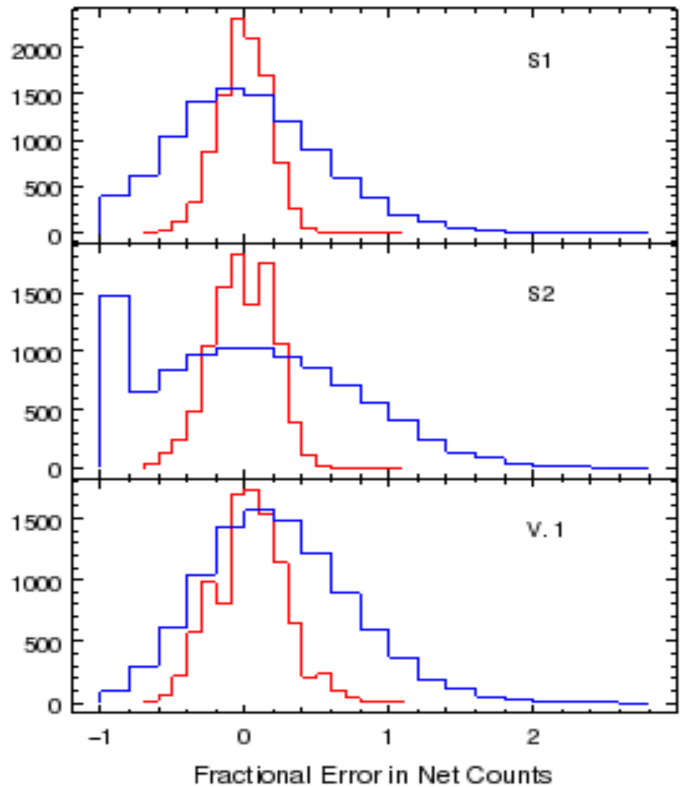


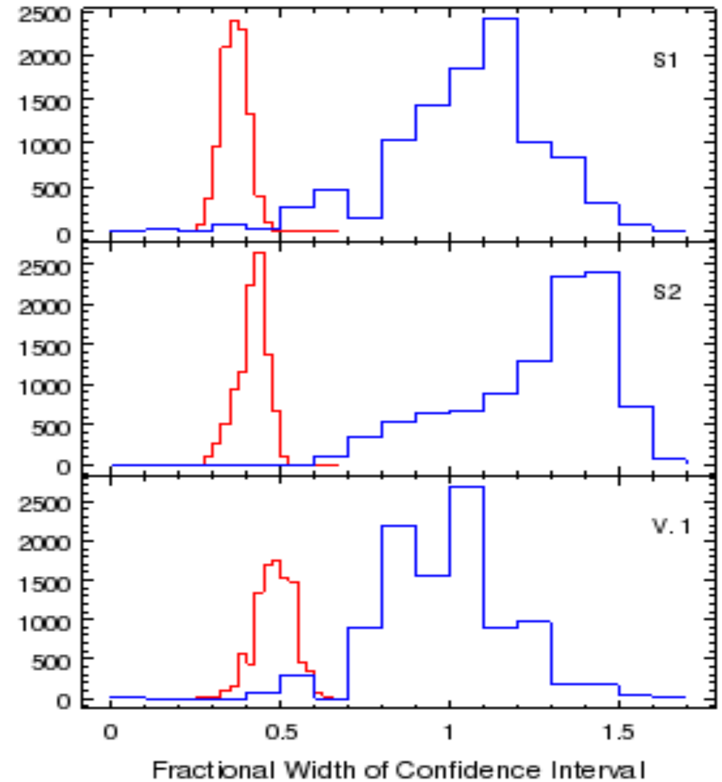




- Simulation of two point sources separated by 0.75″ at $\theta \sim 0.5'$, with true intensities of S1 = 100 and S2 = 30 net counts
- The 90% PSF ECF region is shown for each source
- Counts in the overlap region may be assigned to either source but not both
  - In Rel. 1, counts in the overlap region were excluded from both the source and background apertures (cases S1 ONLY and S2 ONLY)
  - In Rel. 2, sources are analyzed as a bundle and counts in the overlap region are assigned to either S1 or S2 but not both (cases S1 or S2)

# Rel. 2 Aperture Photometry (cont.)

CXC

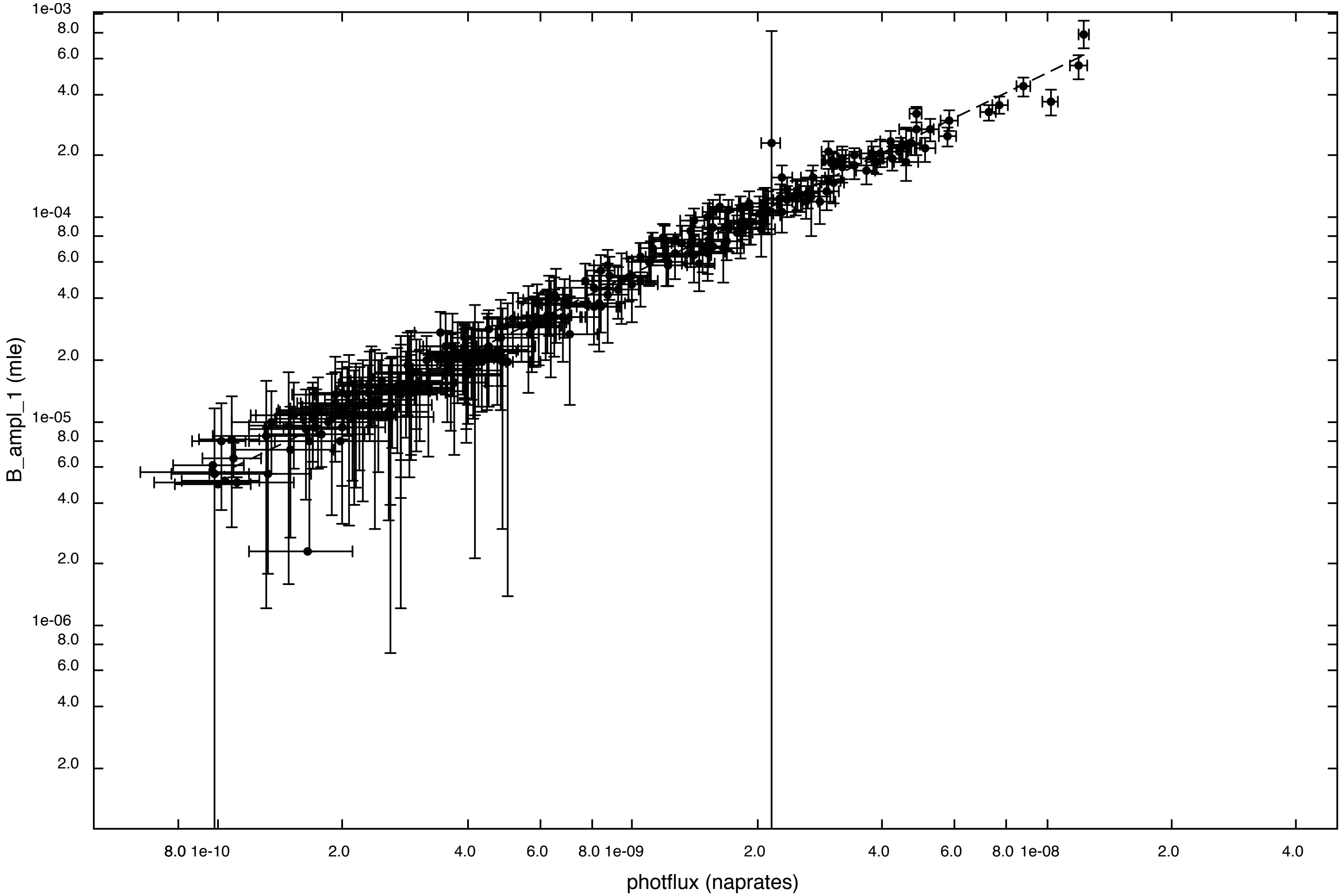



- Distributions of fractional error [i.e., (*mode* − *true value*) / *true value*] and fractional width [i.e., (*upper conf. limit* – *lower conf. limit*) / *true value*] for two point sources separated by $0.75''$ at $\theta \sim 0.5'$ with true intensities of S1 = 100 and S2 = 30 net counts, computed from 10000 simulations each
- Distributions favor assigning counts to aperture of brighter source in this case
- New algorithm recovers true source counts more accurately than Rel. 1 approach for close source bundles

# Rel. 2 Aperture Photometry (cont.)

CXC




- Comparison of Rel. 2 single-ObI source fit amplitudes computed by `mle` with photon fluxes computed by prototype aperture photometry code (`naprates`) for point sources with ‘`b`’ band likelihood > 50 in M31, StackId `acisfJ0042432p411635_001`

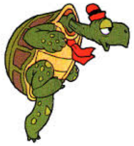
CHANDRA SOURCE CATALOG

# Steps to Catalog Release 2

CXC

## Operations plan

- Take advantage of multi-step architecture to process in distinct phases while subsequent pipeline steps are being developed / tested
  - Each phase needs completed science review and sign-off prior to start
  - Each phase must be completed and verified before the next phase can start
- *Phase 1:* run through *stacker* pipeline
  - Creates all per-ObsId-stack source detections



- *Phase 2:* run *master_match* pipeline
  - Merges source detections from multiple overlapping ObsId-stacks
  - Assigns source names
  - After *master_match* runs, a definitive source list is available

- *Phase 3:* run *source* and *master* pipelines
  - Generates properties for each source detection
  - Merges properties for each unique source on the sky

- *Phase 4:* perform final (human review) QA

# Steps to Catalog Release 2 (cont.)



## Pipelines Already Developed

- *precaldet* pipelines (acis, hrc) — *Operations Phase 1*
  - Performs preliminary calibration and source detection for each observation
  - Pipeline complete
- *fine_astrometry* pipeline — *Operations Phase 1*
  - Computes astrometric corrections to align stacked observations accurately
  - Pipeline complete
- *cal* pipelines (acis, hrc) — *Operations Phase 1*
  - Calibrates observations, removes background flares, computes backgrounds
  - Pipeline complete
- *combodet* pipeline — *Operations Phase 1*
  - Stacks observations, pre-populates limiting sensitivity, performs candidate source detection (`wavdetect`, `mkvtbkg` compact and convex hull sources) in each energy band, filters and merges results
  - Pipeline complete; investigating `mkvtbkg` performance and addressing issues
- *sourcevalidation* pipeline — *Operations Phase 1*
  - Computes convex hull source likelihoods, selects candidate compact source detections to pass to `mle`, creates candidate source bundles
  - Pipeline complete

# Steps to Catalog Release 2 (cont.)



## Pipelines Already Developed (cont.)

- *mle* pipeline — *Operations Phase 1*
  - Creates source region data products (images, exposure maps, backgrounds, etc.), generates local PSF (SAOTrace/MARX), runs `mle` on source bundle
  - Pipeline complete; investigating `mle` performance and addressing issues
- *stacker* pipeline — *Operations Phase 1*
  - Combines output from set of `mle` pipelines (one per bundle) to produce a single merged source list for observation stack
  - Pipeline complete

# Steps to Catalog Release 2 (cont.)



## Pipelines Still To Be Developed

- *master_match* pipeline — *Operations Phase 2*
  - Merges source detections from overlapping stacks
    - Source matching based on existing code from Rel. 1 *master* pipeline UPDATE!
    - New functionality to compute corresponding source regions in stacks where a source is not detected to extract photometric upper limits NEW
- *source* pipeline — *Operations Phase 3*
  - Computes source properties from each observation and from stack
    - Bayesian aperture photometry algorithm (`naprates`) NEW
      - » Science prototype and spec in hand
    - Bayesian Blocks algorithm partitions stacks of observations into groups for combined analysis NEW
      - » Science prototype and spec in hand
    - Hardness ratios computed from aperture photometry PDFs NEW
      - » Straightforward, but spec not yet in hand
    - Spectral model fluxes and spectral fits UPDATE!
      - » Similar to Rel. 1, with additional source model added
    - Temporal variability analysis UPDATE!
      - » Similar to Rel. 1, but based on aperture photometry PDFs
    - Source extent — no longer required (extracted from `mle` fits)

# Steps to Catalog Release 2 (cont.)

## Pipelines Still To Be Developed (cont.)

- *master* pipeline — *Operations Phase 3*
  - Similar to merge processing from Rel. 1 *master* pipeline UPDATE!
  - *Mostly* uses components from the *source* pipeline, with groups redefined

## Quality Assurance

- All pipelines are followed by appropriate automated QA processing
  - Pipelines thru *stacker*: being developed
  - Pipelines post-*stacker*: specs not yet in hand
- Manual QA processing can be invoked after *fine_astrometry*, *combodet*, *stacker*, and *master_match* pipelines as required

# Schedule



## Schedule Impacts

- Catalog pipeline development has been impacted by many issues
  - Competition for limited resources
    - Mission support takes priority (same personnel)
      - (e.g.,) Linux migration; peer review support; archive operations
  - Recent loss of key science/software staff will impact schedule going forward
  - Numerous unexpected issues with background determination, source detection, and source fitting that require investigation and resolution
    - In some cases they are due to bugs, but in many cases they are caused by undocumented limitations in existing software (both internal and external)
    - Bug-fixes have been/will be included in CIAO 4.5 or CIAO 4.6
    - Issues and workarounds have been/will be documented on CIAO website where appropriate
  - Science algorithm development takes longer than expected
    - Some algorithms require fundamental research and extensive prototyping
    - Supporting current catalog pipeline development and testing often takes priority
  - Catalog processing hardware performance is less than expected
    - Increases turnaround time for testing
    - Will impact production time

# Catalog Software Releases



## Catalog-Related Releases

- Software release associated with CSC Rel. 1

| Release | Date | Content |
|---|---|---|
| CAT 3.2.7.2 | 5 Mar | CSCview updates for Java 7; packaged as OS X app |
| CAT 3.2.7.2 | 19 Jun | VO interfaces update; CSCview package update |

- Software release associated with CSC Rel. 2

| Release | Date | Content |
|---|---|---|
| CAT 4.1.1 | 13 Sep | Refinements to PLs (add/validate HRC); archive ingest thru *stacker*; pass basic test — list of 20 observation stacks |
| CAT 4.1.2 | ~End Nov | Additions to PLs (tune algorithm params, crater detection, add ultrasoft band); 2nd tier `mle` updates; data product refinements |
| CAT 4.2 | Spring 2014 | QA thru *stacker*; reprocessing support; interleave-mode; performance tuning; pass big test — fraction of archive<br>**Production run through *stacker* pipeline** |
| CAT 4.3 | Fall 2014 | *master_match* and *source* PLs; limiting sensitivity population<br>**Production run *master_match* and *source* pipelines** |
| CAT 5.0 | Winter 2014 | *master* PL; populate databases<br>**Production run *master* pipeline and final QA review** |