Statistical Characterization of the Second Chandra Source Catalog

Francis A. Primini[1], Ian N. Evans[1], Joseph B. Miller[1], Janet D. Evans[1], Christopher E. Allen[2], Craig S. Anderson[1], Glenn Becker[1], Jamie A. Budynkiewicz[1], Douglas Burke[1], Judy C. Chen[1], Francesca Civano[1], Raffaele D'Abrusco[1], Stephen M. Doe[2], Giuseppina Fabbiano[1], J. Rafael Martinez Galarza[1], Danny G. Gibbs II[1], Kenny J. Glotfelty[1], Dale E. Graessle[1], John D. Grier Jr.[1], Roger M. Hain[1], Diane M. Hall[3], Peter N. Harbo[1], John C. Houck[1], Jennifer L. Lauer[1], Omar Laurino[1], Nicholas P. Lee[1,] Michael L. McCollough[1], Jonathan C. McDowell[1], Warren McLaughlin[1], Douglas L. Morgan[1], Amy E. Mossman[1], Dan T. Nguyen[1], Joy S. Nichols[1], Michael A. Nowak[4], Charles Paxson[1], Menelaos Perdikeas[1], David A. Plummer[1], Arnold H. Rots[1], Aneta L. Siemiginowska[1], Beth A. Sundheim[1], Sinh Thong[1], Michael S. Tibbetts[1], David W. Van Stone[1], Sherry L. Winkelman[1], and Panagoula Zografou[1]

[1]Center for Astrophysics | Harvard & Smithsonian [2]Formerly Center for Astrophysics | Harvard & Smithsonian [3]Northrop Grumman Mission Systems [4]Formerly MIT Kavli Institute for Astrophysics and Space Research

The second major release of the Chandra Source Catalog, CSC 2.0, offers significant improvements over the previous catalog release, CSC 1.1, both in the amount of data included and the analysis procedures followed. CSC 2.0 includes ~315,000 unique compact X-ray sources and ~1000 highly extended sources, and covers ~550 square degrees of the sky. The sensitivity limit for compact sources has been significantly improved to ~5 net counts on-axis. Both the additional data and the improved analysis techniques mandate a full re-characterization of the statistical properties of the catalog, namely, detection efficiency, false source rate, sensitivity, and accuracy of source properties. We present a preliminary summary of that work here. As in CSC 1.1, we use both analysis of real CSC 2.0 catalog results and extensive simulations of blank-sky and point source populations to characterize the catalog. For more detail, please visit https://cxc.cfa.harvard.edu/csc/char.html .

SOURCE DETECTION

Source detection in *CSC 2.0* is a two-step process. After observations have been co-added into stacks, the combined image data are analyzed with two separate source detection tools—the CIAO tool wavdetect and a Voronoi Tesselation based detection tool, mkvtbkg, developed by the CSC team for detecting large extended sources and point sources embedded in diffuse emission. Both tools are run with very low detection thresholds to maximize the number of real sources detected. A point source model is fit to combined image data for all source candidates, and candidates are classified as FALSE, MARGINAL, or TRUE, depending on where their detection likelihoods fall with respect to two likelihood thresholds, corresponding to false source rates of ~1 (FALSE-MARGINAL boundary) and ~0.1 (MARGINAL-TRUE boundary) false sources per stack, respectively.

Thresholds are determined using simulations in which the event lists for actual catalog observations are replaced with blank-sky event lists derived from the background map for the corresponding observation, randomized with Poisson noise. Typically, ~100-200 runs of the same simulation set were generated. The table below summarizes simulation sets used to derive likelihood thresholds.

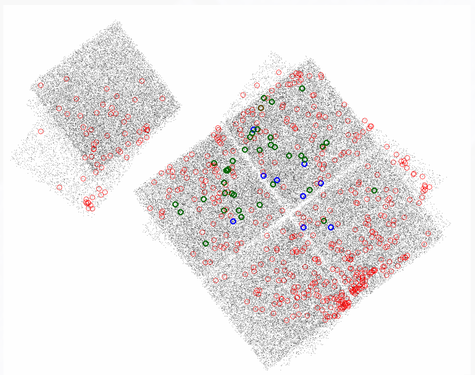
Aimpoint	Obsids	T_{stack} (ksec)	Marginal Source Detections		True Source Detections	
			Detections(Runs)	FSR	Detections(Runs)	FSR
ACIS-I	15164	9	40(225)	0.18	3(225)	0.01
ACIS-I	14024	135	59(194)	0.30	1(194)	0.01
ACIS-I	3251,10413,10786,10797	135	82(153)	0.54	25(153)	0.16
ACIS-I	14022,14023	296	64(158)	0.41	8(158)	0.05
ACIS-S	7921	135	100(199)	0.50	33(199)	0.17
ACIS-S	11688,11689,12106,12119	288	223(178)	1.25	33(178)	0.19
			60(178) (no chip8)	0.34	20(178) (no chip8)	0.11

FALSE SOURCE RATE

We can demonstrate the performance of the likelihood threshold functions by computing the actual false source rates in the various simulation sets. An example simulated event list from the four-ObsID ACIS-I set is shown here. We find 82 detections with likelihoods above the FALSE-MARGINAL threshold, yielding an average false source rate (FSR) of 0.54+/-0.06 per field for MARGINAL sources. Similarly, we find 25 detections above the MARGINAL-TRUE threshold, for an FSR of 0.16+/-0.03 per field for TRUE sources.

In general, FSR are consistent with the desired rates of 1 per field for MARGINAL sources and 0.1 per field for TRUE sources, with the exception of the ACIS-S aimpoint four-ObsID set. We note that, as in CSC 1.1 there is an excess of detections in the vicinity of bad columns in Chip 8. Excluding these detections, the FSR for this set agree with those in the other sets.

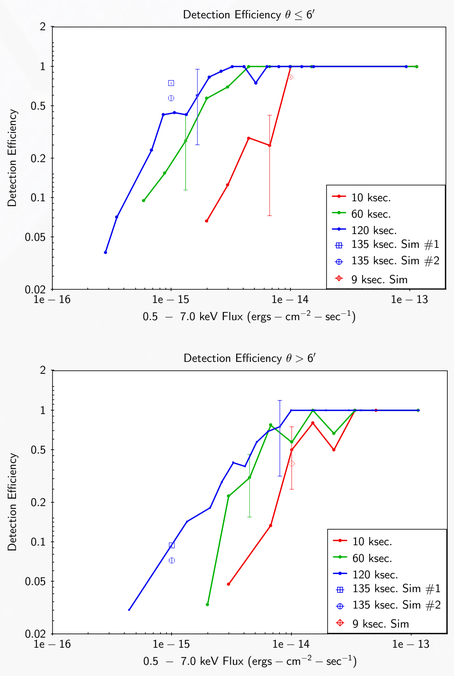
Simulated event list for the 4-ObsID ACIS-I stack. FALSE (red), MARGINAL (green) and TRUE (blue) sources for all runs of this set are shown.



DETECTION EFFICIENCY

We estimate the detection efficiency by comparing the number of detections in individual observations that are part of the Chandra Deep Field South Survey to the number of sources reported in that survey's 7 Msec catalog (Luo et al., 2017 ApJS 228 2). The CDFS sources are assumed complete at the exposures of individual observations. We consider ObsIDs 12047, 12054 and 17535, with exposures of ~10, ~60, and ~120 ksec, respectively, and calculate the ratio of the number of CDFS sources detected in *CSC 2.0* as MARGINAL or TRUE to the total number of CDFS sources, as a function of CDFS flux, for two ranges of off-axis angle.

We also use a number of faint point source simulations with exposures comparable to those in ObsIDs 12047 and 17535. These are based on the simulation sets described in the Table above, but with point sources with a few counts injected into the simulated event lists. Fluxes for these sources are assumed to be similar to those of *CSC 2.0* sources with comparable counts observed in ObsIDs 12047 and 17535. Detection efficiencies from these simulations are in good agreement with those from the CDFS analysis.



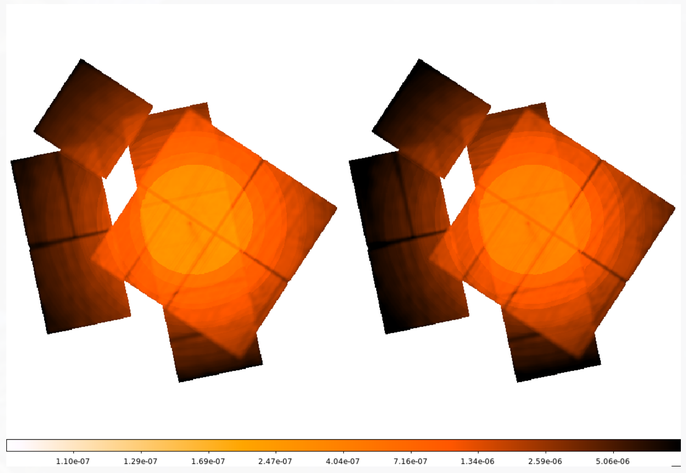
LIMITING SENSITIVITY AND SKY COVERAGE

SENSITIVITY MAPS

Limiting sensitivity maps are computed for each stack, in all source detection energy bands, for both FALSE-MARGINAL and MARGINAL-TRUE detection thresholds. The maps are based on stack-level background maps. Each map value represents the point source photon flux, p_{min}, corresponding to net counts n in a 90% ECF aperture at that location, with background B, such that

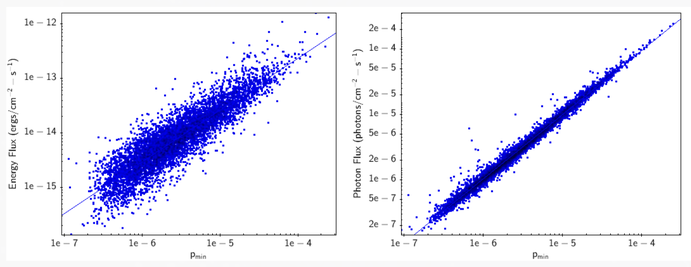
$$P_{poisson}\,(Counts\ in\ aperture \geq n + B \mid B) < P_{threshold}\ ,$$

where $P_{threshold}$ is a threshold probability corresponding to the source detection likelihood threshold at that location. B-band maps for stack acisfJ1509253m585033_001, for FALSE-MARGINAL (left) and MARGINAL-TRUE (right) thresholds are shown at right.



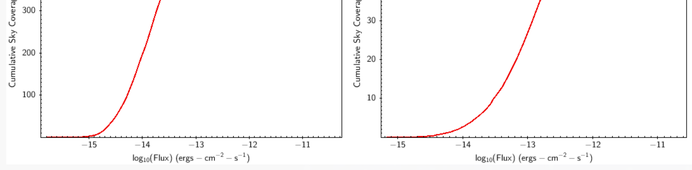
APERTURE FLUXES vs. p_{min}

we selected a sample of isolated CSC Release 2.0 point sources and calculated p_{min} from the available aperture quantities, using the actual detection likelihoods. We then compared these to actual photon fluxes and energy fluxes, from the corresponding photflux_aper90 or flux_aper90 values. Results for the b-band are shown at right, and indicate good agreement between actual source fluxes and sensitivity values for the same detect thresholds. The energy flux relations are used to compute limiting sensitivity in energy flux units.



SKY COVERAGE

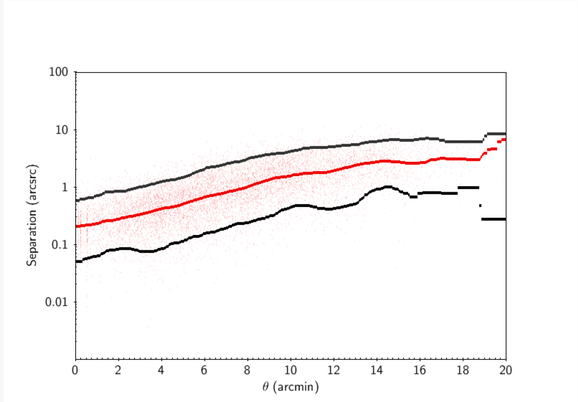
In addition to stack-level sensitivity maps, all-sky maps of limiting sensitivity are constructed by regridding individual maps in a HEALPix nested celestial grid with index = 16 (θpix≈3.22″). All populated HEALPix pixels are collected in the catalog database. For HEALPix pixels that occur in multiple stacks, the highest sensitivity value (i.e., lowest sensitivity value) is used. Users may then query the database for limiting sensitivity values near positions of interest. All-sky maps are generated for s, m, h, b, and w bands, for both FALSE-MARGINAL and MARGINAL-TRUE detection thresholds. The total cumulative sky coverage at the MARGINAL-TRUE threshold is ~520 deg^2 for the b-band and ~55 deg^2 for the w-band, and is shown as a function of energy flux at right.



ASTROMETRY

SEPARATION BETWEEN CSC2.0 AND SDSS COUNTERPARTS

To characterize the astrometric accuracy of *CSC 2.0*, we cross-match *CSC 2.0*, Master Source positions with positions of stars in the SDSS-DR13 catalog, using a technique similar to that used in *CSC 1.1*, (Rots & Budavári, 2011 ApJS 192 8). By considering only *CSC 2.0* sources which derive from a single observation, we can investigate the dependence of astrometric accuracy on off-axis angle, θ. A plot of separation offset δ vs. θ for a sub-sample of ~9000 single-observation matches is shown in at right. The mean offset is ~0.32″ for sources with θ < 3′, ~0.83″ for sources with θ < 10′, and ~1.2″ overall. We note these values are slightly larger than the corresponding values for *CSC 1.1*.



Angular separation between CSC 2.0 sources and SDSS stars for ~9000 single-observation matches, as a function of off-axis angle θ. Solid black lines indicate 5% and 95% quantiles, and the solid red line indicates the median.

This work has been supported by NASA under contract NAS 8-03060 to the Smithsonian Astrophysical Observatory for operation of the Chandra X-ray Center.