Chandra Automated Point Source Processing

Bradley D. Spitzbart and Scott J. Wolk

Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA, USA

ABSTRACT

We have implemented a system to automatically analyze *Chandra* x-ray observations of point sources for use in monitoring telescope parameters such as point spread function, spectral resolution, and pointing accuracy, as well as for use in scientific studies. The *Chandra* archive currently contains at least 50 observations of star cluster-like objects, yielding 5,000+ sources of all spectral types well-suited for cataloging. The system incorporates off-the-shelf tools to perform the steps from source detection to temporal and spectral analyses. Our software contribution comes from wrapper scripts to autonomously run each step in turn, verify intermediate results, apply any logic required to set parameters, decide best-fit results, merge in data from other catalogs and to format convenient text and web-based output. We will outline this processing pipeline design and challenges, discuss the scientific applications, and focus on its role in monitoring on-orbit observatory performance.

Keywords: Chandra, point sources, automation, trending

1. INTRODUCTION

The *Chandra X-ray Observatory* was launched in July 1999 and has performed reliably since, while returning exquisite pictures and detailed data of objects from the Earth to distance galaxies. The *Chandra* mirrors in combination with the two on-board focal plane instruments, the Advanced CCD Imaging Spectrometer (ACIS) and the High Resolution Camera (HRC), yield on-axis spatial resolutions of 0.5 arcsec. Therefore, many previously unresolved sources are now seen.

Typical nuances and judgement calls in data analysis lead to many difficulties in comparing published data from different observers or from different observation dates. The data reduction and analysis software itself changes over time which can affect the final results even on the same dataset. For example, standard tools are now available to compensate for CTI degradation and ACIS filter contamination. Some observers may choose to apply these corrections, others may not based on their specific needs. In most cases different spectral models and classification criteria will be applied based on the observer's preferences and familiarity. Even with the same models, initial parameter settings and statistics, often assigned subjectively, will lead to different final fit values. The goal of our catalog is to provide a uniform (not necessarily optimal) database for the comparison of data from different stellar clusters. Secondarily, the catalog must be easily available and usable, such as via the World Wide Web. This type of catalog provides value-added science return as well as convenient observatory health and performance metrics.

Several similar projects are also available spanning different object classes and wavelengths, all of which we have drawn on for ideas. See for example the Chandra Supernova Remnant Catalog [*] and the WEBDA [†] project on galactic open clusters.

Below, we will describe and give examples of the catalog content in Sect. 2. Section 3 covers how the data is analyzed and compiled. The next two sections discuss the observatory performance and scientific applications, respectively, and Sect. 6 proposes future enhancements.

Send correspondence to B.D.S.; E-mail: bspitzbart@cfa.harvard.edu

[*] http://snrcat.cfa.harvard.edu/

[†] http://obswww.unige.ch/webda/webda.html

2. CATALOG CONTENTS

The point source catalog web archive is located at

http://hea-www.harvard.edu/~swolk/ANCHORS

It is introduced with a home page containing background information, links to help files, and a list of star cluster fields observed by *Chandra*. For each field, a few key parameters including right ascension (RA), declination (dec), target name, exposure time, and date are listed. For the fields with analysis available, links are activated to the individual source characteristics pages. Users may click the target name to continue.

2.1. Source Characteristics

The web display has a cover page with information about the star cluster such as distance and multi-wavelength images with links to the detailed analysis for each source. Also listed will be parameters for the *Chandra* observation. Each source page has multi-wavelength images and data, spectral fit parameters and plots, and lightcurve statistics and plots. For our database, final x-ray source properties including position, net count rates, flux, and hardness ratios in several energy bands are output. Also lightcurve, histogram, and smoothed image plots are created. Spectral fits and parameters are shown for one and two three temperature APEC models and other models as necessary.

2.2. Utilities

Many interactive features are built into the system. The main target table and source tables are easily sorted by the user choosing a column to sort on. Compound searches can be made within a single observation or across all observations to find, for example, all sources with derived temperatures between 1 and 2 keV. Following manipulation, tables can be written out and downloaded in a tab-delimited text format. In addition, the final web interface will allow users to produce fits using other models or initial parameter values. Through a CGI, users may select a model and initial parameter files, and apply any desired filters to the data. The fit is run, which generally takes only a few seconds. Returned will be a plot of the fit and a listing of best-fit parameters and statistics. Like all other results, the plot, data, and listings will be available for download.

3. TECHNICAL METHODS

To accomplish automated analysis we employ several outside software packages. See Appendix A for links to the tools mentioned below.

3.1. Data Preparation

Exposure maps and aspect histograms are required by several steps in the analysis pipeline to accurately compensate for the mirror effective area and telescope dithering effects at different locations across the detector. There are tools called *merge_all* and *asphist* included with CIAO installations to perform these tasks automatically. We generally use the default energy of 1.49 keV (the HRMA/ACIS effective area peak) and settle for a monotone exposure map.

Each step in the automated analysis pipeline is keyed off a list of source identifications and positions. For the prototype analysis runs, we have simply used the standard level 2 source lists provided by the *Chandra X-ray Center* (CXC) pipeline. These are created by celldetect with the default settings, ie. no exposure map applied, no background subtraction, and signal to noise threshold for source detection set to 3. A better approach is to regenerate a source list using PWDetect instead. We choose this algorithm because we can statistically evaluate false detection probabilities and it performs well near chip edges. In the end, all that is needed is a good list of source positions. The actual extraction regions for each source is not important as acis_extract will define its own source and background regions.

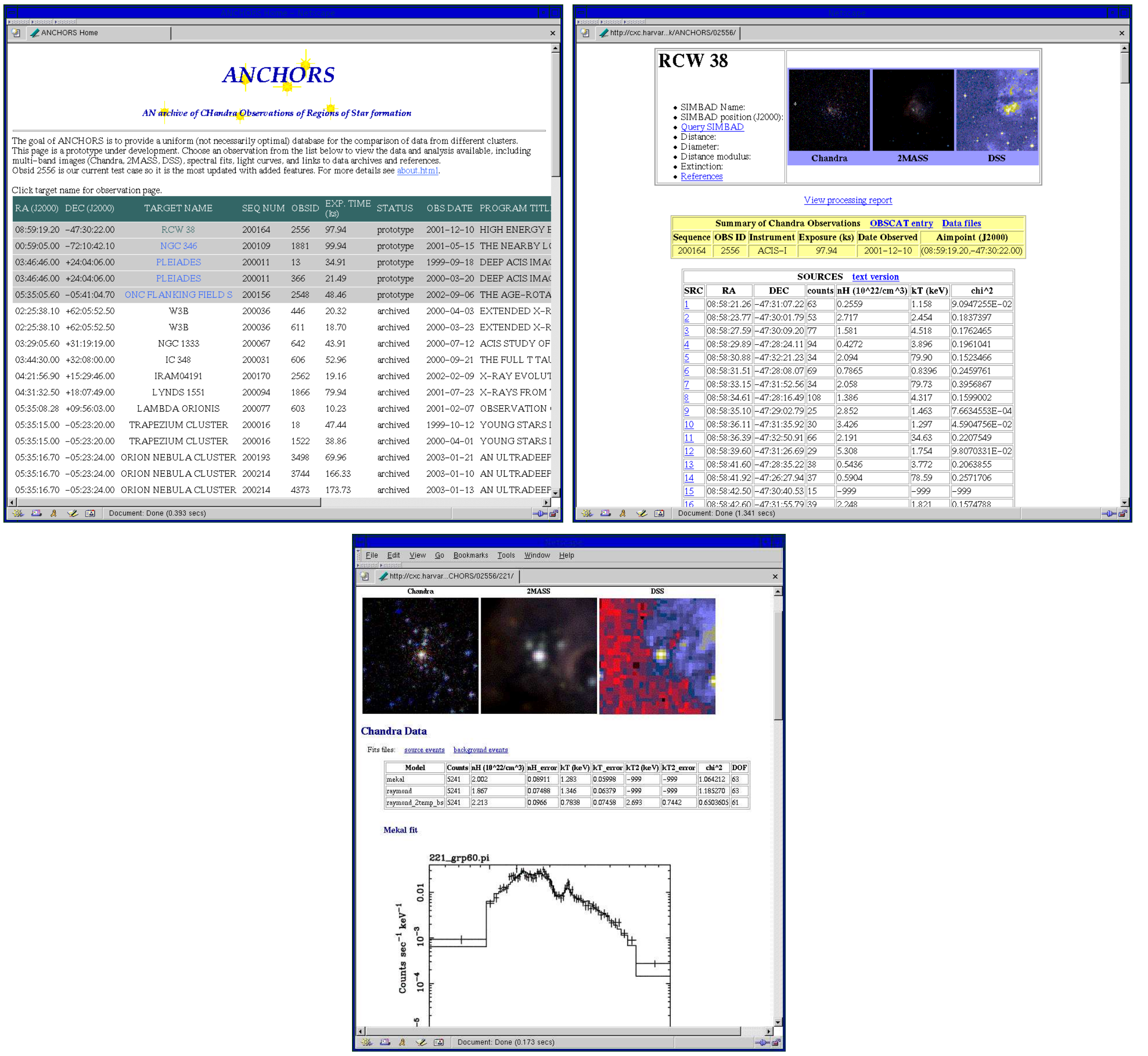

Figure 1. The catalog's three main page types. Links and information about all available observations are listed on the top page (upper left). Each observation has its own sub-page (upper right). And each source has its own page, showing multi-band images and data, spectral fit parameters and plots, and light curve plots and statistics (bottom).

3.2. Acis_extract

Acis_extract (AE) (Broos 2004) is an IDL package which picks up at the end of CXC level 2 event processing. AE does a fairly robust job of choosing appropriate source and background regions, given the centroid position, based on the *Chandra* PSF library and encircled energy parameter (default 90%). AE also creates helpful plots summarizing the analysis process. For example, extraction region area vs. off-axis angle and our own supplements of source region images can be provided to verify and sanity check the automated analysis. Some human intervention is required to confirm the regions or make adjustments. While AE is a very accomplished tool, we are working on modifications to allow it to run Sherpa in addition to XSPEC to take advantage of the available APEC model and improved statistics. The APEC models yield acceptable fits for the widest range of spectral profiles. Based on the reduced χ^2, we discriminate automatically those sources (with enough counts) that should be refit with a different initial guess.

As AE was developed for the Orion Ultradeep Project (E. Feigelson PI) it is well tested on multiple observations of a single field. The output of AE is a series of directories, 1 per point source which contains a set of spectral model fit parameters and light curves (with a variety of variability tests) for each source, and a wide variety of statistical properties of the sources.

In addition, we extend the output from AE with our own code and by plugging in packages developed by others. Some procedures are rewritten to generate GIF or PNG images instead of PostScript for easier web display. To expand the analysis, a promising quantile approach (Hong, 2004) is used to augment spectral fitting and traditional hardness ratios. Instead of working with predetermined energy bands, quantile analysis determines the energy values that divide the detected photons into fractions of the total counts. This is especially useful for faint sources where low statistics preclude meaningful spectral fitting. We also find that the quantile analysis often shows higher fidelity than hardness ratios. Another expansion is the calculation of Bayesian blocks (Scargle, 1998) for light curve analysis. This technique yields a segmentation of time intervals during which the photon arrival rate is statistically constant. Instead of using predetermined time bins, the data itself determines periods of constant representation. The resulting locations, amplitudes, and rise and decay times provide another metric for identifying flaring and variable candidates.

3.3. Data compilation/conversion

The AE output data files in FITS and ASCII are compiled and formatted in HTML by Perl scripts. The tasks are divided into five scripts. 1) the target homepage is created, 2) the homepage is made for each source, 3) 2MASS and DSS images are collected, 4) acis_extract postscript plots are converted to gif format (many of these will be replaced by our own IDL renderings), and 5) *Chandra* images are produced in energy-space color.

Available information from other sources is included when possible. Some of this is generated automatically. For example, a link to query SIMBAD for basic data and name cross-references can be easily created given any source position. Other facts must be gathered manually. Here, a query of the Astronomical Data System (ADS) and scan of the returned articles can turn up mass, distance, and multi-band magnitude values.

4. OBSERVATORY PERFORMANCE APPLICATIONS

The ultimate measure of observatory performance is in the quality of the science data. The catalog will make it possible to treat science quantities in similar ways to how spacecraft temperatures and voltages are treated for monitoring and trending. Using PSF and spectral line measurements from all the best point sources (greater than 200 counts) gives a much better baseline than could be available from a necessarily limited calibration campaign. Using the observatory's standard data products and tools we can best access the final data quality that reaches investigators. This assessment takes into account instrumental, aspect reconstruction, and ground software factors.

The *Chandra* High Resolution Mirror Assembly (HRMA) consists of four pairs of concentric grazing-incidence Wolter Type-I mirrors of polished Zerodur glass coated with iridium with a binding layer of chromium in between. The front mirror of each pair is a paraboloid and the back a hyperboloid. The diameter of the innermost pair is about 0.65 meters, while the outermost pair has a diameter of 1.23 meters. Each shell contributes proportionally to the total HRMA effective area as a function of energy. The inner shells are most effective at reflecting high

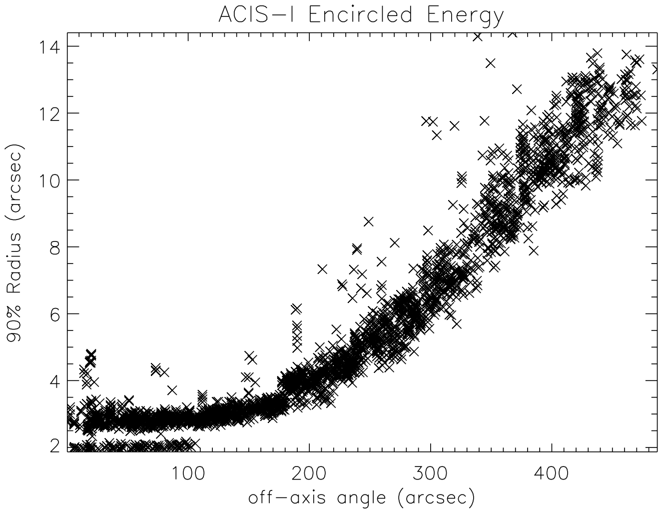

Figure 2. Encircled energy vs. off-axis angle for ACIS-I sources combined from throughout the mission. This metric defines focus and alignment parameters.

energy X-rays, while the outer shells contribute more at lower energies (Schwartz, 2000). The total effective area is about 800 cm^2 at 0.25 keV and about 100 cm^2 at 8.0 keV with a focal length of about 10 m (Chandra Proposers' User Guide, 2003).

Figure 2 shows the 90% encircled energy radius (EE) at 1.49 keV vs. off-axis angle (θ) from the the nominal aimpoint for sources combined from throughout the *Chandra* mission. The ACIS-I CCD is a 4-chip array, arranged 2 X 2, for a total field-of-view of 16.9' X 16.9'. Highly piled-up (bright) sources are apparent with EE radii less than 2.3 arcsec. The curve fits a polynominal of the form, $EE = 8e^{-5}\theta^2 - 0.01\theta + 3.0$. By filtering on time or energy we can detect shifts in this fit which indicate any mirror alignment problems.

The Orion Nebula Cluster was observed for 1 million seconds in January 2003 and provides a good test of observatory performance applications. Figure 3 shows the combined image with over 900 detected sources. This image gives an excellent view of the telescope focal plane. Note the elongation and rotation of sources far off-axis due to the cylindrical shape of *Chandra's* mirrors. The elongation and rotation as a function of chip position is an important quantity to measure and track over time, as any shifts must be explained. Our catalog will enable a "mega image" of all point source photons that have fallen on the focal plane.

ACIS offers the capability to simultaneously acquire high resolution images and moderate resolution spectra. Therefore, we have intrinsic energy information for each photon collected. As shown in Sect. 2, spectral fits, hardness ratios, and spectral quantiles are included in the AE runs. We will build up a list of absorption and emission lines for comparison against time and detector location and not be limited to the Al-Kα, Mn-Kα, and Ti-Kα energies tracked now from the external calibration source. Due to the widely reported degradation in charge transfer inefficiency (CTI) early in the mission, we expect to see reduced resolution and greater centroid offsets over time and at higher row numbers on the detector. Spectral characteristics can be used in many other ways. For example, we will monitor for a systematic appearance of any new absorption features indicating contamination on the mirror surfaces.

Thermal control has become a topic of concern on-orbit. Due to radiation interactions the spacecraft's silver teflon insulation has darkened at a faster rate than expected, leading to higher temperatures. Although mirror and optical bench components can peak several degrees above established norms, no image distortions are reported to date. The point source catalog will confirm this result and provide a method of continued monitoring. In addition, because we use standard data products, distortions introduced by the aspect reconstruction system will be readily apparent. In the past year, *Chandra* operations has swapped gyro units and lowered the optical

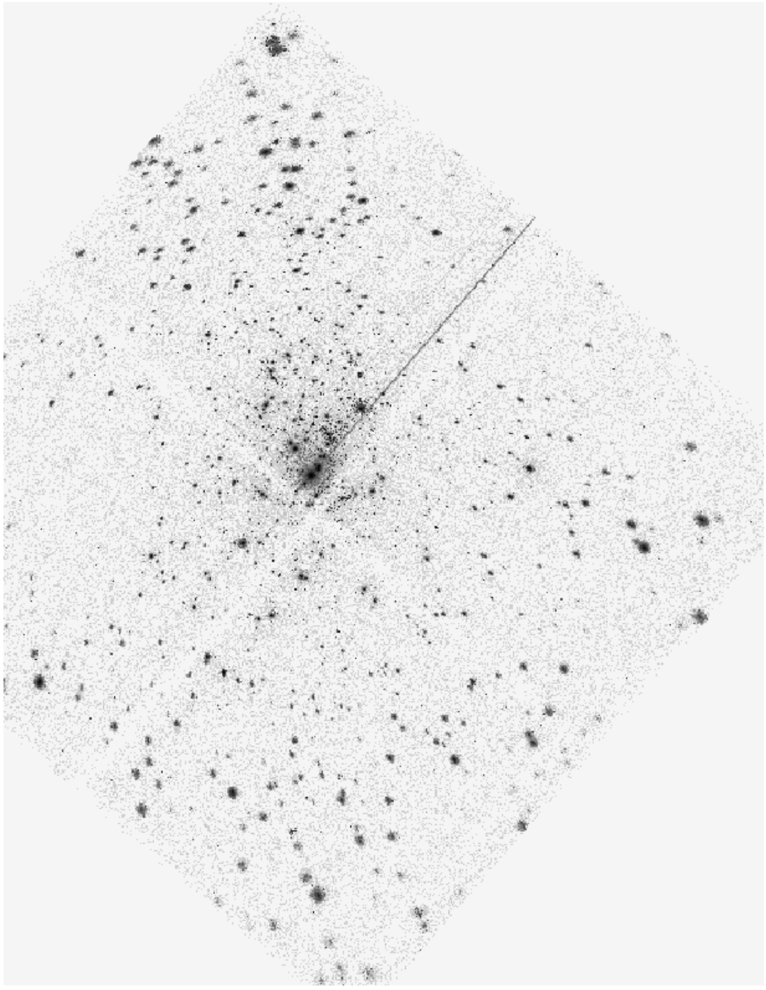

Figure 3. A one million second Orion Nebula Cluster observation. We measure and track the elongation and rotation of sources as a function of off-axis angle and chip position. The streak is an artifact of photons from the bright central source striking the detector while the CCD was being read out.

aspect camera CCD temperature to mitigate the effects of increased current draw and more bad pixels, respectively. No change in image quality is observed, but monitoring now becomes more important to know when more shifts in either hardware or software are needed to maintain *Chandra's* exquisite performance.

5. SCIENCE APPLICATIONS

The archive is designed to aid both the X-ray astronomer with a desire to compare X-ray datasets and the star formation astronomer wishing to compare stars across the spectrum. It brings together *Chandra* data on open clusters and other young stars.

From the earliest *Einstein* observations, it has been clear that young stars are bright, time–variable X-ray sources (Feigelson & Decampli 1981, Montmerle et al.1983). The source of the X-ray emission has been assigned to various physical mechanisms depending on the mass of the stars. These include (from high to low mass) wind-wind/ISM interactions, emission of an unseen companion, an $\alpha - \Omega$ dynamo driven by the interaction of the stellar core with the convective envelope and rotation or an $\alpha - \alpha$ or turbulent dynamo capable of producing emission without core–envelope shearing.

Though only a small fraction of the total luminosity of these stars, the X-ray flux is the main observable difference between young stars without disks and field stars. The high energy emission of the high mass stars not only leads to local shocks, but may also excite diffuse emission (Wolk et al.2002, Townsley et al.2003, Yusef-Zadeh et al.2001). Finally, X-rays induced ionization and melting undoubtably have significant effects within protoplanetary disks (Feigelson et al.2002).

Chandra has proven to be the perfect vehicle for the study of regions of star formation. Its superb spatial resolution allows *Chandra* to resolve stars in crowded regions 2-3 kpc away. With good sensitivity between 2 and 8 keV, *Chandra* can penetrate star forming clouds to levels rivaling near-IR telescopes. These features allow *Chandra* to investigate star formation which is more massive, more embedded and more distant than previously possible.

While much can be learned about stellar evolution from the study of individual open clusters, science return is enhanced when the clusters are viewed as a group. As a pilot study, we examined brown dwarfs observed by *Chandra* during AO1-2. We found almost 70 candidate brown dwarfs had been detected by *Chandra* (Wolk 2003b) (though only 8 bone-fide). Trends indicate that the younger brown dwarfs are hotter in X-rays than the field brown dwarfs, but the total X-ray luminosity of detected brown dwarfs are similar. Another study of a subset of objects was done by Feigelson et al. (2002). They examined only the 43 X-ray sources in the ONC between 0.7 $M_\odot$ and 1.4 $M_\odot$ in order to understand the mean properties of the young Sun at 0.5 Gyr. They conclude that the flares which occur during the protoplanetary phase can cause significant production of unusual nuclieds including ^{26}Al. Using the point source database, one could follow the progression of luminosity and variability for sun-like stars from the birthline to the present day (with the inclusion of AO-4 target NGC 752) without having to weigh the impact of the different analysis assumptions made by each team. Similar studies can be performed on intermediate mass stars. These studies are particularly interesting since flares imply the presence of confined plasmas which should be absent in these stars.

6. FUTURE ENHANCEMENTS

The *Chandra* point source catalog presented here is a prototype of the final version. Due to the long-term nature of the coverage it will always be evolving. Many applications we have not yet envisioned will hopefully emerge as the database is populated and with input from users. We are currently concentrating on processing representative observations through the acis_extract and data compilation stages. Most of the analysis tools to automate the observatory performance applications do not yet exist, but will be forthcoming. We know a more sophisticated web interface design is desired to make navigation easier. Finally, the catalog should be made to be useful and available in the emerging Virtual Observatory paradigm. Specifications for file formats and protocols are still being defined to open up the vast stores of data for optimized science return.

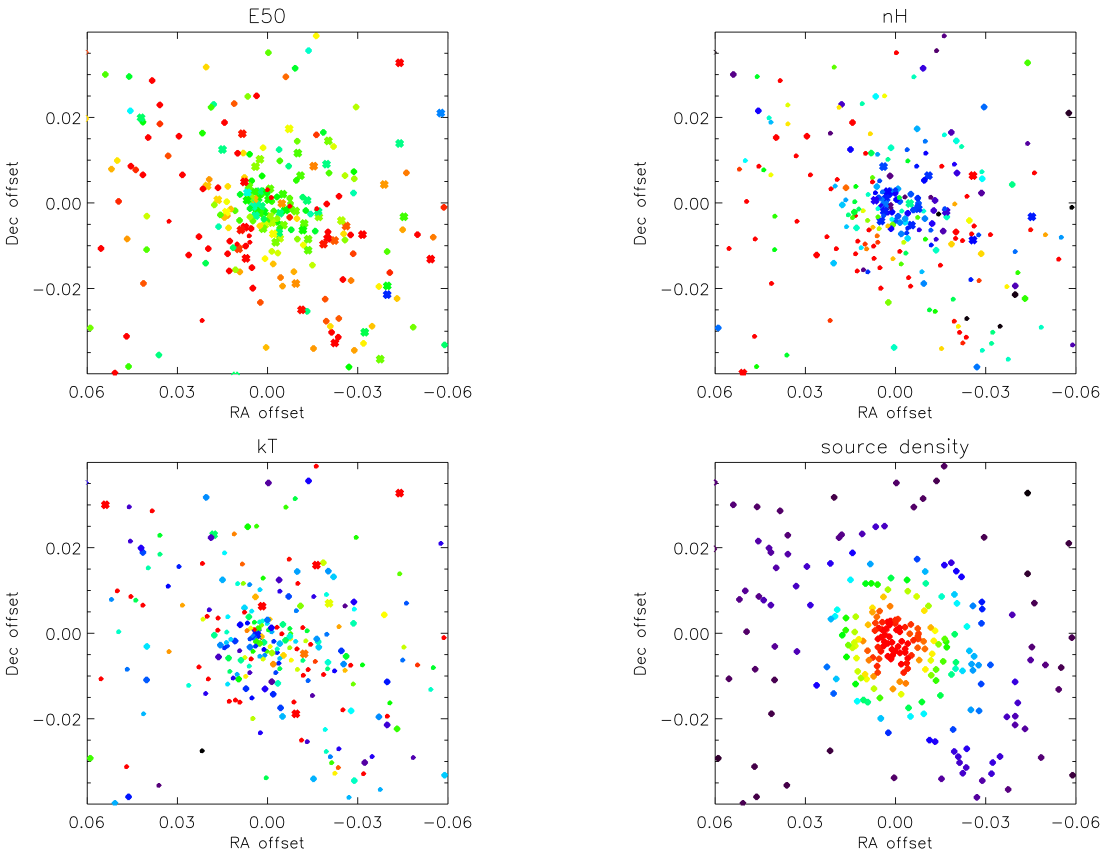

Figure 4. Examples of color-coded maps easily created from the point source catalog. Clockwise from upper left, scaling based on 50% quantile, absorption column (nH), source density (N/min^2), and temperature (kT).

APPENDIX A. LINKS TO SOFTWARE

The following tools are used directly or indirectly in the creation or presentation of the catalog:

CHaSeR (http://cda.harvard.edu/chaser/mainEntry.do)
Chandra Search and Retrieve is used to find available archival data. There is also a java-based version included in CIAO distributions which can return an ASCII list of observations satisfying search parameters.

CIAO (http://cxc.harvard.edu/ciao/)
CIAO is used by the standard Chandra data processing pipeline. It is used directly by the catalog prep pipeline to copy and filter event files and to make exposure maps and aspect histograms.

PWDetect (http://www.astropa.unipa.it/progetti_ricerca/PWDetect/)
PWDetect is used for source detection in place of celldetect. Produces list of source positions, count rates, significance.

Acis_extract (http://www.astro.psu.edu/xray/docs/TARA/ae_users_guide.html)
Acis_extract is the main analysis engine for the catalog, from Patrick Broos et.al. at Penn State University.

XSPEC (http://heasarc.gsfc.nasa.gov/docs/xanadu/xspec/)
XSPEC is the spectral fitting tool called by acis_extract.

Sherpa (http://cxc.harvard.edu/sherpa/index.html)
Sherpa is the CXC-based spectral fitting tool. We will use this tool in place of XSPEC to run APEC [‡] modelling.

Yaxx (taldcroft@cfa.harvard.edu)
Yaxx is a Perl package by Tom Aldcroft to automate Sherpa fitting.

Bayesian Blocks (http://space.mit.edu/CXC/analysis/SITAR/)
We run a suite of S-Lang [§] scripts to create Bayesian block light curve analysis.

Spectral Quantiles (http://hea-www.harvard.edu/ChaMPlane/quantile/)
Quantiles are indicated as another spectral metric, especially in low-count sources. Created with IDL and Perl code by Jaesub Hong.

SkyView (http://skyview.gsfc.nasa.gov/)
SkyView provides images from many telescopes across all wavelengths. Batch execution is available for use in scripts.

ADS (http://adswww.harvard.edu/)
NASA's Astronomical Data System is the bibliography of published Astronomy and Astrophysics, Instrumentation, Physics and Geophysics.

SIMBAD (http://simbad.u-strasbg.fr/Simbad)
SIMBAD provides basic data, cross-identifications and bibliography for astronomical objects outside the solar system.

[‡] http://cxc.harvard.edu/atomdb/sources_apec.html
[§] http://www.s-lang.org/

ACKNOWLEDGMENTS

This work is supported by the *Chandra X-ray Center* and NASA contract NAS8-39073.

REFERENCES

Broos, P., et al. 2003, AAS 35.
Feigelson, E.D. & Decampli, W.M. 1981, ApJ 243, 89.
Feigelson, E.D., et al. 2002, ApJ 572, 335.
Hong, J., et al. 2004, ApJ, *in press*.
Scargle, J. 1998, ApJ, 504, 405.
Schwartz, D., et al. 2000, SPIE, 4012, 28-40.
Townsley, L., et al. 2003, ApJ, 593, 874.
Wolk, S.J., et al. 2002, ApJ 580, L161.
Wolk, S.J. 2003, *Brown Dwarfs* E. Martin ed. p447.
Yusef-Zadeh, F., et al. 2002, ApJ 570, 665.
Chandra X-ray Center, "Proposers' Observatory Guide." 2003.