The Chandra Source Catalog: Automated Source Correlation

R. Hain[1], I. N. Evans[1], J. D. Evans[1], K. J. Glotfelty[1],
C. S. Anderson[1], N. R. Bonaventura[1], J. C. Chen[1], J. E. Davis[2], S. M. Doe[1], G. Fabbiano[1], E. Galle[1], D. G. Gibbs[1], J. D. Grier[1],
D. M. Hall[1], P. N. Harbo[1], X. He[1], J. C. Houck[2], M. Karovska[1], J. Lauer[1], M. L. McCullough[1], J. C. McDowell[1], J. B. Miller[1],
A. W. Mitschang[1], D. L. Morgan[1], J. S. Nichols[1], M. A. Nowak[2], D. A. Plummer[1], F. A. Primini[1], B. L. Refsdal[1],
A. H. Rots[1], A. L. Siemiginowska[1], B. A. Sundheim[1], M. S. Tibbetts[1], D. W. Van Stone[1], S. L. Winkelman[1], P. Zografou[1]
[1]Smithsonian Astrophysical Observatory, [2]MIT Kavli Institute

Abstract

Chandra Source Catalog (CSC) master source pipeline processing seeks to automatically detect sources and compute their properties. Since Chandra is a pointed mission and not a sky survey, different sky regions are observed for a different number of times at varying orientations, resolutions, and other heterogeneous conditions. While this provides an opportunity to collect data from a potentially large number of observing passes, it also creates challenges in determining the best way to combine different detection results for the most accurate characterization of the detected sources.

The CSC master source pipeline correlates data from multiple observations by updating existing cataloged source information with new data from the same sky region as they become available. This process sometimes leads to relatively straightforward conclusions, such as when single sources from two observations are similar in size and position. Other observation results require more logic to combine, such as one observation finding a single, large source and another identifying multiple, smaller sources at the same position.

We present examples of different overlapping source detections processed in the current version of the CSC master source pipeline. We explain how they are resolved into entries in the master source database, and examine the challenges of computing source properties for the same source detected multiple times. Future enhancements are also discussed.

Overview

The Chandra Source Catalog will contain targeted and serendipitous X-ray sources obtained from processing public observations. Automated pipeline processing will calibrate data, detect sources, calculate source properties, and combine individual detections into a single catalog of master sources. The CSC will be continuously updated, and fully characterized snapshots will be released as the latest catalog version.

Automated source detection is currently performed on a single observation (more precisely, an "observation interval", or "obi"). Sources that are detected together within the same observation interval are referred to as the set of “obi” sources. The master pipeline collects source property information from obi sources and (1) determines which obi sources from different observation intervals are really observations of the same actual sky source (a "master" source), and (2) determines how to combine various source properties (position, flux, variability, etc.) from the individual obi sources into the properties for a single master source.

Combining Multiple Observations of an X-ray Source

Multiple Chandra observations of the same part of the sky often result in the same X-ray source being detected with different apparent sizes and position errors. The same sources may appear on-axis in some observations, potentially showing a more compact, point-like source, and off-axis in other observations. An off-axis observation is impacted to a greater extent by Chandra's non-linear point spread function (PSF), which will cause a point-like source not only to appear as a larger ellipse, but also to potentially have internal structure. Uncertainty in aspect reconstruction can also lead to offsets in source positions.

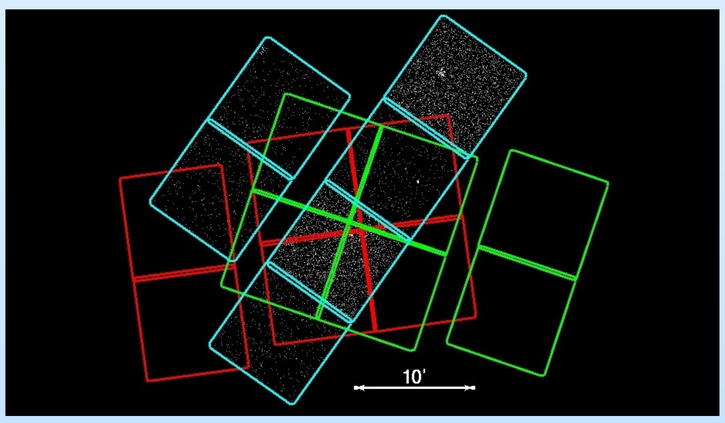

Fig. 1. Sky Region Coverage from 3 Observations of M33

Three different overlapping observations of M33 are shown in Fig. 1, above. Chandra ACIS detector chips are outlined in different colors for each obi. This figure shows the potential for the same source to be detected in multiple observations at different off-axis angles and PSFs . X-ray events are shown for the detector chips outlined in cyan.

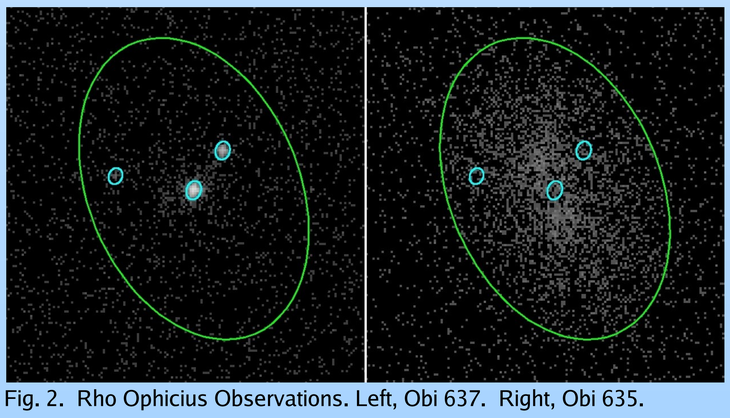

Fig. 2. Rho Ophicius Observations. Left, Obi 637. Right, Obi 635.

Significant differences in source appearances between multiple observations are shown in Fig. 2, where ellipses represent detections in the Rho Ophicius region from two different obis. X-ray events from obi 635 are on the right, and events from the same region as seen in obi 637 are shown on the left. The source detections from 635 are shown on both images as green ellipses, and detections from 637 are shown in cyan. The larger PSF size in 635 is one reason that data from multiple observations cannot simply be co-added prior to source detection. The differing number and size of obi source detections shows some of the complexity of determining how obi sources map to master sources.

Types of Source Overlaps

Individual detections from different obis are combined into a single master source in the catalog by identifying all sources from different obis whose Point Spread Functions (90% ECF) overlap. Empirical analysis indicated that the 90% ECF PSFs were a good proxy for source position uncertainty. Source position uncertainty is a function of several factors, including positional uncertainty and the ability to assign photons into the correct source (more difficult for nearby sources with dissimilar PSFs). The types of overlaps are categorized into three different classes.

1. Merged Obi Sources: Obi sources merge to a single master when multiple detections in different obis mutually and exclusively overlap each other. Fig. 3 shows overlapping obi sources in M33. A single master source will be in the catalog with properties combined from seven overlapping obi sources. The central yellow obi source is the one detected from the X-ray events shown in the image. Larger obi sources, notably in green and magenta, represent detections in obis where this master source was further off axis and at different orientations.

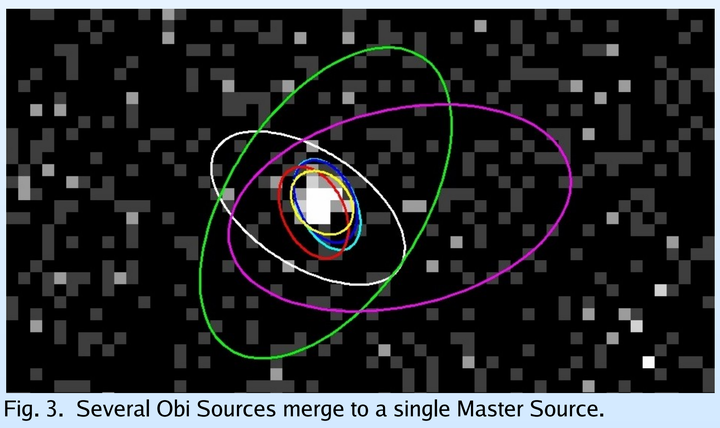

Fig. 3. Several Obi Sources merge to a single Master Source.

2. Confused Obi Sources: The simplest case of a "confused" obi source is a source which overlaps two or more sources in a different obi. An example is shown by the Rho Ophicius sources from Fig. 2. A single, large source, shown in green with its events on the right, is detected in obi 635. The same region resulted in a detection of three distinct smaller sources in obi 637, shown in cyan with their events on the left. Since it is too difficult to determine how events in the larger source map to events in the smaller sources, the green source labeled as confused. The smaller cyan sources will each create a single master source in the catalog. The green, confused source will be identified as overlapping the master sources during catalog information retrievals, but its properties will not contribute to the master sources' properties.

3. Too Hard Obi Sources: Sometimes a set of overlaps cannot be automatically resolved into master sources. An overlap set like this is simply referred to as “too hard”, and an example is shown in Fig. 4. Detections are shown from seven different colored obis in the M33 region, with X-ray events from the obi of the larger, green detection. The two central red sources each overlap two cyan sources, which qualifies them as confused. However, since the cyan sources also overlap both red sources, they also qualify as confused, and no master sources are generated for the confused sources to link to. Recognizing this, automated processing tags them as needing manual review. They and all contiguous sources will be resolved into master sources outside of the master pipeline. The addition of future detections from additional obis may also resolve this prior to manual review, in which case the obi sources are assigned to other categories and no review is needed.

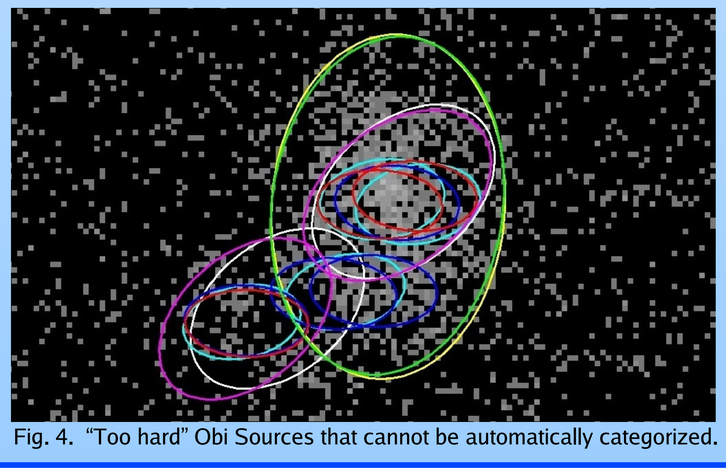

Fig. 4. “Too hard” Obi Sources that cannot be automatically categorized.

Updating the CSC with New Obi Sources

The master pipeline generates a set of eight different actions for the CSC to update its set of master sources and their properties, The two primary actions are to MERGE or LINK an obi source with a master source. In the case of a MERGE, the obi source can only merge with a single master source, and that master source's properties include the obi source's properties. For a LINK, the obi source is still associated with the master source, but its properties do not contribute to the master source's properties. An obi source can be linked to multiple masters.

To determine what actions are necessary to update the CSC when a new obi is processed, the state of the master sources before and after the addition of the new obi sources is compared. In either the old (before new obi) or new (after new obi) state, obi sources will fall into one of the three source overlap categories described above, or be absent entirely. A specific set of instructions is generated for each combination of old and new states, an example of which is shown in Fig. 5, below, for a merged (old) obi source. The instructions are generated in a manner which guarantees that the final state of a master source will be the same independent of the order in which individual obi sources are applied.

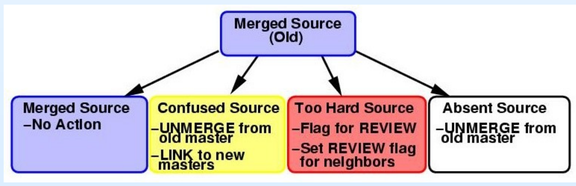

Fig. 5. Actions Generated from Old State (Merged) to New State

The CSC will have fixed releases with corresponding version numbers, but will also be available on a daily basis with the most recently processed obis. Since updates to master sources can happen while the CSC is being accessed by users, a method is needed to prevent partial updates while source data are being retrieved. This is accomplished by implementing all the actions that may affect a master source at the same time (i.e., locking the master source's information, updating that information, and then releasing all of the updated information together).

Merging Obi Source Properties to a Master Source

For CSC Release 1 all obi source properties are determined on a per-obi basis and master source properties are determined by combining individual obi properties, or selecting a specific single obi source's property. CSC Release 2 will feature improved master source properties by co-adding multiple observations of the same field from the same instrument before source detection is performed.

For Release 1, various methods are used to determine master properties from obi source properties. For example, some master source properties will be populated by taking an error weighted average of a property from all obi sources merged to that master. These properties include RA, Dec., and various aperture fluxes for each energy band. Other properties, such as per-energy band spectral fit derived fluxes, are populated by selecting the single equivalent obi source property for the obi source which has the largest value of flux significance.

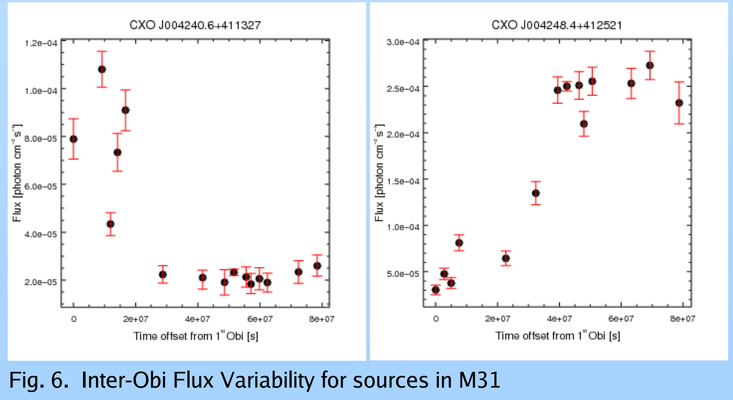

Fig. 6. Inter-Obi Flux Variability for sources in M31

The master source property of inter-obi variability is determined by comparing individual aperture source photon fluxes among all contributing obi sources. For two master sources in the M31 region, the values of aperture source photon flux in the broadband energy band is shown along with error bars in Fig. 6, above. For both of these sources, photon flux varies by about a factor of five during observations made over an approximately two and half year time period. In each case, the broadband inter-obi variability flag is set to TRUE from these data.

Acknowledgments

Support of the development of the Chandra Source Catalog is provided by the National Space and Aeronautics Administration through the Chandra X-ray Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of the National Aeronautics and Space Administration contract NAS8-03060