Chandra Source Catalog Review

Characterization

Ian Evans
Chandra X-ray Center
February 9, 2006

Goals of Catalog Characterization

- Provide quantitative measurements of key statistical properties that will enable prospective users to assess whether the current state of the catalog can support their science goals

- Summary current status:
  - Preliminary simulation studies for source detection mostly complete
  - Ready to start baseline characterization studies using the prototype pipeline
    - Characterization will be used to feed back enhancements for production pipeline

Key Items To Be Characterized

- Limiting sensitivity
- Completeness
- False source rate
  - Above are function of source detection process
- Astrometric accuracy
- Photometric accuracy
- Time variability analysis

Limiting Sensitivity

- Limiting sensitivity
  - Limiting source *flux* that can be detected at a specified confidence level
  - Function of detector, position on detector, off-axis position, background (instrumental and sky), source morphology, energy spectrum, ...
  - Complicated to quantify *as a flux* when using hypothesis test based detection algorithms (**`wavdetect`**)
    - Reviewing quantitative approach developed by Kashyap et al. to determine flux sensitivity limit

Completeness and False Source Rate

- Completeness
  - Fraction of real sources actually detected for a specified detection threshold (most usefully source flux)
  - Similar dependencies to limiting sensitivity
- False source rate
  - Fraction (number) of non-sources incorrectly detected as real sources
  - Function of detector, position on detector, off-axis position, background (instrumental and sky), ...
  - **`wavdetect`** allows the user to specify a false source threshold (e.g., 1 false source per chip)
    - Assumes background determination is completely accurate
    - Instrumental artifacts such as chip edges, ACIS readout streak, chip streaks (ACIS-S4), pile-up, variable background, ... can be detected and do not contribute to the false source threshold

False Source Rate II

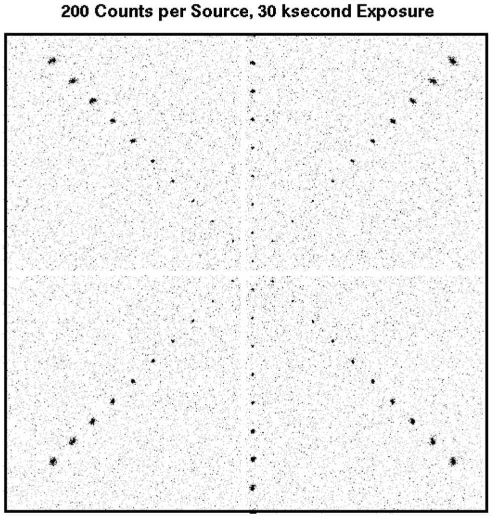
- 2-D fitting post-**`wavdetect`** modifies false source statistics
  - Characterization must quantify the fraction of false detected sources that are rejected by fitting (primarily due to detections of instrumental artifacts)
  - Characterization must quantify the fraction of false sources that are created due to improper fitting (i.e., fitting two point sources when there is only a single real source is present)

Source Detection Characterization

- Limiting sensitivity, completeness, and false source rate are evaluated by characterizing the results of the pipeline source detection process (`wavdetect` + 2-D fitting)
- Preliminary results developed through detailed analysis of well studied datasets and simulations
  - Studied results of simulations and investigations performed by the *ChaMP* project
  - Performed independent simulations where needed to complement above studies
    - Most simulations were not run using the prototype pipeline
  - Manually investigated results of source detection applied to a subset of public observations

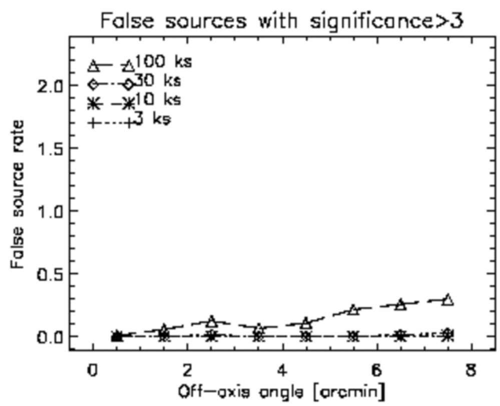
Preliminary Source Detection Characterization

200 Counts per Source, 30 ksecond Exposure



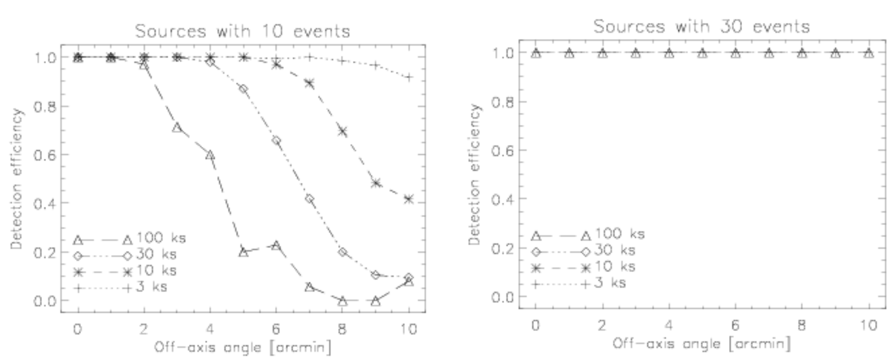
- Simulate blank ACIS-I fields with exposures of 3, 10, 30, 100 ks (assuming nominal background rates)
- Add simulated point sources with 10, 30, 100, 200, 2000 counts at 1' increments in off-axis angle (θ), up to 10' off-axis, for various values of azimuth (ϕ)
- Run `wavdetect` (as in the detect sources pipeline) and compute:
  - False sources detected per field of view as a function of counts, exposure, θ
  - Fraction of sources detected as a function of counts, exposure, θ

Preliminary False Source Rate Characterization



- **`wavdetect`** false source thresholds work as advertised (see also Kim *et al.* 2004, ApJS, 150, 19)
  - Applies only to false sources detected as statistical fluctuations
  - Image artifacts such as readout streaks or unscreened bad pixels can increase this rate
- Plot shows false source rate *per field* versus off-axis angle
  - *Worst case 0.3 false sources per field for long exposures (typically much lower)*

Preliminary Detection Efficiency Characterization



- Preliminary characterization:
  - *Efficiency ≥ 95% for source counts ≥ 10, $\theta < 2'$*
  - *Efficiency ≥ 95% for source counts ≥ 30, $\theta > 2'$*
- Plots show fraction of detected sources versus off-axis angle

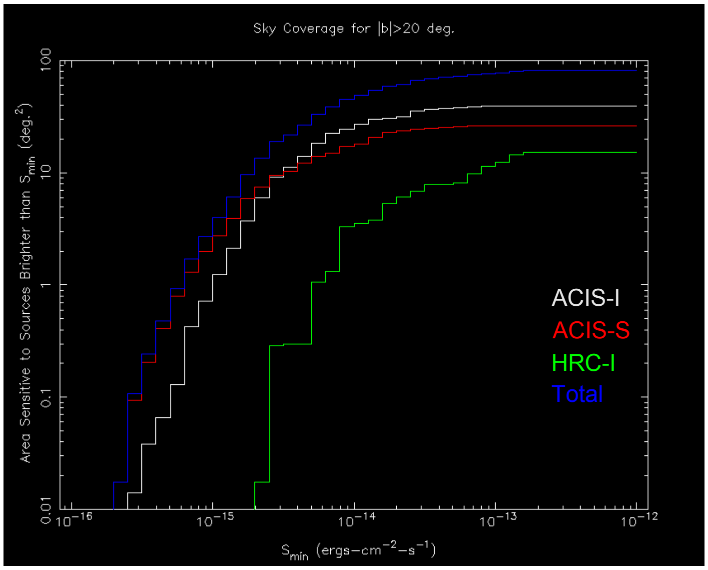
Limiting Sensitivity Characterization

- Preliminary characterization
  - Use two-level "count map", exposure time, instrument geometric areas and PIMMS rate-to-flux conversion factors for assumed spectrum to estimate overall sky coverage as a function of limiting source flux
- Ultimate goal
  - Estimate limiting source counts as a function of background
  - See (*e.g.*), eq. 4 of Lehmer *et al*. 2005, ApJS, 161, 21:

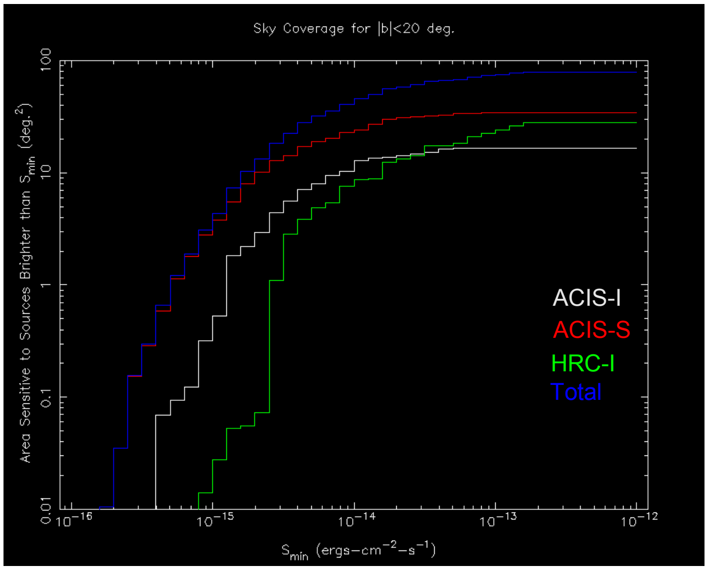
    $$\log(N) = \alpha + \beta \log(b) + \gamma [\log(b)]^2 + \delta [\log(b)]^3$$

    where α, β, γ, and γ are empirically determined constants
  - Use background maps generated from data (not blank-sky) and correlation maps generated by `wavdetect` to estimate minimum source counts over the field
  - Combine with the exposure map to derive the limiting sensitivity map

Sky Coverage of Current Public Archive I: Preliminary



Sky Coverage of Current Public Archive II: Preliminary



Future Work

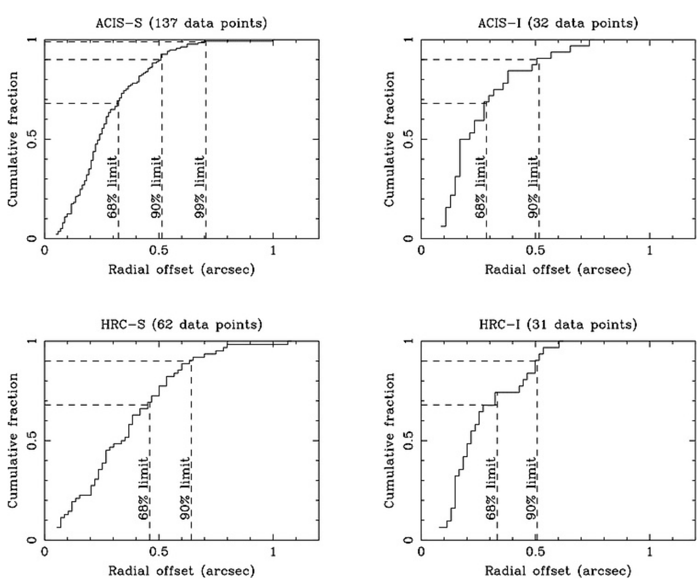
- The following investigations will improve the limiting sensitivity characterization
  - Elimination of false sources due to instrument artifacts
  - Investigation of effect of additional source detection (and elimination) in 2-D fitting on Type 1 and 2 errors
  - Generation of background maps from data for all ACIS chips and for HRC-I and HRC-S
  - Computation of sensitivity maps from **`wavdetect`** output
  - Combination of sensitivity maps into mosaics
  - Computation of sky coverage from sensitivity maps

Astrometric Accuracy I

- Two sources contribute to astrometric uncertainty
  - Uncertainty inherent in the applied aspect solution
    - Error in *Chandra* astrometric frame
  - Uncertainty contributed by source detection process
    - Error in measured source positions within frame

- The uncertainty inherent in the applied aspect solution is a measure of the accuracy of the *Chandra* reference frame
  - Established by comparison of measured X-ray source positions from *Chandra* observations with Hipparcos/ Tycho 2 and ICRS VLBI catalog positions
    - 90% uncertainty radius is 0.6"
    - 99% uncertainty radius is 0.8"

Measured Aspect Solution Astrometric Accuracy



Astrometric Accuracy II

- Uncertainty contributed by source detection process (2-D fit position)
  - Depends on source counts and off-axis angle
    - Will dominate for large off-axis angles
  - Plan to characterize as a function of source counts and off-axis angle through a combination of simulation and study of actual field data, comparing with known source positions determined independently from other wavebands

Photometric Accuracy

- Source photometry is determined in the pipeline by three independent methods:
  - Integrated flux from the 2-D fit
  - Aperture photometry
  - Spectral fitting (for broad-/soft-/medium-/hard-band fluxes)
- Allows cross-comparison consistency estimates
- The accuracy of the photometry of a source depends primarily on the accuracy of the exposure map, PSF fraction, 2-D fit, spectral fit, dither fraction, and pile-up
  - The last two issues are difficult to characterize accurately, but fortunately affect only a very small fraction of catalog sources

Photometric Calibration Dependencies: Exposure Map

- Exposure map
  - Issue
    - Used to convert from counts to photons cm^{-2} s^{-1}; depends on the source spectrum for conversion from to ergs cm^{-2} s^{-1}
  - Plan
    - Characterize error introduced by using incorrect source model when computing exposure map
    - Generate simulated spectra for an ensemble of sources and corresponding exposure maps for a diverse sample of observation configurations and source counts; compare with exposure maps computed with simple power law model spectra

Photometric Calibration Dependencies: PSF Fraction

- PSF fraction
  - Issue
    - Applies correction for fraction of PSF not included in source region
    - Assumes point source morphology
    - Depends on PSF model source spectrum
  - Plan
    - Characterize error as function of source size and source spectrum by generating simulated spectra for an ensemble of compact sources with a range of spectra, source counts, morphologies, and extents; compare source fractions in source region with computed PSF fractions using monochromatic energy spectrum

Photometric Calibration Dependencies: 2-D Fit

- 2-D fit accuracy
  - Issue
    - Determines whether source region includes a single point source, a close double source, or compact but extended sources
    - Computes source positions and integrated source fluxes
  - Plan
    - Characterize ability to identify and differentiate source types by generating an ensemble of simulated source region containing sources with appropriate characteristics as a function of source counts, off-axis angle, background, morphology, separation, and contrast ratio, and perform 2-D fit
    - Characterize robustness of derived source positions and integrated source fluxes by a combination of simulation and studying source regions from real fields with well understood and independently verifiable source properties

Photometric Calibration Dependencies: Spectral Fit

- Spectral fit accuracy (band fluxes)
  - Issue
    - Compute ARF and RMF, spectral fit parameters, and derive band fluxes
    - ARF must be corrected for fraction of PSF outside of source region
  - Plan
    - Characterize error in ARF and RMF, spectral fit parameters, and derived band fluxes by generating an ensemble of simulated sources with various spectra and source counts, applying spectral fit process, and comparing resulting fit parameters
    - Characterize error introduced in ARF correction by generating an ensemble of simulated PSF models with various source spectra and compare with results using monochromatic PSF model for a range of source regions
    - Evaluate suitable count thresholds for spectral fit complexity

Photometric Calibration Dependencies: Dither Fraction

- Dither fraction
  - Issue
    - Correction for fraction of time that source and background regions are dithered across chips
      - Photometric correction via exposure map
      - Spectral fit correction applied by using weighted ARFs
    - *Affects only the small fraction of sources* that have source and/or background regions that dither across chip boundaries
  - Plan
    - See also Exposure Map item
    - Study the error introduced in spectral fitting by using weighted average responses through simulations by generating ensembles of sources with varying spectra and source counts, explicitly computing detailed responses as a function of chip position and comparing with weighted averages
    - Plan to flag affected sources in catalog

Photometric Calibration Dependencies: Pile-Up

- Pile-Up
  - Issue
    - Central "cratering" of heavily piled-up sources confuses source detection and 2-D fitting
    - Photometric and spectral properties of piled-up sources incorrect
    - *Affects only the small fraction of bright sources*
  - Plan
    - Characterize false source rate and 2-D fitting errors by studying actual datasets with piled-up sources
      - Plan to detect and flag potentially false sources detected nearby heavily piled-up sources
    - Plan to flag piled-up sources in catalog with the observed count rate per pixel per frame (ACIS)
      - Need to quantify flag threshold by studying correlation between observed count rate per frame and estimations of the percentage pile-up based on pile-up models

Time Variability Analysis

- Characterization of time variability analysis depends on total counts, source and background time history during the observation interval, and the variability timescale *as a fraction of interval duration*
  - Fractional variability timescale is typically less useful for users who are searching for candidate sources displaying specific variability characteristics, which are independent of the duration of the observation
- Approach will be to simulate temporal characteristics of specific types of sources (e.g., bursters, X-ray pulsars, ...) and characterize properties of computed variability index as a function of the remaining free parameters

Characterization of Merge Processing I

- Characterization of merge processing will be conducted once the prototype merge pipeline is implemented
  - Goals
    - Ensure that merge processing is performing optimally
      - Verify appropriate merge algorithms are used for each source property (weighted average, best per-observation, ...)
        - Revise merge algorithms if necessary
      - Verify that best estimates for sets of merged properties remain internally consistent
        - Example: Baseline plan is to merge broad-band fluxes separately from high-/medium-/low-energy band fluxes, so that the best estimate of the broad-band flux can be derived from a combination of ACIS and HRC observations. We expect that this will *not* introduce significant inconsistencies cf. the merged band fluxes (which can only be determined from ACIS observations). However, if this turns out not to be the case statistically, then we must consider whether consistency of the fluxes is more important than reduced broad-band flux errors

Characterization of Merge Processing II

- Goals (continued)
  - Characterize merged source properties
    - Verify merge improves determinations of source properties, and quantify expected improvement as a function of observation and source properties
      - Compare merged source properties with per-observation properties and verify scatter is decreased
- We will use a combination of simulations and actual data to characterize the merge processing

Characterization of Merge Processing III

- More details of the characterization plan will be developed when the merge processing requirements are more completely defined
- Characterization of merge processing will be refined using production data after the start of catalog production
- Characterization will continue to be improved in future catalog releases
  - These will include enhanced merge processing to provide improved handling of multiple observations of the same field
    - Expected to increase in significance due to exposure limitations resulting from spacecraft thermal constraints

Evolution Of Characterization I

- Current Status
  - Preliminary studies to validate the approaches used to define the prototype per-observation pipeline functionality were performed via simulations and analysis of selected well-studied observations
    - Made use of results of simulation studies and analyses performed by *ChaMP* group
    - Attempted to identify range of observations that could "trip-up" pipeline processing
      - Including crowded fields, extended sources, ...
    - Studies of the behavior of the source detection process are mostly complete

Evolution Of Characterization II

- Baseline (detailed) characterization of catalog properties
  - Goals:
    - Support quantitative characterization of catalog properties appropriate for initial catalog release
    - Provide scientific and technical feedback to prototype pipeline design and implementation
      - Does it break? Does it do the right thing? Is it efficient?

Evolution Of Characterization III

- Baseline characterization follows on from science evaluation testing (recently started)
  - Plan is to run a number of observations (at least ~100) and simulations through the prototype pipeline for science evaluation testing
    - Observations selected to support the overall characterization effort, and to "stress" the pipeline
    - Simulations chosen to address specific characterization items and resolve pipeline issues
    - Combine observations with "injected" simulated data for higher fidelity
- ***Plan to release baseline (detailed) characterization at the same time as the catalog***
  - However, will not hold up catalog release if characterization is largely complete, well-understood, and stable

Evolution Of Characterization IV

- During operations, *spot-checks* will be performed to ensure continued integrity of process and verify characterization has not changed
  - *Manual V&V will not be performed for every observation or source (due to resource limitations)*
- ***Characterization will be verified and updated if necessary following major pipeline, algorithm, or calibration changes***
- Users will be informed about the current state of catalog characterization, and will be able to determine the catalog properties appropriate for any catalog snapshot