
CHANDRA
SOURCE CATALOG

Catalog overview

- **Master (merge) source catalog**
  - Core catalog: 1 entry per source, lots of properties per source
  - For sources in more than one observation, best estimate combined properties
- **Per-observation source database**
  - Same as core catalog, but each source has a separate entry for each observation it appears in. 10 counts on axis, 30 counts far off axis
- **Data products**
  - 'Catalog' includes 'products': event files, PHA files and responses, etc.
- **CAT pipeline**
  - Sofware releases on separate schedule from CIAO. Currently running CAT2.6.1
  - CAT2.7 is pre-production test
  - CAT3.0 is production release
- **Documentation**
  - Requirements doc and internal memos (maintained by Ian)
  - User docs – we need to do this

Source Catalog Energy Bands

- **U (ULTRA-SOFT) 0.2-0.5 keV**
- **S (SOFT) 0.5-1.2 keV**
- **M (MEDIUM) 1.2-2.0 keV**
- **H (HARD) 2.0-7.0 keV**
- **B (BROAD) 0.5-7.0 keV**
- **W (WIDE, HRC) 0.1-10.0 keV**

Source Catalog project

- **Calibrate phase**
  - Screening with tighter background flare removal (like lc_clean)
- **Detect phase**
  - High spatial frequency background derived from 1D streak map
  - Low spatial frequency background made by cutting out sources and blocking
  - Final background map combines these two
  - Wavedetect run; merge bands
  - Quality assurance step to automatically weed out some spurious sources
    - significance > 3
    - flag very-crowded fields, fields with background 5 times normal
  - SAOTrace run to generate PSFs for each source (used for extent matching)
  - Make source region files for each source

Source Catalog project

- **Per-Observation Source properties**
  - Source position and error ellipse
  - count rates, fluxes, colors, spectral fits
  - Extent, time variability
  - Create source product files (per-source events, PHA etc)
  - Create catalog entry in database
  - 2-sided confidence intervals where possible (rather than 1-sided sigma)
- **Merge phase**
  - Source name CXO J140700.0+273420 (ICRS)
  - Best estimate of properties based on combining values and errors
  - Attempt to make results indepedendent of processing order
  - Create catalog entry in merged catalog database
- **Characterization phase**
  - describe systematics of catalog: false source rate, det.effy.,sky coverage
  - limiting sensitivity/upper limit at arbitrary point on sky

Catalog schedule

- **Mid Jan- CAT2.6 Science analysis of test runs**
  - SDS look at data, identify problems...
- **February - CAT2.7 pre production test**
- **March: SDS quick feedback on CAT2.7 data**
- **April (early) revised algorithms and constraints**
- **April: preliminary characterization of CAT2.7 ensemble**
- **May - CAT 3.0 code freeze**
- **June - begin production**
- **July? - First public data access, 3 years processed, prelim. characterization**
- **Oct-Nov - Formal release 1**
- **2009: Release 2 with merged-observation detect**

Catalog resources

- **Requirements doc at https://icxc.harvard.edu/soft/schedule/L3/**
- **Other memos, meeting notes in same place**
- **I'd like to also keep copies of SDS memos at /data/jcm/sds/memos**
- **Catalog is run by AP (Automatic Processing) just like Level 1/2 pipes**
- **Outputs go into archive**
- **Cache data products in /data/L3/**
  - e.g. /data/L3/AP.20071207/00786_001/DET/acisf00786_001N001_evt3.fits.gz
- **FITS file binary table with dump of catalog from CAT2.6 run (35 obsids) in /data/evans_i/L3/CAT2.6.1_2007Dec12_prop3.fits**

Source Catalog project

- **Included in original CXC plan (SE03 doc, CDR)**
- **Endorsed by external science review**
- **Deemed by CUC and HT as a high priority**
- **External review recommended simple release 1 and more detailed version later.**

Source Catalog project

- **What is the minimal catalog?**
  - Must not be dominated by spurious sources
  - Must have positions and count rates and UNCERTAINTIES for those
  - Must be able to characterize whether a given subset may be dominated by spurious results
  - Must include central regions of most (almost all?) public ACIS observations

Source Catalog project

- **What is in and out for release 1?**
  - Out: detect on merged OBSIDs
  - In: merged source properties for sources present in more than one OBSID
  - *So: we miss faint sources that will only be seen in coadded deep fields: this is a major science limitation on release 1*
  - Out: Gratings observations
  - In: (maybe) HRC observations – but may be dropped, *background issues are giving too many false sources*
  - In: large off axis angles in ACIS
  - Out: Medium-extended sources (more than 1 arcmin)
  - Out: Entire fields with Highly-extended sources (*many false detections of knots in Cas A, etc.*)

Properties for each source (1)

- **Where?**
  - Position
  - Source region ellipse RS (and cutouts for nearby contaminating sources)
  - 90% ECF region ellipse RE (based on model PSF)
  - *Off-axis angle, chip coords, chip edge flag, multi chip flag*
- **How bright?**
  - Count rates in RS and RE; *total counts, ECF for RS*
  - Aperture-corrected fluxes in RS and RE (scale each photon by ARF)
  - Hardness ratios (colors (M-S)/B etc)
  - Fluxes based on assumed 1.7 power law and Gal NH
  - Fluxes based on 1.0 keV blackbody and NH=3E20
  - Spectral fits to PL and to BB (if more than 250 counts)
  - Pileup flag, source significance

Properties for each source (2)

- **How big?**
  - Extent flag, Confusion flag
  - Convolved extent: how big a blob on the image; ellipse parameters
  - Deconvolved extent (compared to PSF): how big is the source really (ellipse)
- **How variable?**
  - Kuiper, KS, Gregory-Loredo prob, Var. Index
  - Mean count rate, sigma, min, max in RS
  - Dither warning flag

Limiting sensitivity maps

- Frank Primini developed method
- Add Poisson noise to model background map
- Use tabulated PSF model to determine 90 percent aperture radius at each pixel
- Find count rate that gives SNR=3 for a point source at that pixel given the PSF size and background estimate.

Limiting sensitivity maps

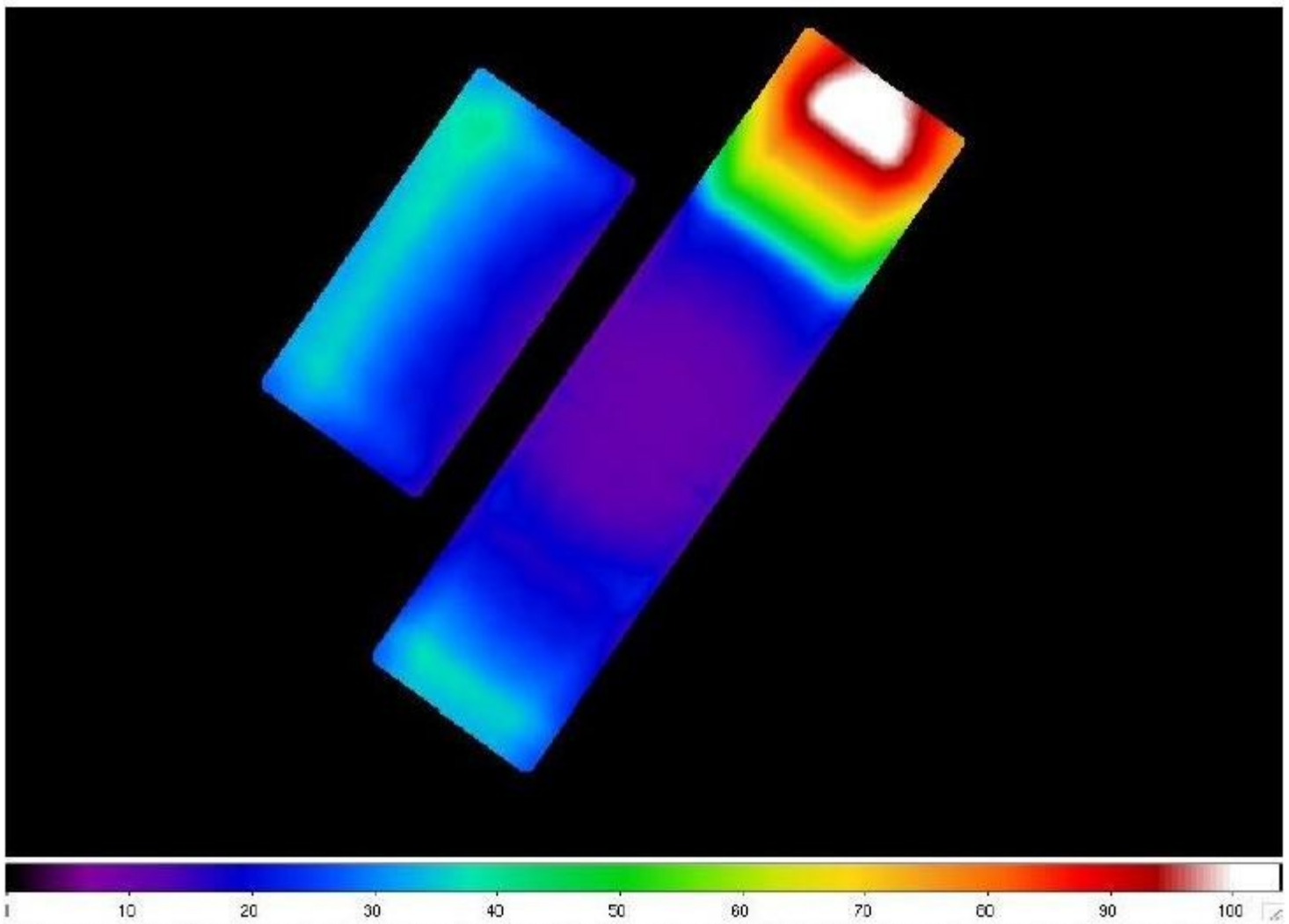
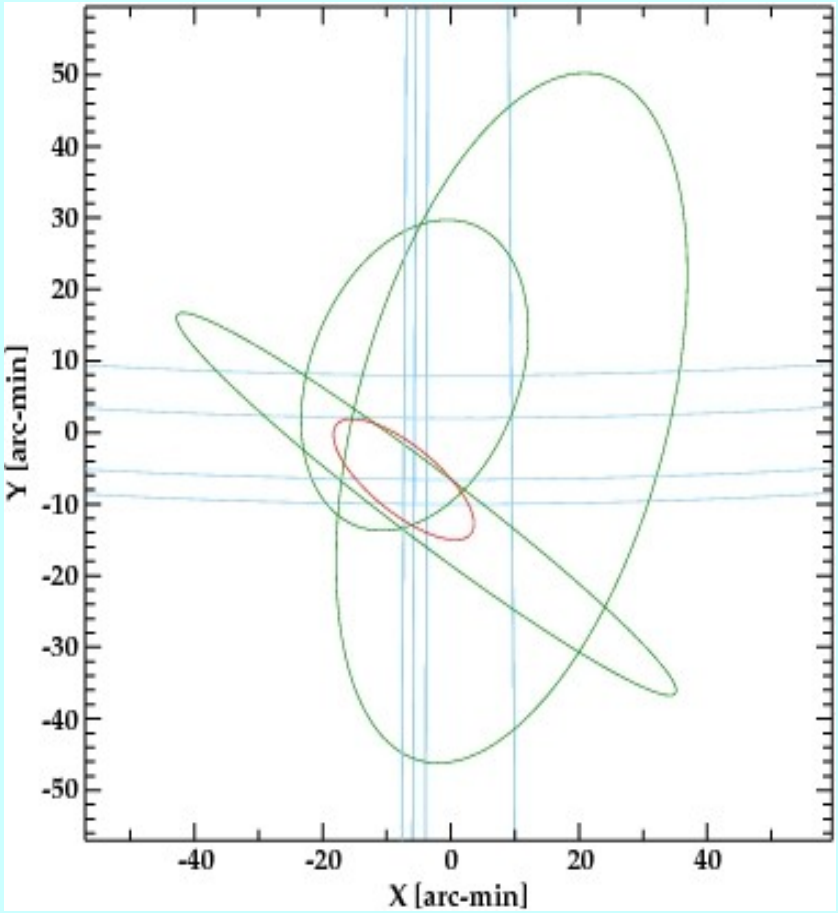


Figure 4.1: Net Counts Limiting Sensitivity Map for OBSID 786, broad band, blocking factor = 4

Combining error ellipses – John Davis

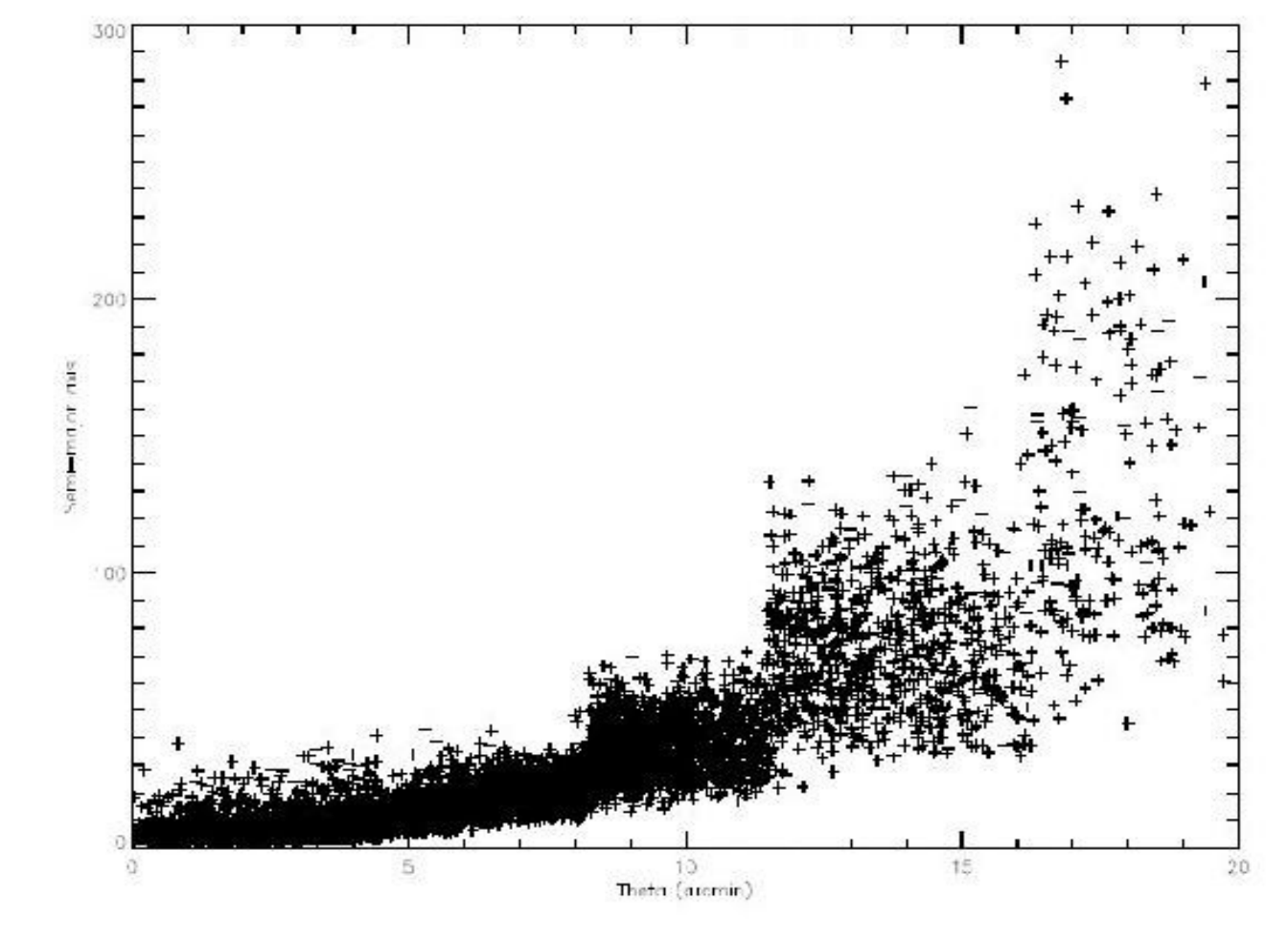
- Improved estimate of position and error ellipse
- Adopted algorithm used for military targeting
- Unbiased estimate assuming no systematic error in positions
- Takes spherical geometry into account (works at poles)
- Math memo available



Source extent comparison – John Houck

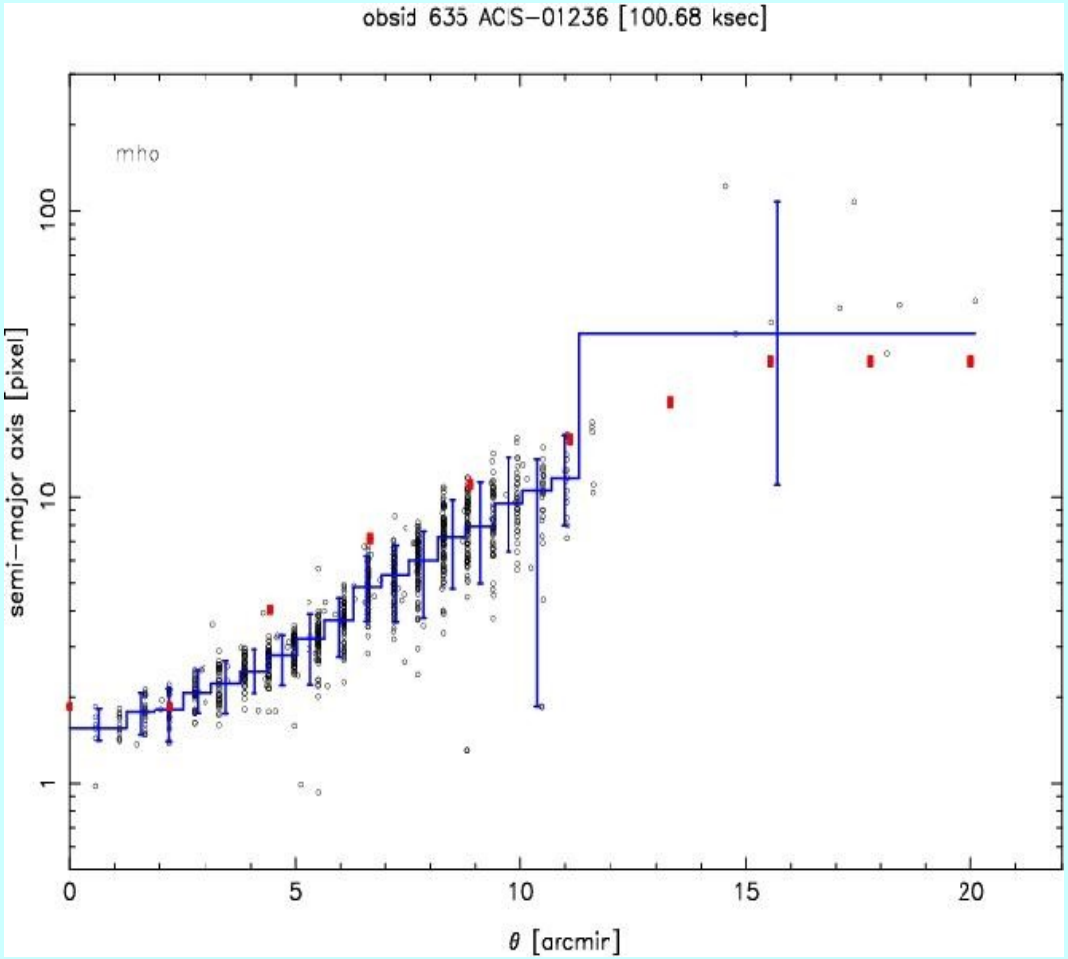
Wavdetect approach introduces steps associated with the discrete wavelet scales used – hence discontinuities vs off axis angle.

Extent estimates are not accurate; discontinuities in other properties too

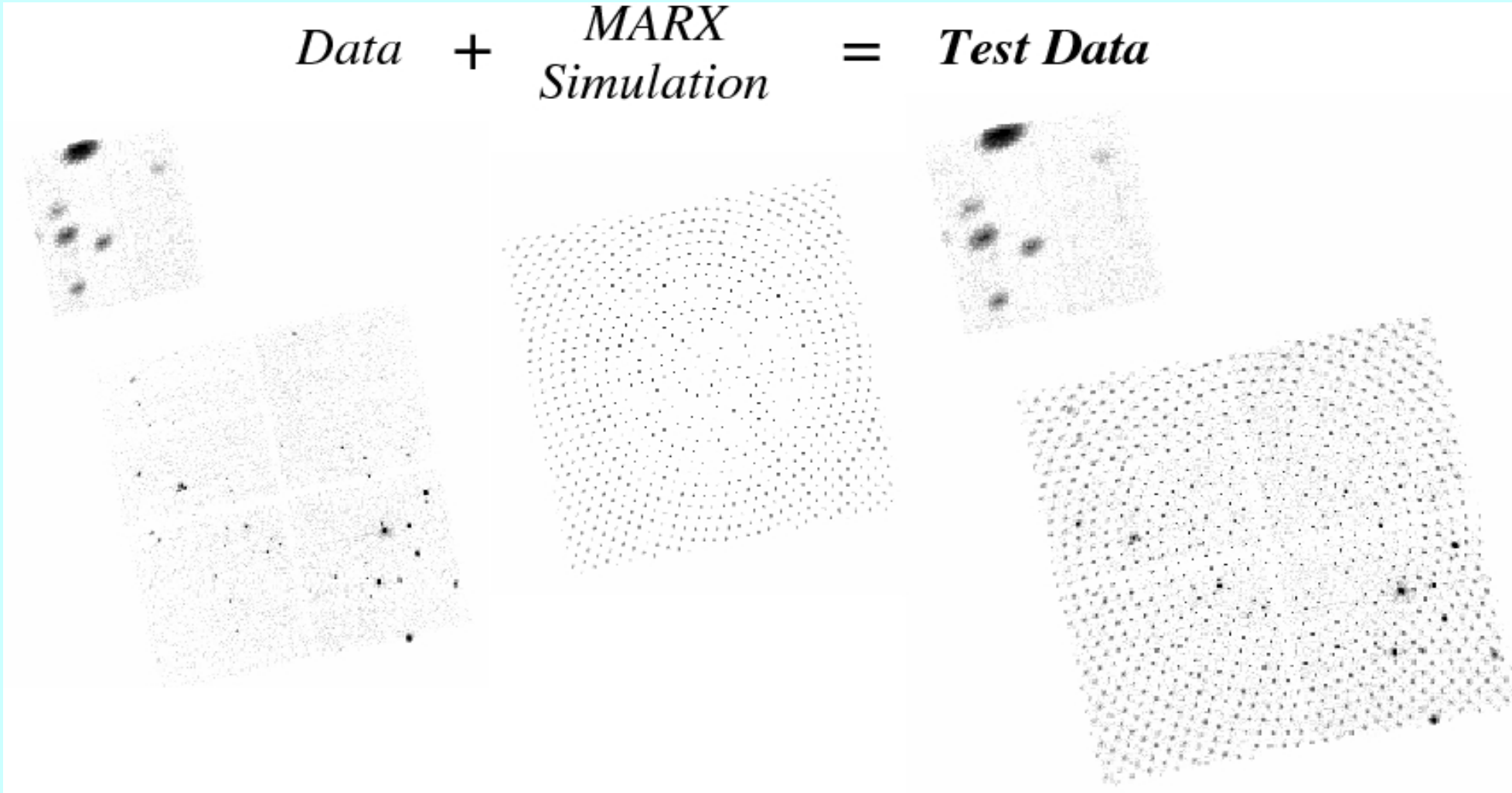


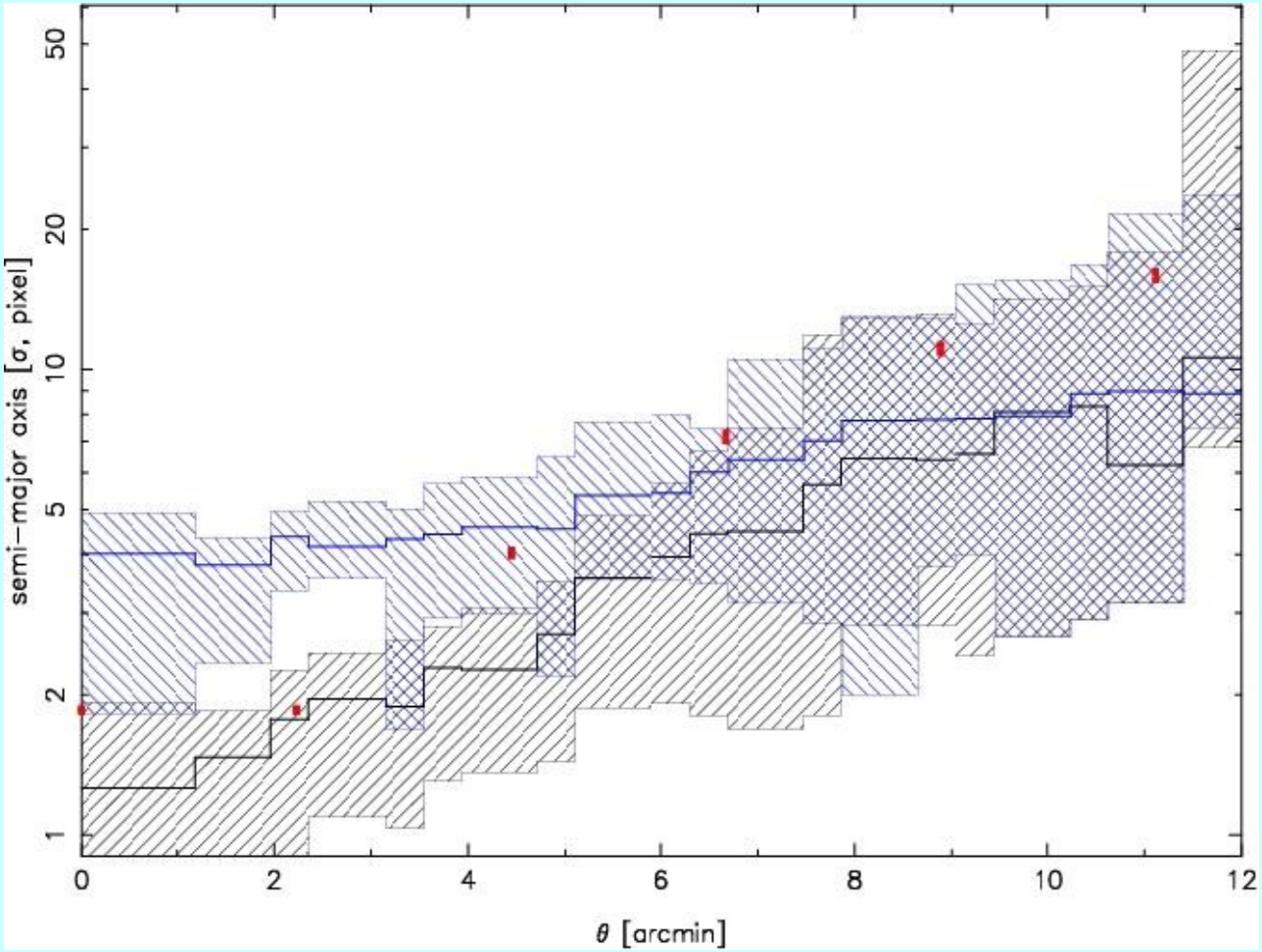
Source extent comparison

New method (below) is scale free and has no discontinuities; figure shows 1000 simulated 150-count sources (blue line) and nominal SAOSAC PSF extent (red). Method post-processes wavdetect source regions, maximizing correlation with an elliptical wavelet



Source extent evaluation – John Houck





Black line: point sources with 35 counts

Blue line: 2” radius disks with 35 counts

Hatched regions give stat uncertainty range

Distinguishable within 5' off-axis

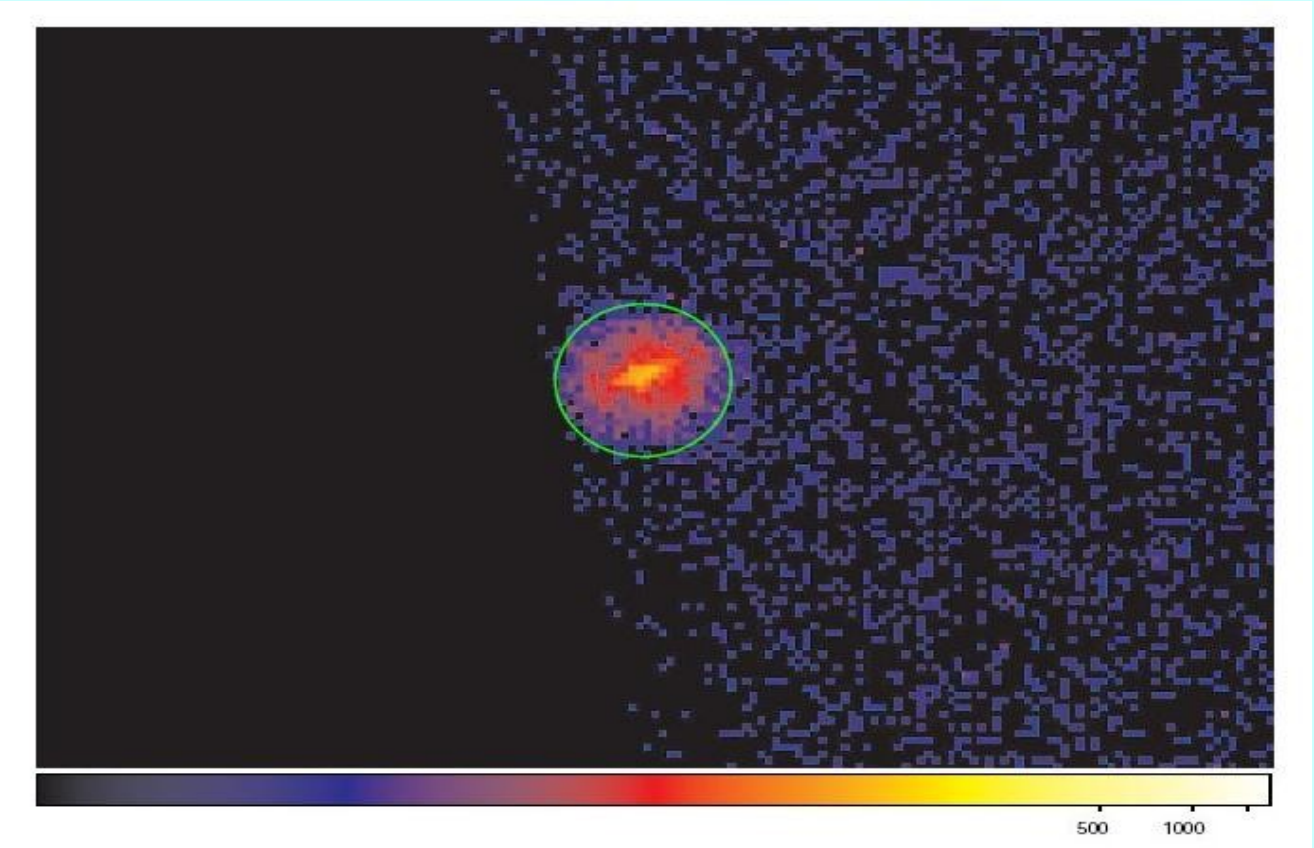
Variability and dither

Catalog will have KS, Kuiper variability tests as well as Gregory-Loredo algorithm

Testing GL with dead-time/dead-area/dither taken into account

Uses new dither_region tool gives fractional source area on chip versus time (including bad pixel correction)

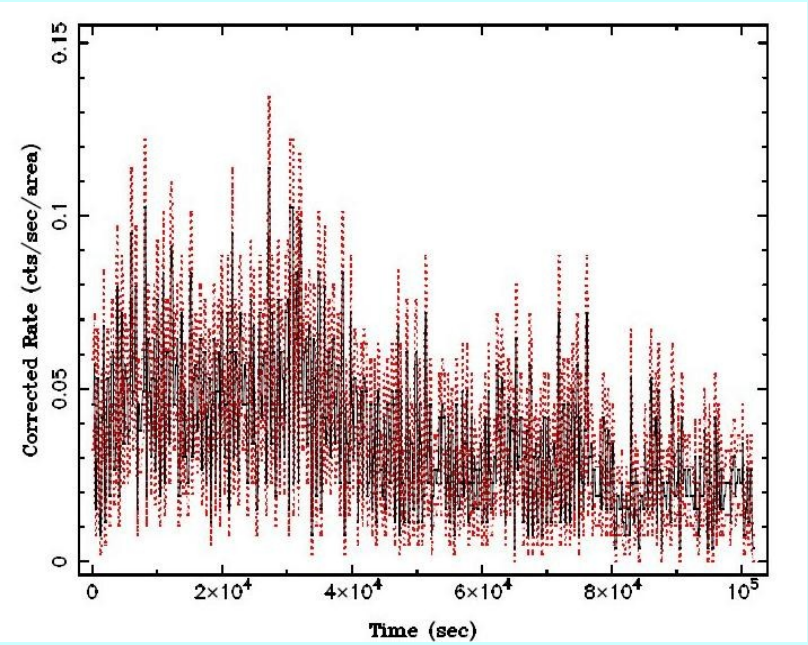
635:

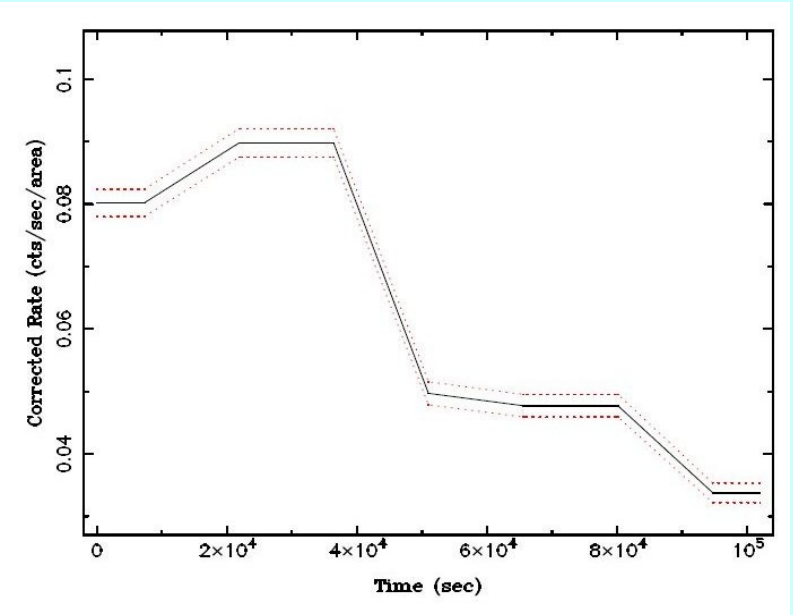


Without dither correction: dither dominates light curve

With dither correction: clear measurement of variability

Algorithm automatically chooses binning for significance





Science issues remaining - examples

Frank is working on combining posterior probability distributions for formally correct errors on merged source count rates [DONE!]

Problem with exposure thresholding at edge of chips –can lead to incorrect wavedetect results (too many sources).

Reviewing cases of spurious source detections (e.g issue with wrong background normalization on subarrays) – check if significant contribution to source catalog (must fix) or just small absolute number of cases (can live with for release 1?). QA will remove most of these false detections anyway

Tweak variability tests to take exposure variations into account and incorporate background variability (M Nowak, A Rots)

Using Monte Carlo simulations to further characterize and calibrate position and extent uncertainties (Houck, Primini)

New science issues

- **Background flare filtering – KJG needs algo to find source-free regions (action Mike**
- **Overlap logic for master catalog – needs review (volunteer?)**
- **Spectral fit hard limits – verify (action Aneta)**
- **Support for John's extent algorithm in new Sherpa (action Aneta and John H)**