# The Chandra Source Catalog: Background Determination and Source Detection

**Michael L. McCollough[1], A. H. Rots[1], F. A. Primini[1], I. N. Evans[1], K. J. Glotfelty[1], R. Hain[1],**
**C. S. Anderson[1], N. R. Bonaventura[1], J. C. Chen[1], J. E. Davis[2], S. M. Doe[1], J. D. Evans[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], 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], B. L. Refsdal[1],**
**A. 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

The Chandra Source Catalog (CSC) is a major project in which all of the pointed imaging observations taken by the Chandra X-Ray Observatory will be used to generate one of the most extensive X-ray source catalog produced to date. Early in the development of the CSC it was recognized that the ability to estimate local background levels in an automated fashion would be critical for essential CSC tasks such as source detection, photometry, sensitivity estimates, and source characterization. We present a discussion of how such background maps are created directly from the Chandra data and how they are used in source detection.

The general background for Chandra observations is rather smoothly varying, containing only low spatial frequency components. However, in the case of ACIS data, a high spatial frequency component is added that is due to the readout streaks of the CCD chips. We discuss how these components can be estimated reliably using the Chandra data and what limitations and caveats should be considered in their use.

We will discuss the source detection algorithm used for the CSC and the effects of the background images on the detection results. We will also touch on some the Catalog Inclusion and Quality Assurance criteria applied to the source detection results.

This work is supported by NASA contract NAS8-03060 (CXC).

## Background Map Creation in the Chandra Source Catalog

Early in the development of the CSC it was recognized that the ability to estimate local background levels in an automated fashion would be critical for essential CSC tasks such as source detection, photometry, sensitivity estimates, and source characterization. Presented is a discussion of a method where such background maps are created directly from the *Chandra* data.

The general background for *Chandra* observations is varying rather smoothly, containing only low spatial frequency components. However, in the case of *Chandra* ACIS data, a high spatial frequency component is added that is due to the readout streaks of the CCD chips. As a result the background maps for *Chandra* ACIS data consist of two components. One is a high spatial frequency component due to readout streaks and the other is a low spatial frequency component. A discussion of the algorithm used to generate both components and how they are combined is given below.

## High Spatial Frequency Maps (Streak Map)

X-Ray sources observed with *Chandra* ACIS will produce high spatial frequency structure in the background caused by readout frame transfer (readout streak). This steak is the result of the finite amount of time (40 $\mu$sec per row) that it takes to read out the CCD after an integration period. Consequently, all pixels along a given readout channel are exposed to all points on the sky that lie along that readout channel. The result is that for readout channels which contain bright sources there will be a streak along the length of the readout channel.

The key to producing the streak map is the determination of source free regions of a given observation for each binning/chip/energy band. These regions should then give a measure of the background + readout streak contribution. For ease of computation we take the X-axis to be along the chip rows (across readout channels) and Y-axis is taken be along the direction of the readout channels. The chip rows which are deemed to be source free then will be used in the computation of the streak map.

### Determination of Source Free Regions

Sum along the X-axis ($X_{sum}$), across readout channels. From the $X_{sum}$ values a histogram is created (excluding off-chip and dither regions). Determine the median, mode, and standard deviation ($\sigma$) from the histogram. These quantities should be representative of the basic characteristics of the background. $X_{sum}$ values which greatly exceed these quantities represent rows which have a substantial source contribution. From an examination of many data histograms the maximum value of $X_{sum}$ which is still considered background dominated is best determined by:

$$\mathbf{X_{sum}(max) = min(median + n \times \sigma, 2 \times mode)},$$

where n is a parameter that one can vary; currently a value of 1 is used. All rows with $X_{sum} \leq X_{sum}$ (max) (excluding off-chip and dither regions) are then considered source free regions and are used to calculate the streak map.

### Creation of the High Frequency Background Map

Using only rows which have been determined to be source free sum along the Y-axis (direction of the readout channels) and divide by the total number of source free rows. The result is a distribution of counts/pixel vs. readout channel. This represents the average background + Readout Streak contribution per readout channel. This is replicated (copied) for each row of the chip to create a background image which also contains the Readout Streak contribution.

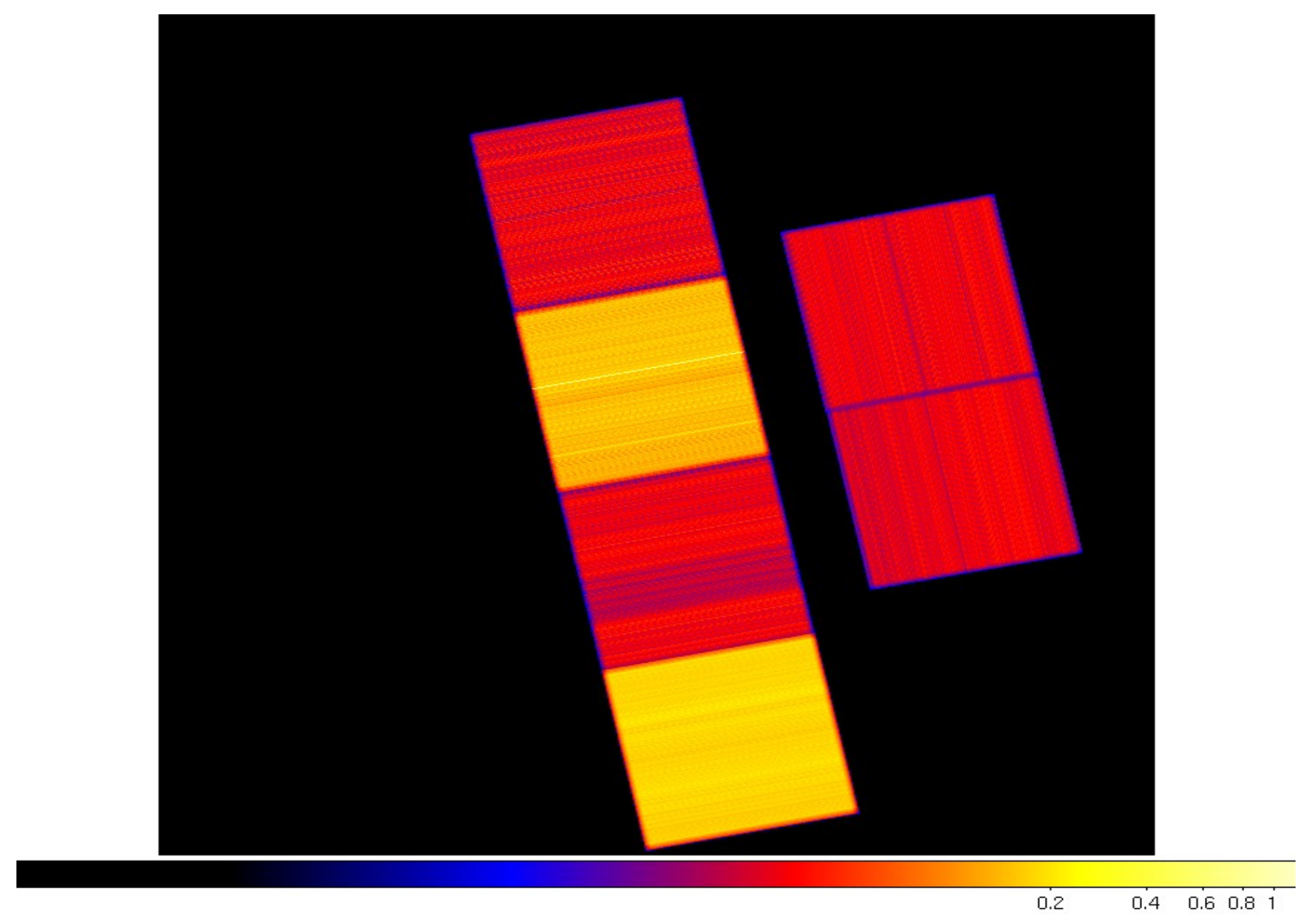

Fig. 1: A streak map created for OBSID 735. Note the higher background in the two back illuminated chips.

## Low Spatial Frequency Maps (Modified Poisson Mean)

As noted in the previous section the streak map contains a 1-dimensional low frequency component which represents an average of the low frequency background over the rows used to create the streak map. What is needed is a 2-dimensional low frequency map which accounts for variation of the low frequency component across the image.

### Sampling Region

The streak map is first subtracted from the image from which it was created. This removes the high frequency component from the image as well as the low frequency component that is contained in the streak map. For each point in the subtracted image a box of *nsize* by *nsize* pixels centered on that point is defined as the sampling region. The edges of the box are *nsize*/2 away from the point. The value *nsize* represents the low frequency filter size (sizes smaller than this are filtered out and attenuated). Near the edges of the maps the area is truncated and some higher frequency information may propagate into the map. But this has not been found to create any major impacts on the low frequency map.

### Modified Poisson Mean

For the sampling area of each point in the image a histogram (h) is created of the count distribution. Since the readout streak has been subtracted there will be some negative count pixels. As a result bin 0 will typically represent the range between -0.5 and +0.5. From the histogram the highest bin (a) and its higher neighbor (b) are identified. The background is calculated as:

$$\mathbf{b_{lf} = mean(h[a] \cup h[b])}.$$

This is a modified form of a Poisson Mean. That is essentially the mean value of all of the pixels contained in the histogram columns of h[a] and h[b] . This process is done for each point in the map. It is expected that there will be negative values in this map. This map serves to modify the low spatial frequency component that is contained in the high spatial frequency map (streak map). For the ACIS observations *nsize* = 129 . This corresponds to a scale size of ~1 arc minute.

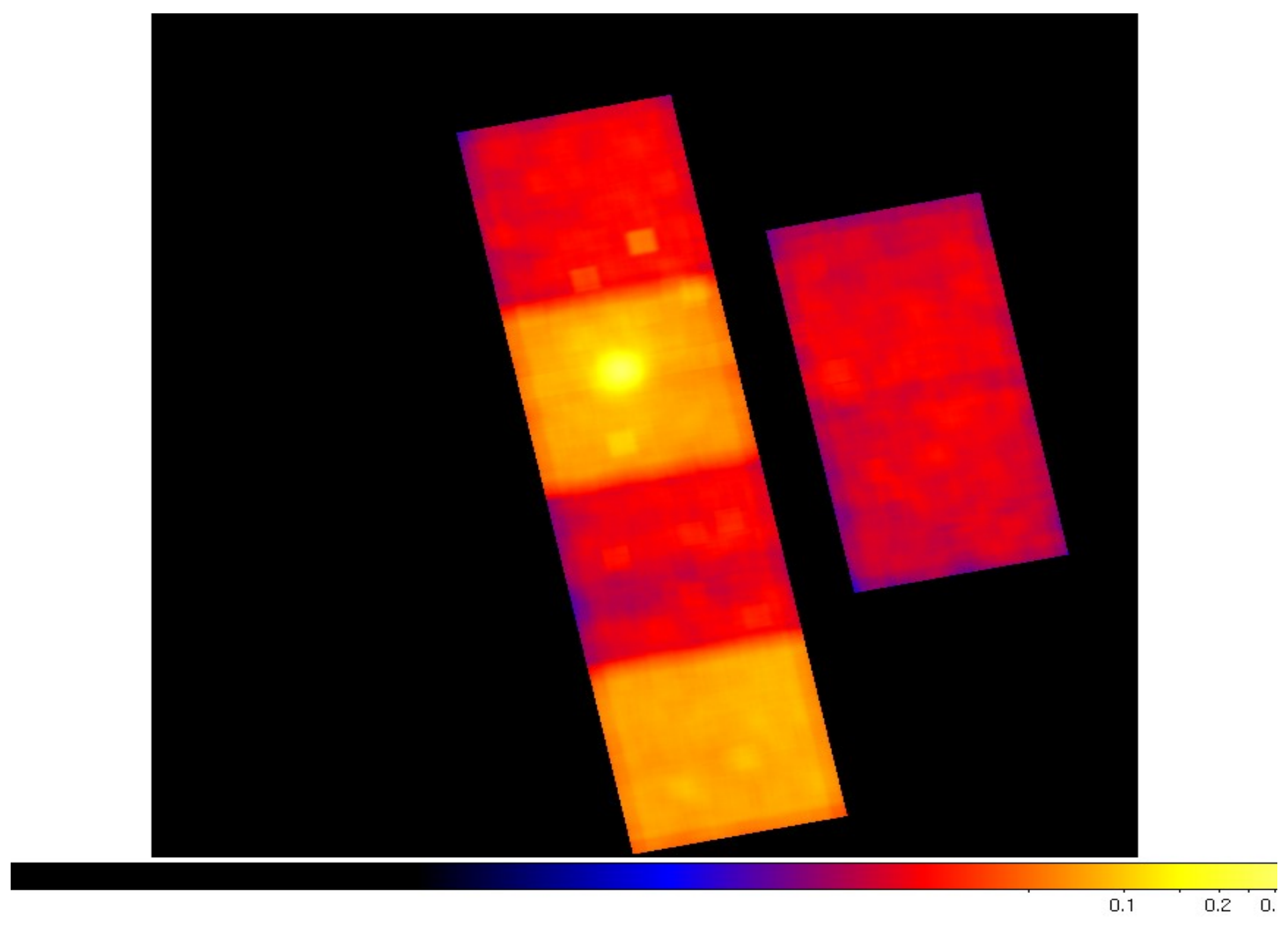

Fig. 2: A low spatial frequency map created for OBSID 735. Note the higher background in the two back illuminated chips and the low frequency spatial component of the bright sources.

## Total Background Map

The final background map is created by: **(1)** Adding the streak map to the low frequency map; **(2)** Creating an exposure map for the image and normalizing it so that its maximum value is one; **(3)** Dividing the combined map from first step by the normalized exposure map. This last step is to make exposure adjustments to the edge of the background where the effects of dither come into play.

One can generate similar background maps for *Chandra* HRC data. These data are dominated by the low spatial frequency maps for the HRC since this detector does not suffer from the read-out streak problem found in *Chandra* ACIS data. Hence one follows the procedure for creating the low frequency map but without the need for a streak map.

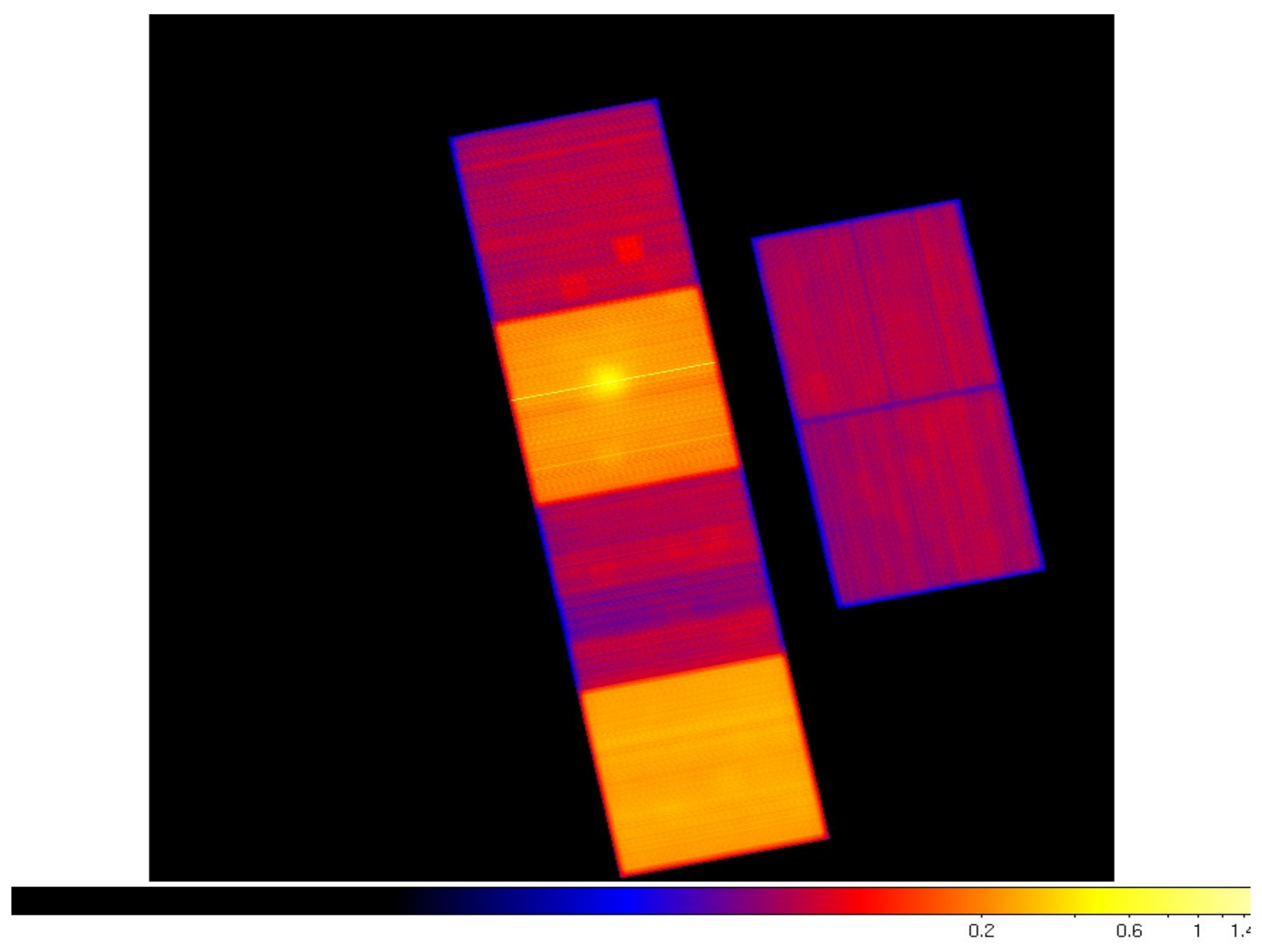

Fig. 3: A total background map created for OBSID 735. The maps created in Figs. 1 and 2 were added and the an exposure map created for the observations was applied.

## Source Detection (WAVDETECT)

The wavelet-based source detection algorithm WAVDETECT (Freeman et al. 2002, ApJS, 138, 185) was used for the CSC. It was found to work well in crowded fields and identified point sources superimposed on extended emission.

Only an approximate PSF shape is needed and the edge of field effects are not a major problem. A point source source is most easily detected when a wavelet function of similar size to the PSF is used; since the PSF varies substantial across the field a set of wavelet scales are used. The source detection in the CSC is sensitive to source sizes up to ~ 0.5 arc minute.

WAVDETECT is run on 2048 square images for each detection band (ACIS: b, s, m, h; HRC-I: w ) and each binning (ACIS: 1, 2, 4; HRC-I: 2, 5, 12). The largest binning in each case encompassed the entire observation. See Rots et al. (Poster 472.3) for a definition of the different energy bands.

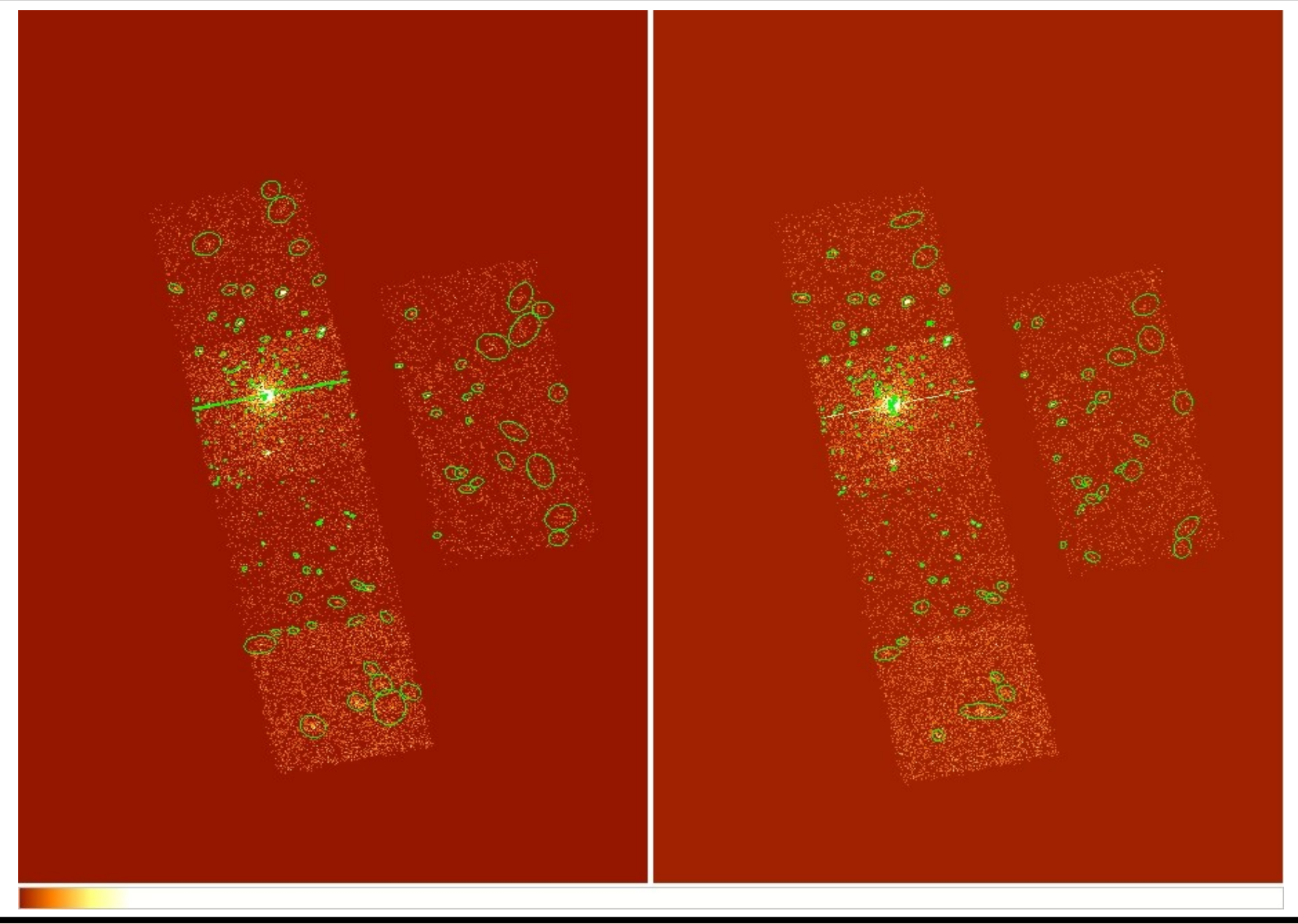

Fig. 4: *left:* A broad band image for OBSID 735 with WAVDETECT source detection without using a background map. Note the false sources due to the readout streak. *right:* Same observation but using a background map.

## Catalog Inclusion and Quality Assurance

After detection the sources are filtered through a catalog inclusion (CI) step and a quality assurance (QA) step. The main CI step guarantees that all accepted sources have at least 7 net counts and have a significance greater than 2.7 $\sigma$. The QA step tends to be more restrictive (significance of 3.0 $\sigma$ or greater) and also places restrictions based on source properties (region too small, too elliptical, etc.).

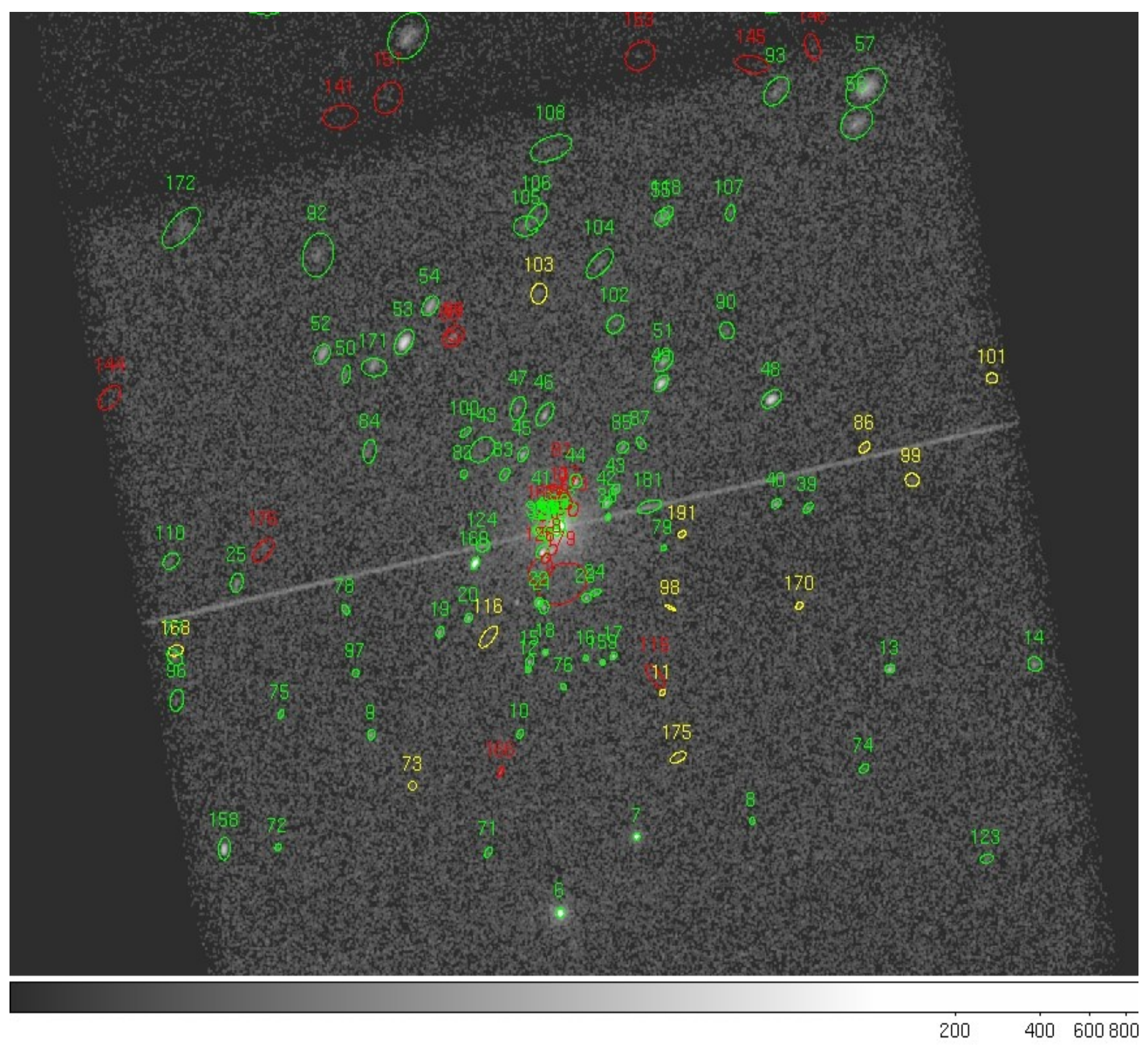

Fig. 5: Source detections for the central region of the OBSID 735. The green represent good sources. The yellow and red represent sources which fail various CI and QA criteria.

More Information about the Chandra Source Catalog can be found at the URL http://cxc.cfa.harvard.edu/csc/