CSC Release 1: Missing sources in crowded fields

Jonathan McDowell and Amy Mossman

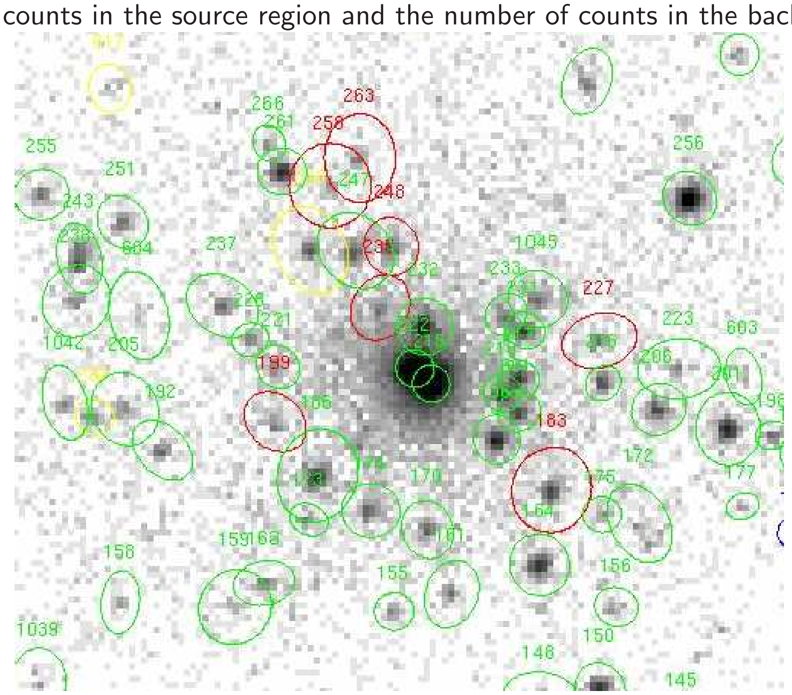
2008 Sep 16 (Rev 1)

During the course of CSC data review it was reported that some clearly visible sources were being missed. This note reports an investigation of such sources in obsid 6420.

ObsID 6420, globular cluster M17, has a bright central confused region with many point sources.

Some of these sources are detected by our initial wavdetect run but do not make it through our quality assurance (QA) and catalog inclusion (CI) filters. In particular, the CI filter imposes a source significance threshhold.

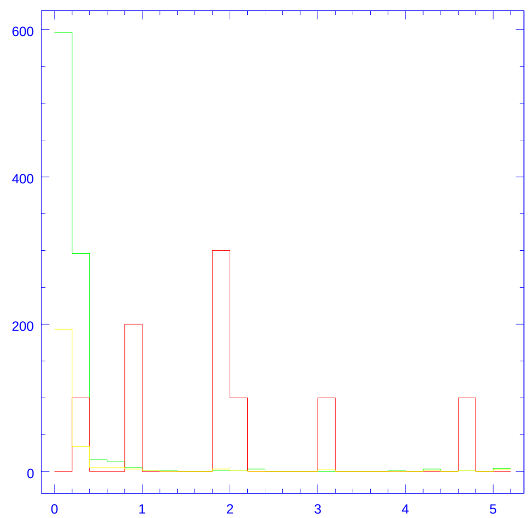
We began with region files provided by Ian Evans. We make three categories: 'red' are sources rejected even though a visual inspection suggests they should be included; 'yellow' are sources which were rejected and are not obviously real; 'green' are sources which make it into the catalog. The red sources have plenty of counts, but their source significance is low and they appear relatively unconfused. The source significance involves, among other things, the number of counts in the source region and the number of counts in the background region.



Hypothesis: red sources are being rejected because their background regions include a lot of photons from the central confused region, and so the backgrounds are not representative of the true background in the source region.

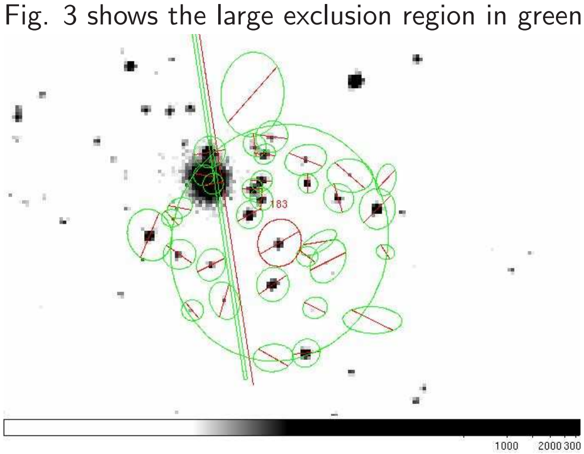
We tested this hypothesis by looking at the pixel value histograms in the background regions. For each source, including rejected sources, we retrieved the region file (srcreg3) from the archive and used dmcopy to extract the background region extension into a separate file, for ease of use. We then filtered the event file on the background region, binned to an image, and made a histogram of the resulting image pixel values. In all cases, these histograms were strongly peaked at a mode of 0 counts per pixel, with a sharply falling tail to higher count densities.

For distributions like these, the mean count density seems the most robust measure of the distribution width. We calculated the mean of each histogram, and made a histogram of these means:



In this diagram, the red histogram shows the means for the red (incorrectly rejected) sources. This histogram has been rescaled by a factor of 100 relative to the green (true detected) and yellow (spurious source) histograms. It can clearly be seen that the means for the red sources are mostly much higher than typical. This agrees with a visual inspection of the data, where bright extended emission is clearly visible in several of the red-source background regions.

Fig. 3 shows the large exclusion region in green around source 183 (source region in red).



Visual inspection also suggests that the problem could be avoided by the use of smaller source regions (e.g. perhaps 70 percent ECF regions). The existing source regions include significant area with little source contribution, unneccessarily reducing the significance. For a smaller source region, our pipeline also creates a correspondingly smaller background region, which would be less likely to overlap with the contaminating emission. Such a change would make the detection less sensitive to extended sources in regions of low background, so in future releases a multiple detection criterion should be considered (e.g. source is deemed to be detected if EITHER the wavdetect-region significance OR the 70-percent-ECF-region significance is sufficiently high).