False Detection Rate of Source Finding

W. D. Cotton (NRAO) and W. Peters (NRL) August 8, 2011

***Abstract*—Astronomical images always contain a randomly distributed component to pixel values which is unrelated to anything on the celestial sphere. The statistics of this distribution must be taken into account in order to establish the significance of features in the image. The nature of this distribution depends on the imaging instrument; in this memo we consider that derived from radio interferometers, in particularly those operating at low frequencies. In this regime, the non–celestial component of the distribution of pixel values is dominated by deficiencies in the calibration due to the ionosphere and the poorly known antenna pattern and is distinctly non–Gaussian. We develop a technique of modeling the pixel statistics directly from the distribution in a given image in order to estimate the probability that a given level feature is a random event. This technique is applied to a reprocessing of the VLA 74 MHz VLSS survey and validated using the VLA 1400 MHz NVSS survey. The method does not produce acceptable results for this data.**

***Index Terms*—Radio Interferometry, Source cataloging**

I. Introduction

WIDEFIELD astronomical images, especially those produced by sky surveys, are frequently decomposed into a list of discrete components to produce a catalog of objects. Due to a random component of the pixel values not of celestial origin, some criterion must be adopted to distinguish features in the image which correspond to plausibly real objects from those unlikely to be real. In the following this non–celestial component of the pixel value distribution will be called "noise". Such tests must invariably involve a trade off between the possibility of missing real objects and contamination of the catalog by fictitious ones.

Some measure of the statistical probability of a feature being due the the noise distribution is required. Such tests are well established in cases where the noise has a Gaussian distribution. The simplest such test is a cutoff at some multiple of the Gaussian σ of the distribution where values in excess of this are expected to be sufficiently rare. For more sophisticated applications of Gaussian statistics to astronomical images, see [1], [2].

In some applications, the noise distribution is not well modeled by a Gaussian and Gaussian statistics underestimate the number of false detections. One such case is low frequency radio interferometry where difficulties with modeling the effects of the ionosphere and the antenna patterns result in a nontrivial fraction of the celestial power being scattered into bogus features. The highly variable nature of the ionosphere further aggravates this problem.

In the following, we develop techniques of estimating the false detection probability directly from images statistics and apply them to a reprocessing of the VLA 74 MHz VLSS survey [3]. Here, false detection rate (FDR) is the likelihood that a feature could be a random occurrence due to the noise distribution. Since this is a relatively shallow survey, the vast majority of detected objects are expected to appear in the deeper VLA 1400 MHz VLSS survey[4]. The catalog derived from the VLSS images can be compared with the NVSS catalog to derive a statistical measure of the false detection rate actually achieved. The techniques discussed are implemented in the Obit package ([5], http://www.cv.nrao.edu/∼bcotton/Obit.html).

W. Cotton is with National Radio Astronomy Observatory, 520 Edgemont Rd., Charlottesville, VA, 22903 USA email: bcotton@nrao.edu

W. Peters is with Naval Research Lab, Along some river, Washington DC, ????? USA email: wendy.peters@nrl.navy.mil

II. Low Frequency Interferometric Image Statistics

At low radio frequencies accurate calibration of the data is compromised by the poorly known antenna pattern and frequently seriously degraded by ionospheric fluctuations. Discussions of ionospheric calibration techniques are given in [6], [7], [8].

Two examples of pixel distributions from reprocessed VLSS images are given in Figure 1. These illustrate a well behaved field and a poorly behaved one. The positive tail in excess of the negative tail of the distributions are dominated by real sources; however, the negative tail is expected to be a combination of thermal noise and calibration and imaging artifacts. In the 1240+150 field, the positive and negative tails of the distribution are nearly the same indicating that few sources can reliably be identified at low flux density values. Clearly, Gaussian statistics are not a good model for the distribution of values well away from zero, even in well behaved cases.

III. Estimation of Actual Noise Distribution

Real sources undoubtedly contribute much of the positive wing of the pixel distributions seen in Figure 1, at least for the 0800+650 field, but the negative wing of the distribution should be an indication of the true noise distribution. If the true positive wing of the noise is symmetric with the negative wing, the negative half of the distribution can be used to estimate the corresponding positive half. Thus, the excess positive values in a given interval over the negative values in the corresponding negative interval give the fraction of the positive values related to real sources. In this simple case, the false detection rate for positive values in this interval can be expressed as:

$$FDR_x \;=\; 1 \;-\; \frac{n_+ - n_-}{n_-}$$

where FDR_x is the false detection rate at flux density level x, n_+ is the number of pixels in the positive x bin and n_- is the number of pixels in the negative x bin.

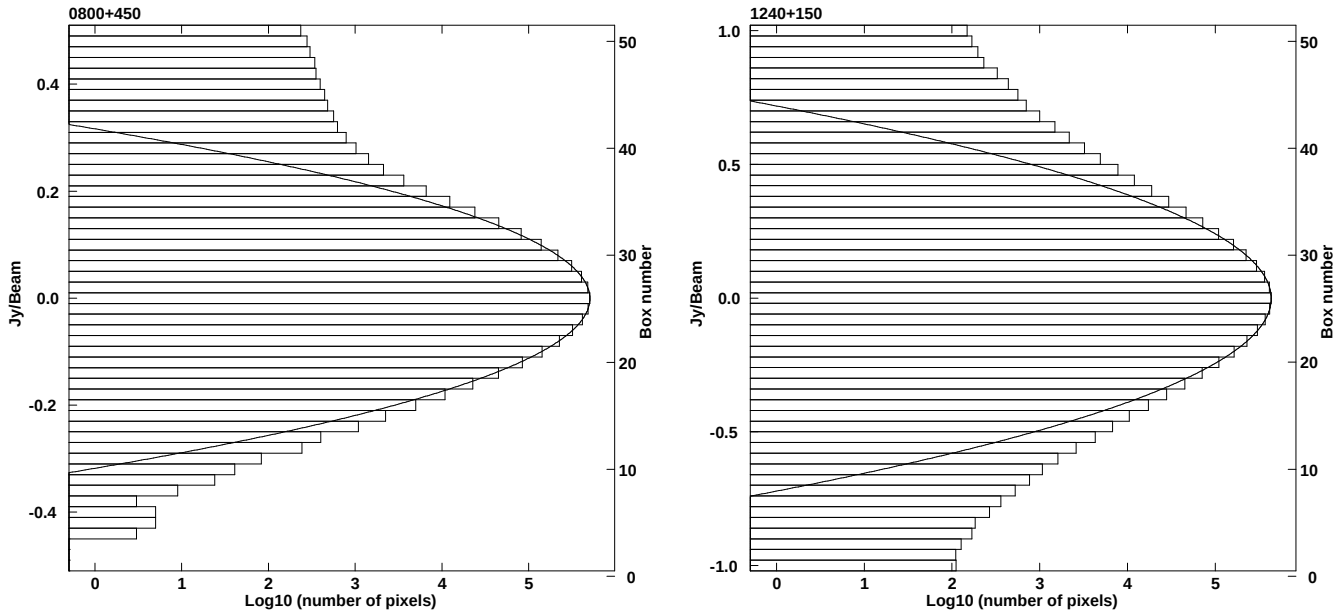

Fig. 1. Pixel histograms for two VLSS fields. Solid lines are the Gaussian fitted near the peak of the distribution. On the left is a well behaved field and on the right, one containing a very bright and extended source. The negative tail of the distributions indicate the non-Gaussian nature of the distribution of values.

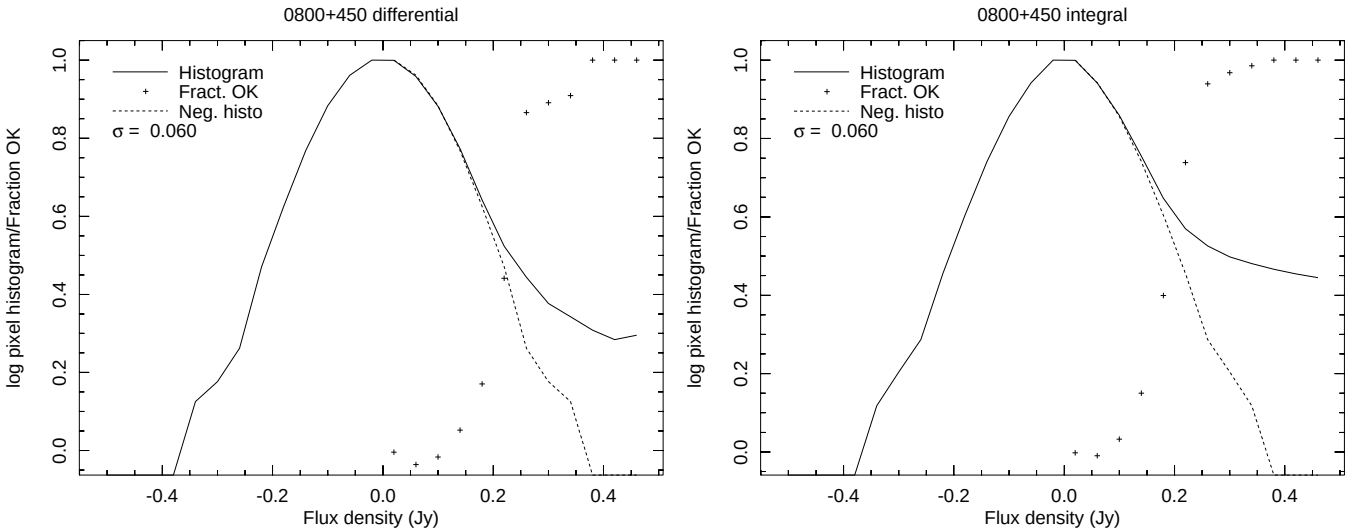

Fig. 2. Differential and integrated histograms from the same set of pixels. The solid line is the histogram normalized to 1, the dashed line is the negative half folded to the positive half and the "+"es are 1-FDR. On the left is the differential histogram and on the right the integrated version. The region sampled is a subset of that shown in Figure 1.

Determining flux density levels for low false detection rates requires good statistics well out in the wings of the distribution. This generally means sampling large numbers of pixels and including large areas. On the other hand, the character of the noise may change across the image meaning statistics over more limited areas are preferred. Images like the VLSS fields are composites of images derived at different pointings and the statistical properties of the noise can be quite variable. One compromise to make the statistics more robust is to integrate the pixel histogram; each bin then includes the counts in that bin plus all bins further from zero. This makes a subtle change in the meaning of the false detection rate given above; it is the probability that a pixel at a given level, or more. has a given probability of being a random event. The difference between differential and integrated histograms is illustrated in Figure 2. The flux density corresponding to a given false detection rate can be interpolated from the values shown as "+"es in Figure 2 right.

IV. Implementation

A histogram analysis of the false detection rate as described in the previous section was implemented in the Obit package. In the c library, the ObitPixHisto class implements the functionality with bindings in python as the PixHistFDR class. The

source finding task, FndSou was modified to allow selecting sources by estimated false detection rate with statistics derived from a box of a given size centered on the source in question.

V. Testing on VLSS

The processing techniques for analyzing low frequency data have improved significantly since the original processing of the VLSS survey data, especially in the areas of ionospheric calibration and the excision of interfering signals. The techniques discussed in this memo were derived to be used in the cataloging of the reprocessed VLSS images.

A. VLSS Reprocessing

The reprocessing of the VLSS survey followed that in the original except improved ionospheric calibration techniques were used to the exclusion of self calibration and interfering signals were estimated and subtracted from the data. Linear mosaics were formed from the overlapping single pointing images.

B. Cataloging

The mosaic-ed VLSS fields were used by Obit task FndSou to generate catalogs. This program does fitting to islands in the image of at least 2×2 pixels in excess of a given threshold. Elliptical Gaussians are then fitted to the islands. The significance of each fitted component was based on the peak value by one or the other of two statistical tests. The first of these tests is that the peak value exceed 5 times the RMS in the pixels within a 121×121 pixel box centered on the component. The alternate test from the estimate of the false detection rate as described above for a given target false detection rate.

C. Comparison with NVSS

The VLSS and NVSS surveys are at different frequencies and different resolutions so a comparison, especially of resolved sources is not always simple. However, in the comparisons made, the test for an NVSS match to a VLSS component was the presence of a valid NVSS component whose centroid was within 60” of the centroid of the VLSS centroid.

A histogram analysis as shown in Figure 2 can be used to select features in the image with a given probability of a false detection. Such a test was performed on the 0800+650 field and features with a false detection rate estimated to be less then 3% in a 401×401 pixel box were fitted by Gaussians and subtracted from the image. A comparison with the NVSS shows that 148/1302 or 11% of the derived catalog entries do not have a nearby counterpart in the NVSS catalog and are unlikely to be real sources. This is substantially higher than the target of 3% false detection rate. The histograms of the residual image are shown in Figure 3. The effects of the inclusions of false sources can be seen in this figure where the positive wing of the distribution is reduced to below the negative wing. For comparison, the 5 σ test identified 890 source of which 3 or 0.3% had no NVSS match within 60”.

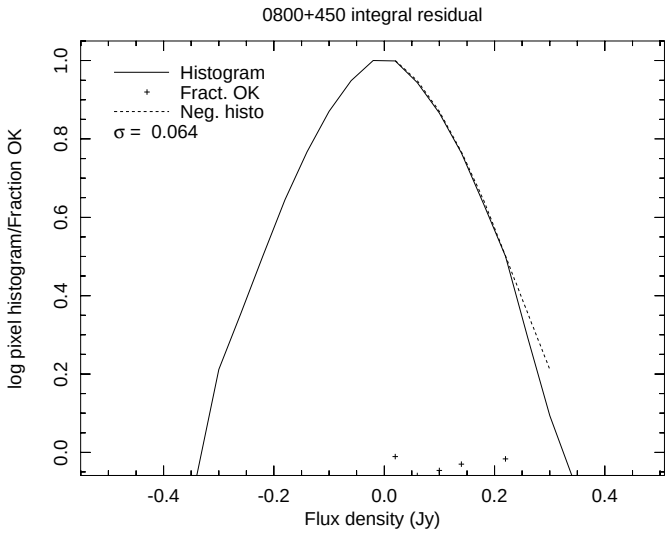

Fig. 3. Like Figure 2 but from the residuals image after the sources in the image used in Figure 2 were subtracted. Spurious sources have reduced the positive wing of the distribution below that of the negative wing.

When the entire survey is cataloged using the 5 σ test, there are 104459 entries of which 5205 (5%) have no NVSS counterpart. The entire survey when cataloged using a 1% false detection threshold produces 117481 entries of which 10907 (9%) have no NVSS counterpart. The 1% FDR catalog has 13000 more entries that the 5 σ catalog of which 44% are unlikely to be real.

D. Effects of Image Point Spread Function

In interferometric images the noise has been passed through the same spatial frequency filter as the sky, so the covariance of the noise will be the same as the point spread function (psf) of the dirty images. However, due to the CLEANing of the images, the dirty psf of the emission is replaced by a Gaussian fit to the core of the dirty psf. In the case of the VLSS, the CLEAN restoring beam is a round Gaussian of 75” FWHM which is approximately that of the “typical” single pointing image. Thus, the noise will not have the same covariance as the CLEANed emission.

Furthermore, there are positive sources well above the noise whose Gaussian representations will add significantly to the number of positive values pixels.

The ratio of the number of positive pixels to the number of sources at the same flux density level can be obtained by comparing histograms of the peak values in cataloged sources to those of the pixels used in the derivation of the catalog. For this test three well behaved fields were used. Source catalogs were derived and a histogram of the peak value formed. A combined histogram of the pixels used was also formed and summarized in table I

The ratio of pixels to sources in flux density bins in Table I in the flux density bins for which the source finding is essentially complete have a ratio of pixels to sources clustered around 16 which is also the beam area in pixels. Note, this is the low flux density tail of distribution of 2948 sources, the maximum of which is 115 Jy.

TABLE I
SOURCE/PIXEL HISTOGRAM COMPARISON

cell	Flux density Jy	pixels	sources	ratio
31	0.2157	83771	66	1269.
32	0.2549	32935	204	161.
33	0.2941	14771	255	57.9
34	0.3333	8061	244	33.0
35	0.3725	5021	229	21.9
36	0.4118	3744	186	20.1
37	0.4510	3033	148	20.5
38	0.4902	2577	135	19.1
39	0.5294	2050	102	20.1
40	0.5686	1750	108	16.2
41	0.6078	1493	105	14.2
42	0.6471	1343	75	17.9
43	0.6863	1137	74	15.4
44	0.7255	1029	52	19.8
45	0.7647	905	65	13.9
46	0.8039	835	46	18.2
47	0.8431	700	42	16.7
48	0.8824	710	40	17.8
49	0.9216	598	40	14.9
50	0.9608	547	36	15.2

VI. DISCUSSION

When applied to VLSS data, the false detection rate achieved is substantially higher than the target value used for the test. The effect is apparently caused by pixels in the skirts of the Gaussian psf of strong sources which increases the ratio of positive to negative pixels at a given flux density level causing the underestimation of the false detection rate. For the whole VLSS survey, cataloging with a target FDR of 1% resulted in an actual FDR of 9%; however, the results show substantial variation among different VLSS fields. Overall, using a false detection threshold in the source cataloging resulted in an increase of 12% in the number of entries of which 44% were likely unreal. For purposes of the VLSS, the 5 σ test is preferred. For better behaved data, this technique may result in more stable results.

REFERENCES

[1] A. M. Hopkins, C. J. Miller, A. J. Connolly, C. Genovese, R. C. Nichol, and L. Wasserman, "A New Source Detection Algorithm Using the False-Discovery Rate," *Astron. J.*, vol. 123, pp. 1086–1094, Feb. 2002.

[2] D. A. Friedenberg and C. R. Genovese, "Straight to the Source: Detecting Aggregate Objects in Astronomical Images with Proper Error Control," *ArXiv e-prints*, Oct. 2009.

[3] A. S. Cohen, W. M. Lane, W. D. Cotton, N. E. Kassim, T. J. W. Lazio, R. A. Perley, J. J. Condon, and W. C. Erickson, "The VLA Low-Frequency Sky Survey," *Astron. J.*, vol. 134, pp. 1245–1262, Sep. 2007.

[4] J. J. Condon, W. D. Cotton, E. W. Greisen, Q. F. Yin, R. A. Perley, G. B. Taylor, and J. J. Broderick, "The NRAO VLA Sky Survey," *Astron. J.*, vol. 115, pp. 1693–1716, May 1998.

[5] W. D. Cotton, "Obit: A Development Environment for Astronomical Algorithms," *PASP*, vol. 120, pp. 439–448, 2008.

[6] W. D. Cotton, J. J. Condon, R. A. Perley, N. Kassim, J. Lazio, A. Cohen, W. Lane, and W. C. Erickson, "Beyond the isoplanatic patch in the VLA Low-frequency Sky Survey," in *Proceedings of the SPIE, Volume 5489.*, 2004, pp. 180–189.

[7] W. D. Cotton, "Lessons from the VLA Long Wavelength Sky Survey (VLSS)," in *Science with Human Wavelengths, Astronomical Society of the Pacific Conference Series, vol, 345*, 2005, pp. 337–340.

[8] H. T. Intema, S. van der Tol, W. D. Cotton, C. A. S., I. M. van Bremmel, and H. J. A. Röttgering, "Ionospheric Calibration of Low Frequency Radio Interferometric Observations using the Peeling Scheme, I. Method Description and First Results," *A&A*, vol. 501, pp. 1185–1205, 2009.