Revised version: submitted to Ap.J.

A Catalog of 1.4 GHz Radio Sources from the FIRST Survey

Richard L. White
Space Telescope Science Institute, 3700 San Martin Dr., Baltimore, MD 21218, rlw@stsci.edu

Robert H. Becker
Physics Dept., University of California, Davis, CA 95616
and IGPP/Lawrence Livermore National Laboratory, bob@igpp.llnl.gov

David J. Helfand
Astronomy Dept., Columbia University, New York, NY 10027, djh@carmen.phys.columbia.edu

Michael D. Gregg
IGPP/Lawrence Livermore National Laboratory, gregg@igpp.llnl.gov

ABSTRACT

We present a catalog of 138,665 radio sources derived from the initial 1550 square degrees of the FIRST survey. The survey parameters are reviewed and a map depicting the coverage for the first two observing sessions is presented. We then describe in detail our algorithm for radio source detection and parameterization, as well as our procedures for constructing the final catalog. The results of extensive tests for astrometric and photometric accuracy, as well as for uncertainties in source extent and morphological characterization are presented. Source positions are all good to better than $1''$, and the flux density scale is accurate to 5%; our elliptical Gaussian fitting procedure is shown both to provide a high-fidelity representation of source morphology for most objects, and to characterize correctly bright sources extended on scales down to $\sim 1/3$ the synthesized beam size of $5\rlap{.}''4$. We use the catalog to construct a $\log N - \log S$ relation for 20-cm radio sources over four decades of flux density, finding excellent agreement with previous determinations from smaller surveys. As one example of the catalog's utility, we present comparisons with the Guide Star Catalog, the IRAS Faint Source catalog, and the ROSAT WGA catalog, identifying radio counterparts to thousands of the objects. Complete instructions for access to this FIRST catalog and subsequent updates thereto are provided.

Subject headings: catalogs — galaxies: quasars: general — infrared: galaxies — X-rays: galaxies — radio continuum: galaxies — radio continuum: stars — surveys

1. Introduction

The FIRST survey of the north Galactic cap began in the spring of 1993. Using the VLA[1] at a frequency of 1.4 GHz, the survey is slated to cover 10,000 square degrees to a sensitivity of $\sim$ 1 mJy with an angular resolution of $\sim 5''$; details are presented in Becker, White & Helfand (1995; hereafter, BWH95). The primary data product of the survey is a set of overlapping images, which are, in turn, coadded to achieve uniform sky coverage. The coadded images are available over the Internet from several sources (see the FIRST home page at `http://sundog.stsci.edu` for details.) However, each such image requires $\sim$ 7 Mbytes of storage in full 32-bit format, hence the set of 5000 images already available is beyond the storage capacity of most users. To make the results of the survey more accessible, we have derived from the images a source catalog which will meet most of the needs of most astronomers.

This paper describes the methodology for generating the radio source list and discusses the reliability and accuracy of the catalog. We begin by presenting the basic parameters of the FIRST images and the coverage the survey has achieved to date (§2). In §3, we describe the algorithm we have developed to extract sources from the images; this is followed (§4) by a description of the catalog generation procedure and an annotated sample catalog page (the catalog itself is available on the World-Wide Web — see below). Section 5 includes a brief summary of the tests we have performed to characterize the accuracy of the astrometric, photometric, and morphological parameters of the catalog entries, and a discussion of the catalog's reliability and completeness. Subsequent sections explore the $\log N - \log S$ relation we observe (§6), and, as an illustration of the catalog's utility, the rate of matches we obtain between FIRST sources and various catalogs from other wavelength regimes (§7). We conclude with a summary that includes instructions for access to the current catalog and subsequent updates (§8).

2. The FIRST Survey Parameters

FIRST was designed to produce the centimeter-wavelength equivalent of the Palomar Observatory Sky Survey. To achieve this aim, it must provide positional accuracy at the sub-arcsecond level for all sources detected, morphological information on scales down to a few arcseconds, and an RMS sensitivity of $\sim$ 0.1 mJy. These goals are met by observing in the B-configuration of the VLA and integrating for 165 s on a grid of densely packed pointing positions designed to provide an efficient, uniform tiling of the sky. Bandwidth synthesis mode is used to achieve the requisite wide-field mapping; the center frequencies for each of two 7-channel bands are 1365 MHz and 1435 MHz. The data obtained are edited, calibrated, self-calibrated, mapped and CLEANed using a set of automated scripts based on the AIPS data reduction package. Maps of the individual point-

---

[1]The National Radio Astronomy Observatory is operated by Associated Universities, Inc., under cooperative agreement with the National Science Foundation.

ings (the "grid" images) are weighted and summed to yield a set of final "coadded" images with 1550×1150 $1''.8$ pixels. (The coadded image sizes will change at the far northern edge of the survey area.) The CLEAN beam size is $5''.4$ FWHM and the peak-to-peak sensitivity variations over each image are $\sim 8\%$. The median RMS for the coadded images is ~ 140 μJy. A complete description of the survey observational and data reduction procedures may be found in BWH95.

FIRST is scheduled to cover the same 10,000 deg^2 of the north Galactic cap slated for inclusion in the Sloan Digital Sky Survey (Gunn & Knapp 1992). Observations began in 1993 April; the catalog described here is derived from the 1993 and 1994 observing sessions which cover a total of $\sim$ 1550 deg^2 of the high Galactic latitude sky between declinations of $+28°$ and $+42°$. The boundaries of this region are slightly irregular and a small fraction ($\sim 3\%$) of the maps are missing owing to various problems with the data acquisition; data to fill these small gaps was acquired in the most recent (autumn 1995) observing session. In addition, a few fields around very bright radio sources suffer a reduction in sensitivity as a result of incompletely CLEANed sidelobes. In order to calculate accurate source surface densities and the survey $\log N - \log S$, as well as to alert users to lower sensitivity regions and gaps, we have calculated a coverage map which provides an effective sensitivity for the survey region with $\sim 3'$ resolution.

The coverage map, including all data obtained in 1993 and 1994 (upon which the catalog presented here is based), is shown in Figure 1(a); an expanded section of the map near the bright radio source 3C286 is shown in Figure 1(b). The coverage map was calculated by establishing an arbitrary grid of positions with an origin at RA(2000) = $07^{\rm h}$, Dec(2000) = $28°$. The database is then queried to find the coadded image which contains a given pixel. The list of grid images which contributed to that sky pixel is extracted from the table attached to each coadded image, and the grid image RMS values, adjusted as described in section 4, are summed in quadrature with the same weighting used to construct the coadded image (BWH95). The result is an estimated RMS sensitivity at each location in the survey area.

The method used to calculate sensitivity in the coverage map is also used to establish a 5σ threshold below which sources are discarded from the catalog (§4)). When there is no source at a given position, the coverage map can be used to determine the maximum brightness of any uncataloged source near that position; including the CLEAN bias (§4.3), $S(lim) = 5\sigma + 0.25$ mJy. Note, however, that this noise estimate is based not on a calculated value for the local RMS in each coadded image, but on the weighted combination of noise values derived from the whole-image RMS for each grid map that contributes to that image. Since the effective noise in a grid map can vary across the image, owing, for example, to incompletely CLEANed bright and/or extended sources, the actual RMS in a particular coadded image may differ by a small amount from the coverage map value. Thus, while this map is a good representation of the overall survey coverage, it should not be used to establish a definitive upper limit to the radio flux density from a given location in the sky; rather, the flux density in the relevant coadded image should be measured directly.

The coverage map RMS sensitivity ranges from 0.075 to 1.86 mJy/beam with a median value of 0.141 mJy/beam over the 1550 deg^2 in the completed survey region; 95% of this area has a

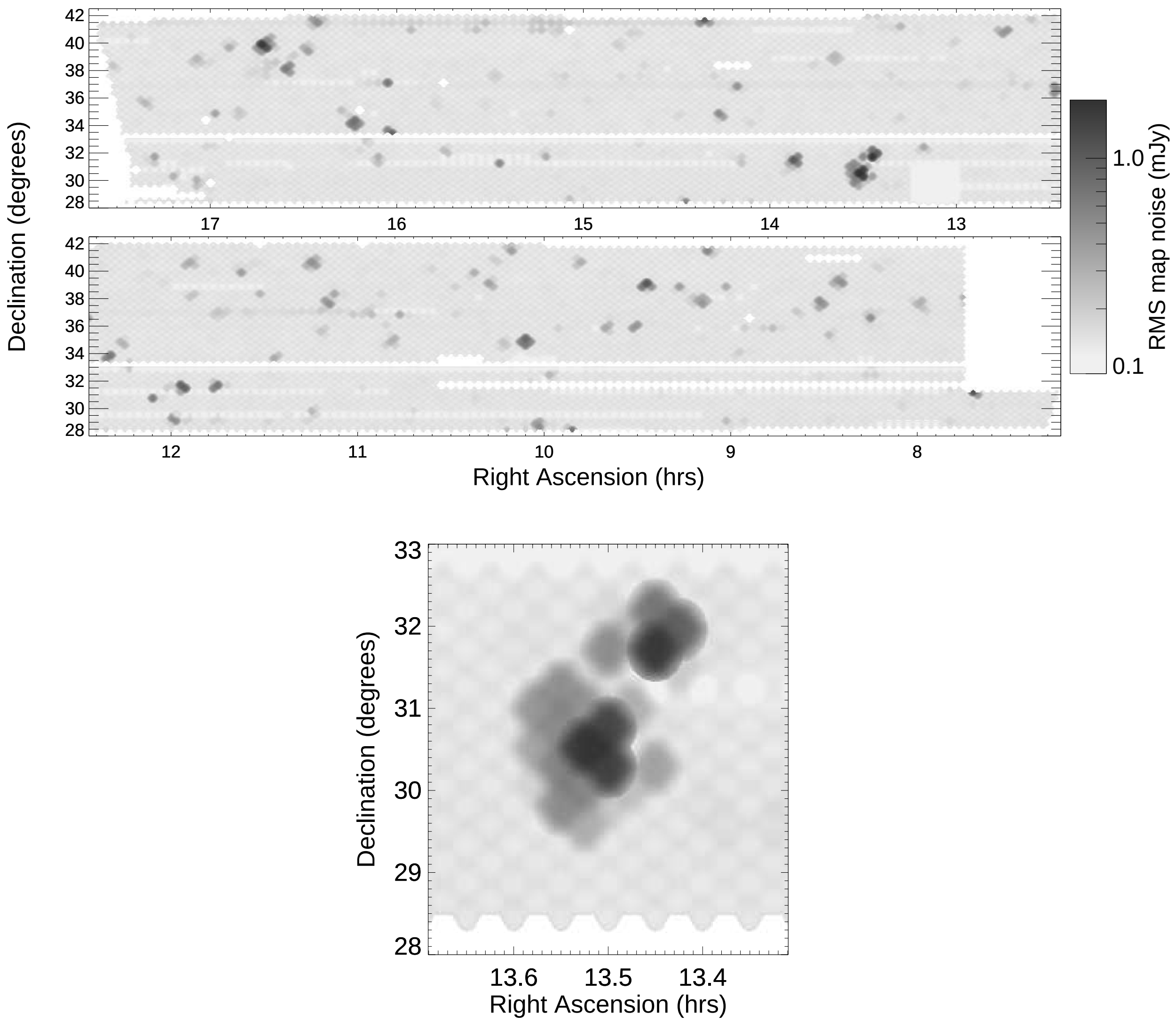

Fig. 1.— RMS sensitivity in the FIRST survey area. Regions of higher noise (dark areas) occur near very bright sources due to incompletely CLEANed sidelobes. (a) Entire area covered during 1993 and 1994 observing campaigns. (b) Blowup of the region around the 14 Jy source 3C286. Note also the scalloped edges of the survey area.

1σ sensitivity of < 0.17 mJy/beam. The map is available through the FIRST WWW homepage (`http://sundog.stsci.edu`).

3. Source Extraction

Deriving a catalog of discrete sources from astronomical images is not a new problem and a number of well-documented computer programs already exist for this purpose (e.g., DAOPHOT). However, most of these programs were developed specifically for extracting point sources from optical images. They are generally not well suited for use with high-resolution radio images in which a significant fraction of the sources are clearly extended. In addition, the noise in radio images has peculiar characteristics that cause trouble when the optical source-finding procedures are used (e.g., it varies from pixel-to-pixel and has some large-scale residual structures remaining after the CLEANing is complete). Furthermore, since all of the images in the FIRST survey are generated in AIPS, we considered it essential to have an AIPS-based source extraction system. The 15JUL93 release of AIPS contained a source extraction program named SAD (for Search and Destroy) which included a number of serious errors, but which served as a good starting point for our own efforts to create a functional AIPS-based routine for source extraction. We call the new task which resulted from our efforts HAPPY (for how we feel about this survey.)

HAPPY starts by searching an image for pixels which exceed a user-provided threshold. It then defines the minimum-size rectangle around each contiguous set of threshold-exceeding pixels. The resulting rectangles, or 'islands', are padded with a border 3 pixels wide. If this border touches another island, any border-strip pixels exceeding the threshold are ignored in the subsequent fit. The pixels in each island are searched for local maxima, which are used as initial estimates for a two-dimensional Gaussian fitting algorithm (similar to the AIPS task JMFIT). Up to four Gaussian components can be fitted to each island. The program is permitted to find up to 10,000 islands per image, although no more than the brightest 2000 islands are actually fit. A typical FIRST survey coadded image has 35 islands.

The individual islands are analyzed in order of decreasing brightness. As the components in each island are fit, the results of the fitting algorithm must pass several acceptance criteria:

- all fitted components must be positive and exceed half the original threshold;
- all fitted peaks must be within the boundary of the island;
- two components may not be separated by < 2 pixels $= 3\farcs6$ (recall that the FIRST survey resolution is $5\farcs4$ FWHM);
- no component can have a peak flux density more than twice that of the highest pixel in the island;

- the sum of the integrated flux densities of all components in the island cannot exceed twice the total flux density in the island or 6 mJy, whichever is greater.

If any of these criteria are not met, the island is refit with one fewer component. Islands are also rejected if they intersect the edge of the image; such areas of sky are always covered more completely in the adjacent image. The results of a successful fit are subtracted from the original image before the next island is fit.

The end product of HAPPY is a list of elliptical Gaussian components. For each component HAPPY gives an RA and Dec, peak and integrated flux densities, major and minor axes, and the position angle of the major axis measured east from north. In addition, an error estimate is calculated for each parameter, and an island number (where island 1 contains the brightest pixel in the field, island 2 the second brightest non-contiguous pixel, etc.) and the number of components within each island are recorded. The peak and integrated flux density for the island are also saved; these quantities may be useful for identifying complex islands that are poorly modeled by the Gaussian components. Some of this information is not included in the currently distributed version of the catalog, but it has not been discarded and may be made available in future versions.

The position angles in the catalog have a somewhat peculiar distribution. Sources that are not detectably elliptical (through a combination of circular symmetry and faintness) have position angles concentrated near 0°, 90°, and 180°. These "magic values" are determined by the initial guess for the position angle for each source and do not have any special significance.

4. Catalog Construction

4.1. Removing duplicates

The FIRST survey creates two types of images, the original grid images from each telescope pointing and the coadded images. HAPPY is run on both sets of images. The primary catalog is based on the coadded images which provide the greatest sensitivity and uniformity. A secondary catalog made from the grid images is used for self-consistency checks between overlapping images and makes possible a search for variable sources (Helfand et al. 1996). We report here on the public version of the catalog which is derived from the coadded images.

The catalog creation process involves deleting duplicate sources and sources likely to be spurious from an initial list of nearly 500,000 sources created by HAPPY. A local RMS noise at each source location is computed via the same procedure used to construct the coverage map (§2): the RMS value for each grid image that contributed at that location is appropriately weighted and summed in quadrature. All sources with fitted peak flux densities fainter than 5 times this RMS value are deleted from the catalog. Sources with fitted peak flux densities fainter than 0.75 mJy/beam are also deleted; this produces a catalog with a more uniform lower limit that is not affected by the small RMS variations among the vast majority of the images.

Sources with fitted minor axes smaller than $3\farcs5$ are also deleted from the catalog. Examination of these sources revealed that essentially all such "skinny" sources, which are considerably narrower than the $5.4''$ beam, are sidelobes of nearby bright sources.

The next step is to delete duplicate entries of sources fitted in overlapping regions of the coadded maps. Since the overlapping regions of coadded images are practically identical, having been formed from the same set of grid images with at most a fraction of a pixel shift, the identification of duplicate sources is in most cases straightforward. In a small number of cases, a source may wind up with substantially different parameters in different fields (e.g., because it is fit as a single Gaussian in one field and as a double Gaussian in the other). Groups are formed by finding pairs of sources that match within $5''$, and then merging all pairs with common sources into larger groups.

Once a group of duplicate sources has been identified, two numbers used for selection are required for each source in the group. The "island count" is the number of components (1 to 4) that were simultaneously fitted in the source's island by HAPPY. (Note that usually all components from an island do not fall in a single group.) The "field count" is the number of sources in the group that come from the same field (coadded image) as this source. We adopt the following criteria to select the source(s) to retain from a given duplicate group:

1. The smallest island count for the group is determined. All sources with larger island counts are discarded from the catalog.
2. The smallest field count for the group is determined. All sources with larger field counts are discarded from the catalog.
3. Finally, the distance of each source from its field center is determined. The closest source to field center and all other sources from the same field are retained; all sources from other fields are discarded from the catalog.

For the vast majority of sources, there is only a single source from each field and only the third criterion listed above has any effect: we simply choose the source closest to field center. For more complicated cases, this procedure was designed to favor simple source models over complex models.

4.2. Identifying spurious sources

Applying these procedures to the complete HAPPY output list results in a catalog containing 138,665 discrete entries. This catalog has no duplicate entries and is complete to the greater of 0.75 mJy/beam or the 5σ flux level indicated by the coverage map. It does, however, contain a small number of spurious sources that are sidelobes of imperfectly CLEANed, nearby, bright sources. Examining all of the sources by eye is clearly impractical and error prone, so we have adopted a machine learning method to identify and flag as many of the spurious sources as possible. We examined several hundred fields containing bright sources and by eye marked catalog sources that appeared to be sidelobes. The complete list of sources in these fields, marked and unmarked, were used as a training set for the oblique decision tree program OC1 (Murthy, Kasif, & Salzberg 1994; available via anonymous ftp from `http://www.cs.jhu.edu/šalzberg/announce-oc1.html`). For each object in the training set, we supplied the catalog parameters (peak flux density, sizes, etc.) and also parameters giving the brightness, distance, and direction to nearby bright sources. The decision tree program then attempted to identify a series of tests based on linear combinations of the parameters that would partition the training set objects into sidelobes and non-sidelobes. See Murthy et al. for more details on oblique decision tree methods; Salzberg et al. 1995 describe another application of this method to astronomical data.

After a reliable, reasonably simple decision tree had been developed, it was applied to all the sources in the catalog to determine which objects might be sidelobes. In the training set, the selected tree (which required only 5 decision planes) successfully identified $\sim 80\%$ of the sidelobes while incorrectly flagging $\lesssim 1\%$ of real sources. Even at this level of accuracy, though, a substantial fraction ($\sim$10–20%) of the flagged objects are *not* sidelobes; consequently, we chose not to delete such objects from the catalog but rather to include a sidelobe warning flag for all selected sources. Examination of the images is necessary to determine with certainty whether any particular flagged source is a sidelobe or not.

In all, 4,813 sources (3.5% of the total) are flagged as possible sidelobes in the catalog. We estimate that $\sim$ 500–1000 of these are real radio sources and that an additional $\sim$ 500 unflagged sidelobes remain in the catalog. The resulting FIRST catalog, presented below, contains 138,665 discrete entries, of which 133,852 are unflagged. We discuss the accuracy of the source parameters and the completeness of the catalog in subsequent sections.

4.3. CLEAN bias correction

All images made from VLA snapshots suffer from a photometric defect which has come to be known as 'CLEAN bias' (see Condon et al. 1994, BWH95). As a consequence, the flux density of all sources in FIRST images are underestimated. One feature of the CLEAN bias not recognized in our initial paper (BWH95) is that the magnitude of the effect varies with off-axis angle within a single grid image, peaking at field center and decreasing monotonically with increasing distance from field center. The functional form of this radial dependence is remarkably similar to the shape of the primary beam response (Fig. 2). At field center, the peak flux density of sources is decreased by ~ 0.25 mJy/beam, independent of source size or brightness. The integrated flux density is reduced by the same *percentage* as the peak flux density, making the effect more serious for extended sources which cover several beam areas. In the preliminary catalog released in 1995 January, the reported flux densities were not corrected for this effect. In all versions of the catalog released after 1995 October, the flux densities are corrected.

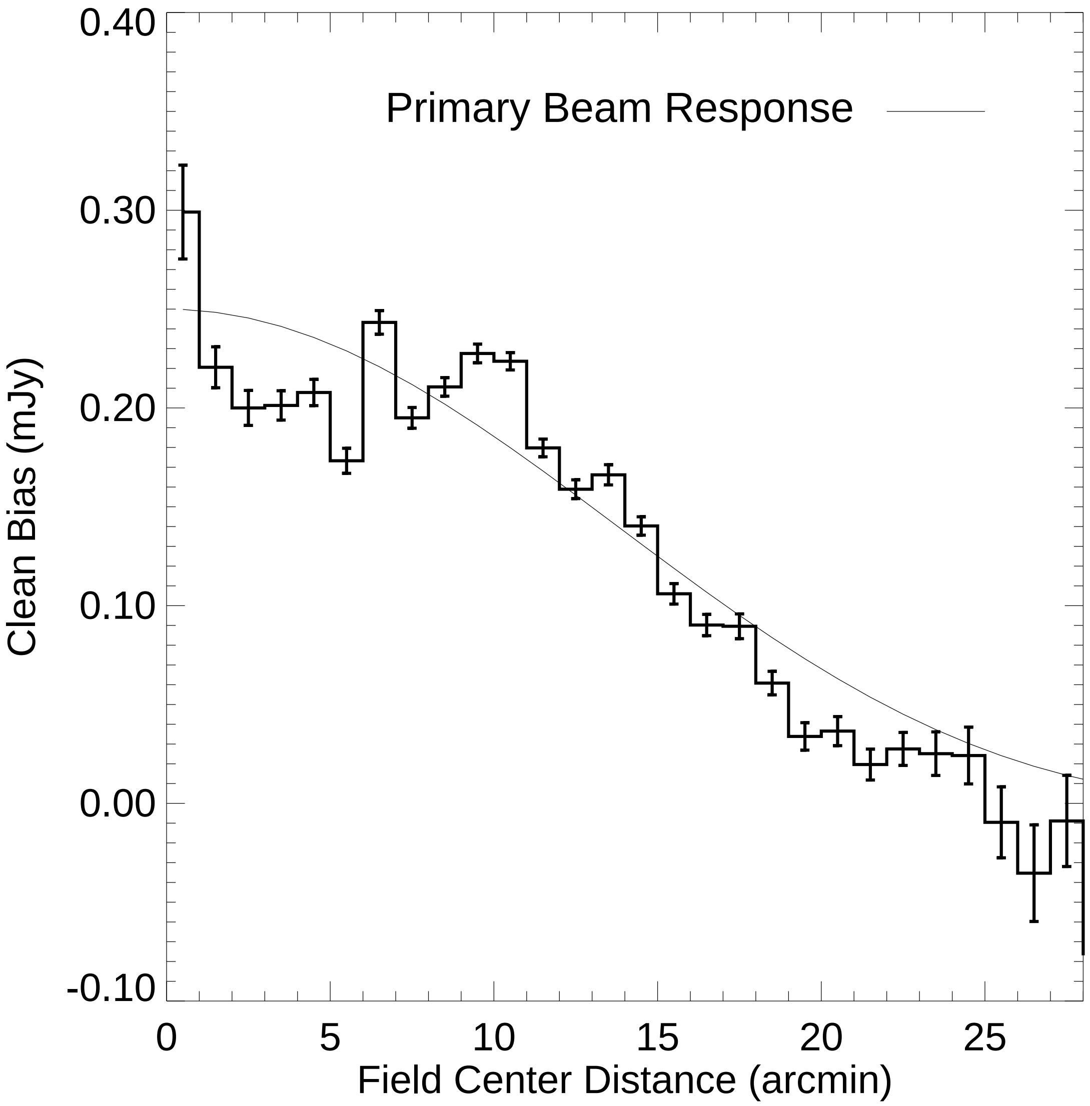

Fig. 2.— Bias in the peak flux density as a function of the distance of a source from field center in the CLEANed grid images, derived from artificial sources inserted in the data and comparisons of sources observed in overlapping grid images. The bias is the value that must be added to the peak flux density to get the correct flux. The shape of the primary beam response is also shown; because the bias is essentially identical in shape to the beam, the bias in coadded images is uniform across the field.

This correction is made simpler by the radial dependence of the CLEAN bias. When the individual grid images are corrected for the primary beam response to create the coadded images, the radial dependence disappears; consequently the bias in the coadded images of ~ 0.25 mJy/beam (BWH95) is independent of source position.

Unfortunately, the magnitude of the bias also depends on the image RMS and can be twice as large in very noisy images. High RMS fields are most often the result of the presence of a very bright source ($S > 1$ Jy) whose sidelobes are not adequately removed from the image. Similar problems are associated with very extended sources which are also CLEANed incompletely. The distribution of grid image RMS values, shown in Figure 3, demonstrates that significant CLEAN bias underestimates are rare. Only 2% of our 11,000 grid images have RMS values as much as 30% higher than the median RMS. These problematic sky positions can be readily identified from the survey coverage map.

4.4. The catalog format

Table 1 is a sample from the on-line FIRST catalog. The column entries are as follows:

1,2: The right ascension and declination of the source (J2000); position errors are dependent on source size and extent (see §5.4).

3: A 'W' in this column is a warning flag that the source may be a sidelobe of a nearby bright source (see §4.2).

4,5: The peak and integrated flux densities (S_p, S_i, quoted in mJy/beam and mJy respectively) derived from the elliptical Gaussian fits (§3) and corrected for CLEAN bias (§4.3).

6: The local noise rms estimate ($\sigma(S)$, mJy/beam) derived as described in §2. Note that the source detection significance is $(S_p - 0.25)/\sigma(S)$, not $S_p/\sigma(S)$, because of the CLEAN bias correction to the peak flux density.

7–9: The major and minor axes (FWHM in arcseconds) and the position angle (degrees east of north) of the fitted Gaussian with the $5\farcs4$ synthesized beam size subtracted in quadrature. Noise can cause the fitted values of the axes before deconvolution to be smaller than the beam; to preserve this information in the catalog, the listed sizes are negative when the fitted size is smaller than the beam. When the deconvolved size quoted in the catalog $\psi \geq 0$, the fitted size was $\psi_F = (\psi^2 + 5.4^2)^{1/2}$; when $\psi < 0$, $\psi_F = (5.4^2 - \psi^2)^{1/2}$. For most purposes, the best estimate of the true source size is either the listed size or zero, whichever is greater.

10: The field name of the coadded image from which the catalog entry was derived (note that a given source may also be found in other images); the field name encodes the position of the field center as `hhmmm+ddmmm` where RA = hh mm.m and Dec = dd mm.m. Field names ending

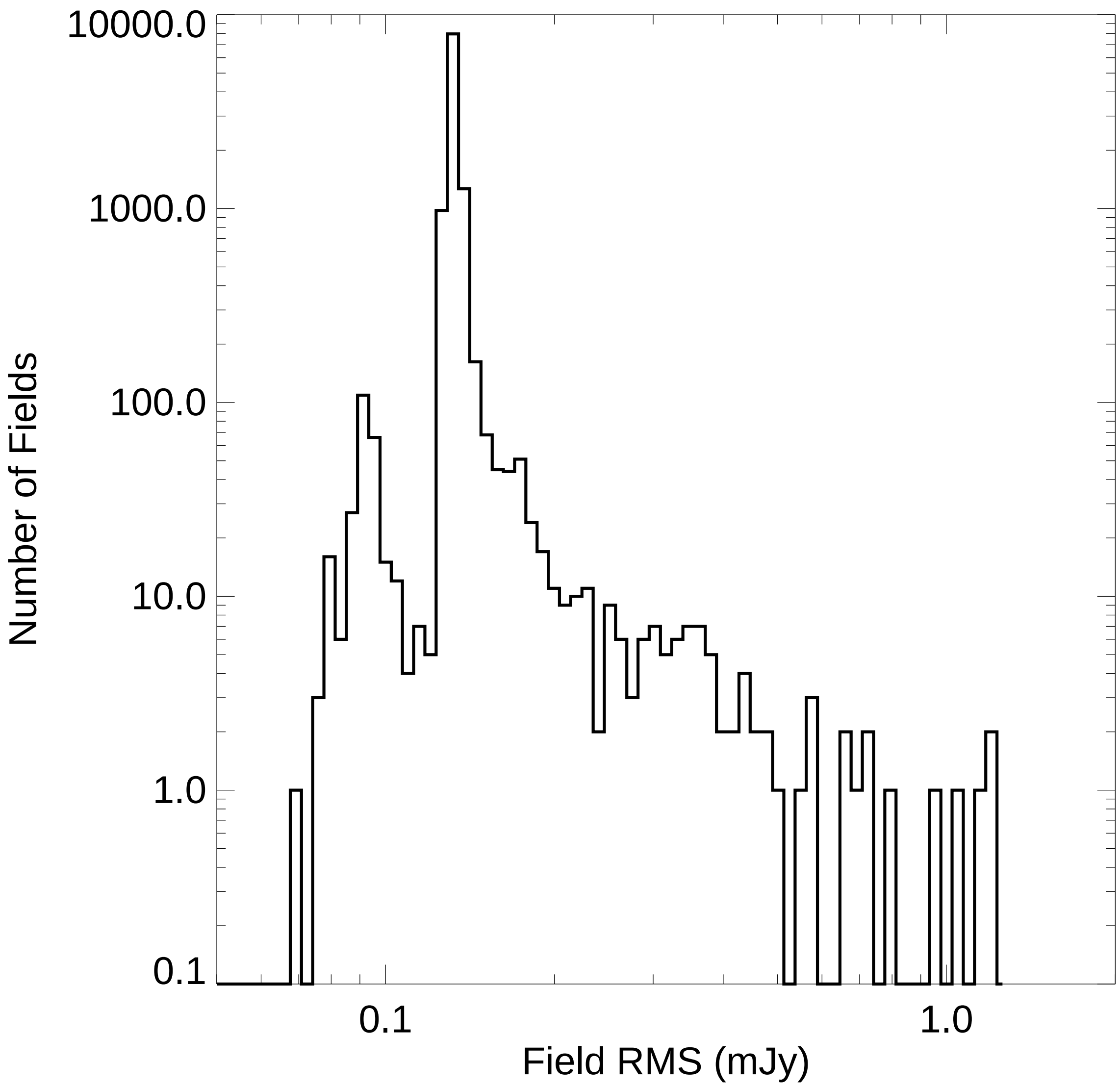

Fig. 3.— Histogram of the RMS noise in grid images. The vast majority of fields have RMS values ~ 0.13 mJy/beam. The secondary peak around 0.09 mJy/beam is due to fields that were observed twice. Fields containing very bright sources account for the tail at higher RMS values.

TABLE 1

Sample Page from the FIRST Catalog

RA (2000) (1)	Dec (2000) (2)	W[a] (3)	S_p[b] (4)	S_i[c] (5)	RMS[d] (6)	Maj[e] (7)	Min[e] (8)	PA[f] (9)	Field[g] (10)
08 00 00.099	+32 06 09.83		1.11	1.07	0.138	2.8	−2.8	173.7	08000+32118B
08 00 00.250	+38 14 21.10		8.99	12.35	0.143	4.1	2.4	23.8	08000+38071B
08 00 00.453	+34 52 18.24	W	1.13	2.17	0.157	11.9	−3.3	91.9	08000+35071B
08 00 00.591	+39 50 49.00		2.03	1.31	0.148	−2.3	−3.8	163.1	08000+39393B
08 00 00.653	+37 33 50.90		22.01	21.91	0.266	0.0	−1.0	34.4	08000+37367B
08 00 00.752	+32 46 50.88		4.24	3.95	0.106	−1.0	−1.8	66.2	08000+32407B
08 00 02.306	+31 17 20.94		2.19	2.75	0.095	3.3	2.0	170.2	08000+31143B
08 00 02.515	+35 49 27.15		1.11	0.85	0.153	1.5	−3.6	133.8	08000+36065B
08 00 03.247	+36 29 42.13		46.53	58.70	0.138	3.7	1.5	21.6	08000+36364B
08 00 03.683	+33 58 24.35		1.50	1.08	0.144	−2.8	−2.8	0.3	08000+34082B
08 00 04.079	+32 02 09.58		40.37	54.07	0.155	4.3	1.8	133.9	08000+32118B
08 00 04.348	+31 23 15.24		4.16	4.98	0.115	3.6	0.0	126.8	08000+31143B
08 00 05.466	+32 11 23.16		2.95	3.17	0.133	2.6	−1.5	26.1	08000+32118B
08 00 05.859	+38 19 21.50	W	1.04	9.24	0.155	19.3	11.8	102.0	08000+38071B
08 00 06.096	+34 04 34.60		1.30	1.62	0.132	3.6	1.5	79.5	08000+34082B
08 00 06.436	+34 58 02.24		1.11	1.69	0.146	5.6	2.0	107.4	08000+35071B
08 00 06.555	+32 01 48.05		1.11	0.75	0.157	−2.0	−3.6	149.9	08000+32118B
08 00 06.574	+34 04 36.10		1.39	1.45	0.132	4.5	−3.1	98.9	08000+34082B
08 00 06.591	+39 38 05.90		1.32	2.02	0.132	6.7	−1.5	32.7	08000+39393B
08 00 07.046	+29 36 56.93		8.97	13.42	0.151	4.2	3.4	145.2	08000+29486A
08 00 07.118	+34 15 07.26		1.17	0.89	0.137	0.0	−3.5	18.1	08000+34082B
08 00 07.145	+33 40 05.92		7.69	9.06	0.129	2.7	1.8	139.7	08000+33389B
08 00 07.371	+38 07 34.51		7.20	8.05	0.138	2.3	1.3	176.0	08000+38071B
08 00 07.745	+34 06 14.68		1.28	1.07	0.133	2.1	−3.4	133.0	08000+34082B
08 00 07.872	+41 13 06.70		3.76	3.71	0.126	0.0	−1.0	0.0	08000+41132B
08 00 07.954	+36 04 08.82		4.83	5.08	0.136	1.5	0.9	62.0	08000+36065B
08 00 08.817	+38 46 23.73		1.15	0.96	0.142	1.5	−3.3	17.2	08000+38377B
08 00 09.145	+36 04 07.45		2.00	4.64	0.137	7.1	5.3	56.4	08000+36065B
08 00 09.623	+38 06 49.86		1.08	1.30	0.138	2.6	2.1	180.0	08000+38071B
08 00 09.689	+33 59 56.95		6.11	8.32	0.140	4.1	2.3	53.3	08000+34082B
08 00 10.125	+30 31 51.56		1.31	0.88	0.156	−1.5	−3.8	151.5	08000+30456B
08 00 10.262	+41 20 36.81		1.36	0.80	0.133	−3.4	−3.6	94.1	08000+41132B
08 00 10.335	+35 57 41.31		1.09	1.54	0.145	5.8	−1.5	79.8	08000+36065B
08 00 10.567	+34 02 20.85		1.59	3.45	0.137	8.9	2.8	178.8	08000+34082B
08 00 10.613	+28 51 23.53		1.88	4.94	0.095	11.8	2.4	159.4	08000+28521B
08 00 10.794	+30 36 06.29		9.51	10.03	0.148	1.5	0.9	66.0	08000+30456B
08 00 11.197	+30 31 48.01		25.66	25.99	0.156	0.8	0.2	117.0	08000+30456B
08 00 11.282	+34 00 07.72		3.23	3.77	0.140	2.7	1.6	89.7	08000+34082B
08 00 11.479	+30 14 34.11		3.13	2.72	0.136	1.0	−2.8	141.0	08000+30170A
08 00 11.495	+35 50 17.56		1.10	0.80	0.154	−2.0	−3.4	176.0	08000+36065B
08 00 12.075	+30 48 48.79		3.02	2.74	0.138	−1.5	−2.0	0.9	08000+30456B
08 00 12.089	+30 05 16.34		9.89	10.43	0.151	1.4	1.1	90.0	08000+30170A
08 00 12.833	+29 23 46.99		1.09	1.35	0.138	5.2	−2.5	67.2	08000+29203A
08 00 14.079	+41 46 12.10		3.09	3.05	0.133	1.5	−1.5	17.7	08000+41450B
08 00 14.111	+41 30 47.54		1.71	2.48	0.146	5.0	2.0	97.9	08000+41450B
08 00 15.829	+35 44 29.46		3.79	4.41	0.141	2.8	1.3	135.9	08000+35367B
08 00 16.065	+40 29 56.01		189.37	199.72	0.145	1.3	1.3	90.0	08000+40417B
08 00 16.990	+32 14 50.34		3.47	3.42	0.135	1.8	−2.0	163.3	08000+32118B
08 00 17.075	+35 42 42.29		3.88	3.93	0.139	1.0	0.0	156.5	08000+35367B
08 00 17.567	+37 02 59.81		33.55	34.94	0.138	1.2	0.9	180.0	08000+37065B

[a] Warning flag 'W' indicates source is possibly a sidelobe of a nearby, bright source.

[b] Peak flux density (mJy/beam) at 1.4 GHz for Gaussian fit to source.

[c] Integrated flux density (mJy) for Gaussian fit to source.

[d] Estimated RMS noise (mJy/beam) in coadded image at source position.

[e] Major and minor axes FWHM (arcsec) for source after 5.″4 synthesized beam size has been deconvolved. Negative values indicate that source was smaller than beam and allow fitted size to be recovered (see text.)

[f] Position angle of major axis (degrees east of north.)

[g] Name of coadded image including this source.

with an 'A' were constructed entirely from 1993 observations, while those with a 'B' include at least some data from the 1994 session. This nomenclature will be extended in the obvious way in the years to come, with 'C' referring to data from 1995, etc. This suffix provides a rough indication of the epoch of the observation, although some 'B' fields do contain data from both 1993 and 1994.

The catalog from the coadded images obtained in 1993 was released in preliminary form (95jan06 version) through our WWW page (`http://sundog.stsci.edu`). The release described herein (95oct16 version) includes 138,665 sources from an area of 1550 deg^2.

5. Catalog Accuracy

Sources which appear in more than one grid image (multiply observed sources) have been used extensively for image quality control (BWH95). In our earlier paper, we demonstrated that source positions, flux densities, and morphologies are reproducible from image to image. In general, source positions down to the survey threshold were found to be internally consistent to better than $1''$ (90% confidence). For brighter sources (flux density > 5 mJy) fitted source sizes were self-consistent to better than $0\farcs5$. Flux densities for point sources were also found to be highly reproducible and, when corrected for CLEAN bias as described above, gave results consistent with a deep VLA pointing which fell within the survey area. Below, we present new information on the accuracy of catalog parameters, concluding with a summary of the uncertainties which should be adopted when utilizing the FIRST catalog.

5.1. Astrometry

The utility of the high quality astrometry we infer from the multiply observed source analysis is dependent on tying FIRST positions to an absolute reference frame. The best set of radio astrometric standards currently available are the MERLIN calibrators (Patnaik, private communication). In our original paper, we reported a mean offset for 46 objects which fell within our initial 300 deg^2 of $\lesssim 0\farcs05$. However, the MERLIN calibrators were originally chosen to be bright and predominantly point-like at 3.7 cm. Closer examination reveals that a significant fraction of these objects are extended and/or have complex morphologies in the 20 cm FIRST images. Of the 115 MERLIN calibrators which fall within the current survey area, 79 have measured extents of $\leq 2\farcs0$ and are the only component within their HAPPY island. Employing only these objects in the comparison eliminates all but one object[2] with a discrepancy of $> 0\farcs50$ (Fig. 4); the mean offsets in RA and Dec are $0\farcs015 \pm 0\farcs019$ and $0\farcs024 \pm 0\farcs016$, respectively. The RMS position error

[2]This object is near the edge of the survey area and was observed far off-axis in only a single field. The position error can be attributed to a recently discovered, very small rotation and scale error in the FIRST grid images that

for this set of 79 bright, point-like sources is $0\rlap{.}''22$, and there are no systematic trends with RA or Dec with an amplitude as large as $0\rlap{.}''1$.

Johnston et al. 1995 have recently published a definitive sample of 436 extragalactic radio source position calibrators with positional accuracies of better than $0\rlap{.}''003$ in both coordinates. Twenty-eight of these sources lie within the FIRST coverage region and 24 have FIRST sizes of $< 2''$, only 9 of which are in the MERLIN sample. For this independent set of calibrators, we find mean offsets in RA and Dec of $0\rlap{.}''012$ and $0\rlap{.}''003$, respectively, and an RMS position offset of $0\rlap{.}''10$. Thus, we are confident that the FIRST positions are tied to the radio reference frame with a systematic uncertainty of $< 0\rlap{.}''03$, and that systematic position errors are $< 0\rlap{.}''1$ throughout the survey region.

The linkage between the radio and optical reference frames has been the subject of considerable recent research (Argue et al. 1984; Lindegren & Kovalevsky 1995; Johnston et al. 1995 and references therein). We will discuss an empirical comparison of FIRST sources to the APM POSS-I plate scan positions in a subsequent paper (McMahon et al. 1996). As we have found a number of systematic discrepancies at the subarcsecond level when comparing FIRST sources to a variety of plate-scan catalogs, we advise workers requiring this level of accuracy to investigate empirically any systematic differences which may be present. We also caution stellar radio astronomers that the possibility of significant proper motions must be considered when searching for stellar counterparts to FIRST sources (Becker et al. 1996).

5.2. Morphology

As noted in BWH95, the multiply observed sources demonstrate that our procedures to determine source sizes are consistent from image to image. Figure 5 displays the distribution of major axis values for all 138,665 sources in the catalog. Over 29,000 of the sources are larger than the $5\rlap{.}''4$ beam size and are clearly resolved. For comparison, this figure also shows the expected distribution based on the flux-dependent model of 1.4 GHz source sizes from Windhorst et al. 1990 (see §6 for the details of this model.) The observed FIRST size distribution is in reasonably good agreement with the model; the discrepancy at larger sizes is due to double sources that appear in our catalog as two smaller components rather than as a single extended object. Pairs of objects separated by $< 50''$ have been merged in the Windhorst et al. 1990 distribution.

All extended sources are represented in the catalog as one or more elliptical Gaussians; in some cases this may not provide a high-fidelity representation of the complete source morphology,

---

can produce offsets of $\lesssim 0\rlap{.}''3$ at the edges of grid images. In the coadded images, the resulting position errors cancel and so are extremely small except at the edge of the survey area. The source of the rotation and scaling is under investigation; we speculate that they are due to a combination of frequency shifts from bandpass shape effects and VLA clock errors.

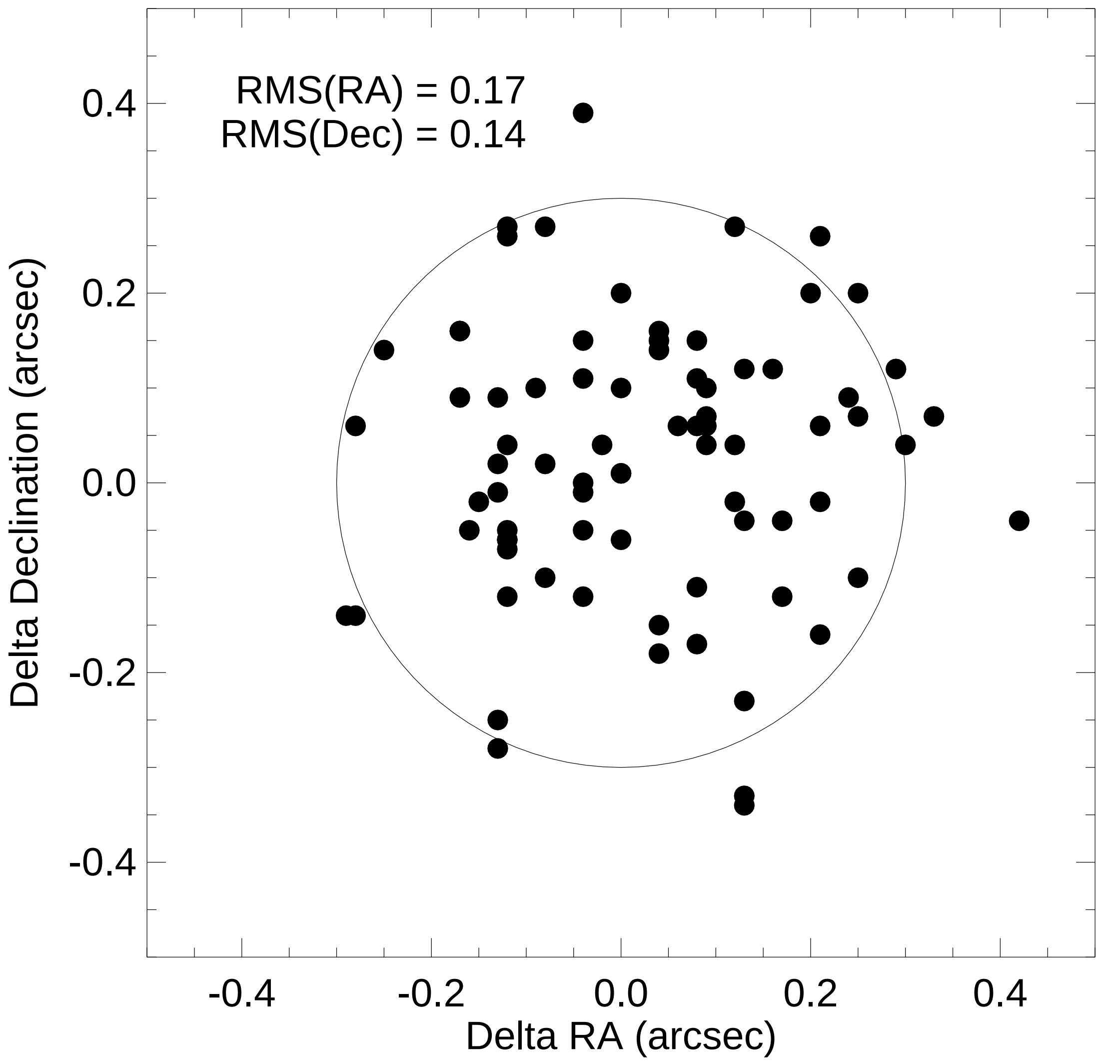

Fig. 4.— Differences between FIRST positions and MERLIN astrometric calibration standard positions. The FIRST positions show excellent astrometric accuracy with no signs of any systematic offsets compared to the MERLIN standards.

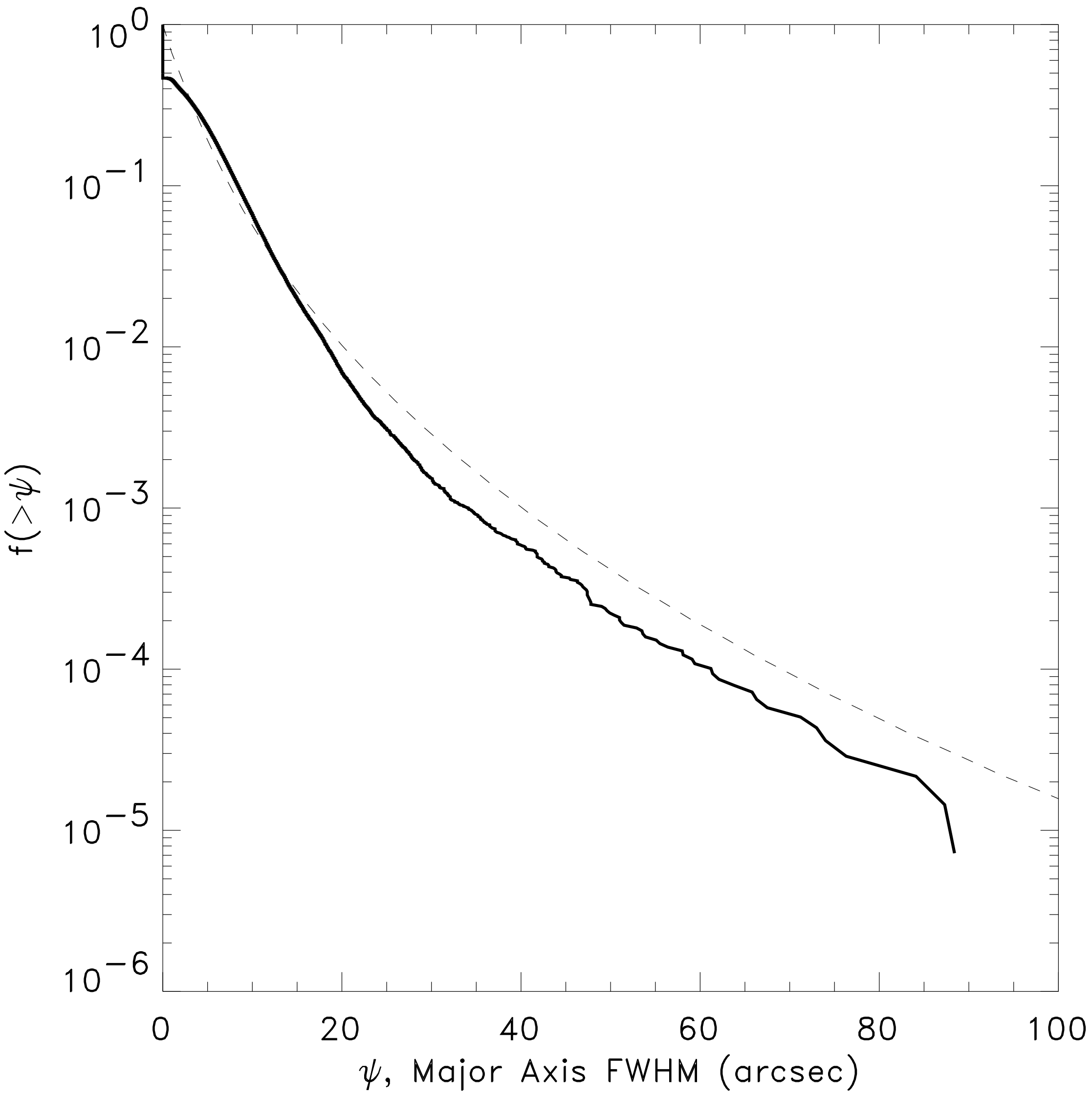

Fig. 5.— Solid: Cumulative distribution of major axis FWHM values for all 138,665 sources in the FIRST source catalog. Dashed: distribution expected from the Windhorst et al. (1990) flux-dependent size distribution. The discrepancy at larger sizes is due to double sources that appear in the FIRST catalog as two smaller components rather than as a single extended object.

but in most instances such a representation is adequate. Figure 6 displays an example of a pair of complex sources. The left panel is the FIRST survey image, while the right panel shows the Gaussian representation of these sources as they appear in the FIRST catalog. The Gaussian deconvolution does an excellent job of capturing the morphology of these complex sources, including the bent-double morphology of the southeastern source and the peculiar ring-like morphology of the brighter object. From this and other examples we have examined, a computer vision approach could probably be used to select such objects directly from the catalog. However, certain detailed features, such as the slight asymmetry of the southernmost lobe, could only be revealed through visual inspection of the original images.

Large extended and complex sources such as those discussed above make up a relatively small fraction of the total catalog. A larger fraction is comprised of sources having measured sizes between $1''$ and $5''$ after deconvolution of the $5\farcs4$ restoring beam. Both for purposes of deriving radio source statistics and for optical identification programs, it is important to determine the degree to which the sizes and position angles of the fitted ellipses for these sources reflect the true underlying source morphology. To investigate this, we have utilized the A-configuration snapshot survey of unresolved Green Bank 6 cm radio sources brighter than 50 mJy carried out by Lehar (1991; hereafter L91) as part of a search for gravitational lenses. Contour plots were generated for each of the 178 sources from that project which fell within the FIRST survey region and a detailed comparison of source positions, extents, and orientations was carried out. Despite the wide disparity in observing frequency (20 cm vs. 3.6 cm) and spatial frequency sensitivity (the maximum baseline was 9 times greater as measured in wavelengths for the Lehar survey), this clearly demonstrates that the FIRST catalog can accurately describe structure on scales smaller than the $5\farcs4$ beam.

Eighty-one of the 178 sources are listed as point-like in L91. We find that 52 of these are, indeed, isolated objects substantially smaller than our beam, and we use them to determine empirically the limit to which we can resolve source structure (the other 29 L91 point sources are the cores of complex radio sources with total angular extents in our maps of $0\farcm2$ to $2'$ and are not useful for this purpose). Of the 52 isolated point sources, 44 have measured major axes $< 1\farcs5$ (51 of 52 have minor axes in this range), and all but three are smaller than $1\farcs8$. The sources with the three largest measured sizes all show evidence of a core-jet morphology, and it is likely that the steeper spectrum extended emission is undetected in the L91 high-frequency maps; i.e., the indications of extent in the FIRST images are probably real. Thus, in all future work, we will adopt a major axis upper limit of $2\farcs0$ as defining the point-source population of the FIRST catalog.[3]

Of the remaining 92 objects for which high resolution maps are available, 15 are classified as faint or undetected in L91. We find that all of these correspond to extended components of complex sources, and have simply been resolved out in the 3.6 cm observations. For the other 77 objects,

---

[3]The position offsets of all but one of the sources are substantially less than $1''$, yet another confirmation of the FIRST astrometric accuracy. In addition, 24 of the 29 sources corresponding to the cores of complex objects also have offsets $< 1''$, suggesting optical identification of such objects from FIRST positions will be straightforward.

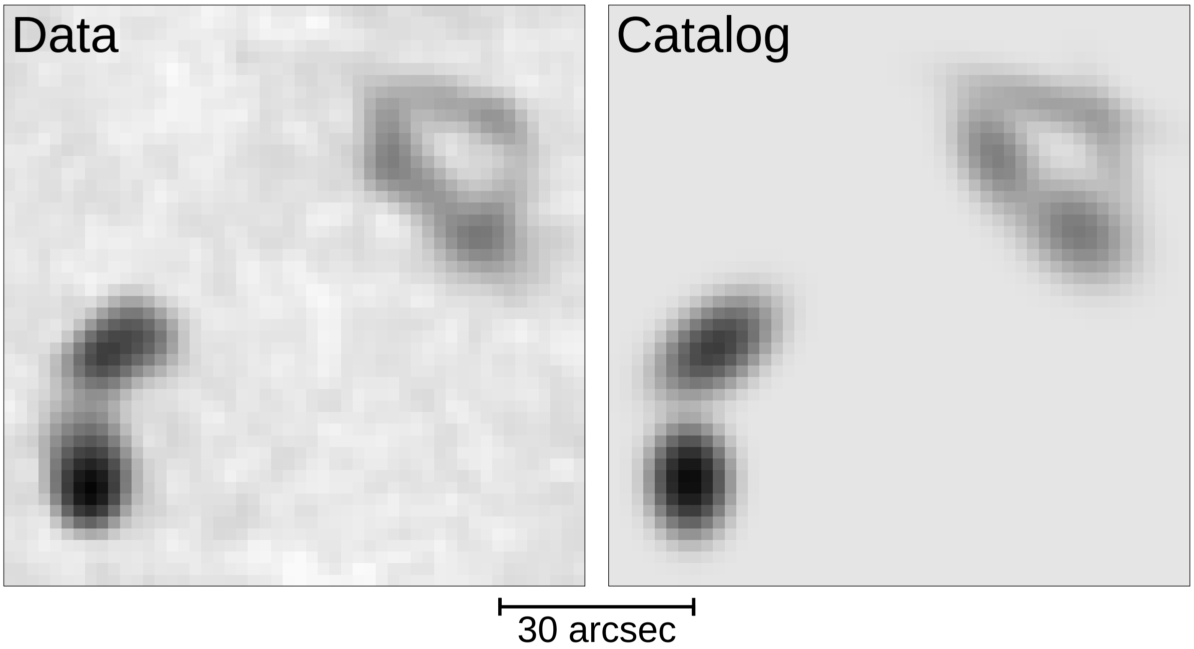

Fig. 6.— Comparison of a FIRST image (left) with the elliptical Gaussian model from the FIRST catalog (right) for two complex sources. The field is centered at RA = $10^{\rm h}50^{\rm m}08.5^{\rm s}$, Dec = $+30°40'15''$ (2000). The catalog contains 6 components, two for the bent double and four for the ring-like object. While not a perfect model for this complex source, the catalog components clearly give a very useful, high-fidelity representation of the objects.

L91 presents maps with a resolution of $0''.24$ which reveal a host of complex source morphologies; detailed comparison of these with the corresponding FIRST images allows us to assess further the utility of our catalog's morphological parameters at size scales smaller than $10''$.

Most of these objects exhibit the classical double-lobe radio source structure with or without a core component; the remainder can be best described as having core-jet morphologies. In 57 of the 77 cases, the FIRST sizes as determined from our component fitting are within 25% of the extents measured in the high frequency images which have twenty times the resolution. In most of the remaining cases, the FIRST sizes are somewhat larger, most likely reflecting the presence of additional steep-spectrum extended emission which is undetected in the 3.6 cm maps. In addition, the FIRST position angles provide an excellent indication of source orientations: in 31 of 34 cases with linearly aligned components displaying a maximum extent of $< 6''$, the FIRST position angle differs from that of the resolved linear structure by $< 20°$. This excellent agreement of source extent and orientation extends down to sizes of $2''$. Our ability to extract accurate morphological parameters down to this size scale both provides another confirmation of FIRST map fidelity and assures that efficient optical identification programs can be carried out for both point-like and extended objects.

5.3. Photometry

In BWH95 there was an extensive discussion of FIRST's photometric accuracy based on internal comparisons using multiply observed sources and external comparisons using a few previously published deep VLA pointings. In §4.3, we have updated our understanding of the 'CLEAN bias', an effect that appears to reduce the peak flux density of all sources in the FIRST catalog by ~ 0.25 mJy/beam. Here, we extend our evaluation of FIRST photometric accuracy by comparing FIRST flux densities to the NRAO VLA Sky Survey (NVSS; Condon et al. 1994). which is also now being carried out using the VLA at a wavelength of 20 cm. At the current time, there is an overlap of ~ 400 deg^2 between the two surveys in the RA range $7.5^{\rm h}$–$10^{\rm h}$.

The primary difficulty in comparing flux densities between the NVSS and FIRST surveys is the factor of 8 difference in angular resolution ($5''.4$ vs. $45''$ FWHM). The high resolution FIRST survey will partially resolve extended sources that appear point-like in the NVSS, so the flux density of extended sources will be smaller in FIRST than in the NVSS. On the other hand, the NVSS will fail to resolve distinct sources closer together than $50''$, further confusing the comparison between the two surveys. To eliminate this latter effect, we restrict the comparison to isolated FIRST sources whose nearest neighbor in the FIRST catalog is at least $100''$ distant. This reduces the number of sources common to the two surveys to 7,318.

In Figure 7 we plot the ratio of the flux densities from the two surveys as a function of the source major axis FWHM as measured by NVSS. For the smallest sources in the survey ($< 2''$ FWHM), the mean ratio is 1.000 ± 0.010. This includes corrections for CLEAN bias of 0.25 and

0.40 mJy/beam for FIRST and NVSS respectively. The mean in each bin was calculated by weighting each of the source pairs according to the noise in the ratio, with the noise parameters taken from this paper for FIRST and from the 1995 August 16 draft of the NVSS source catalog description (`ftp://gibbon.cv.nrao.edu/pub/nvss/catalog.ps`). As expected at the outset of the survey, FIRST partially resolves sources much larger than our $5''\!.4$ beam and underestimates their flux densities.

Many (perhaps most) extended NVSS sources are resolved into several components by FIRST, with the total flux lost from the sum of the components being less than would be lost from a single, very extended Gaussian component. Such multi-component matches are precluded in this comparison by our requirement that the FIRST source be isolated. As a result, for a randomly selected NVSS source having a given size, Figure 7 represents an upper limit for the amount of flux lost in the FIRST catalog. On the other hand, this distribution should accurately describe the flux lost from each FIRST component as a function of the component extent.

In Figure 8 we display a scatter plot comparing the FIRST and NVSS flux densities source by source for objects that appear to be point sources in both surveys. Over three orders of magnitude, the flux scales of the two surveys are in superb agreement.

5.4. Summary

Combining the results of BWH95 and the work presented above, we recommend the following uncertainties be adopted in utilizing the FIRST catalog:

Source flux densities — The uncertainty in the peak flux density S_p is given by the rms noise $\sigma(S)$ at the source position (col. 5). The CLEAN bias for extended sources has been under-estimated in noisy fields (§4.3) and may vary due to other effects, so the true noise in S_p may be larger than $\sigma(S)$. The uncertainty in the integrated flux density S_i can be considerably greater depending on the source size and morphology; for sources not well-described by our elliptical Gaussian model, S_i can either be under- or over-estimated. For bright sources, the accuracy of the flux density determinations are limited by $\sim 5\%$ systematic effects. (These are due to a combination of many effects, including small calibration errors, uncertainties in the primary beam correction factor, mismatches between the Gaussian model and the actual source shape, etc. 5% is a conservative estimate; the actual errors are smaller for most sources.) Analysis of multiply observed sources has revealed that in a few percent of all fields, systematic flux calibration errors produce a larger offset in the flux scale. As these fields are reobserved, the coadded images which contain them will be replaced; a current list of names of questionable fields can be found on the FIRST homepage.

Source positions — Point sources at the detection limit of the catalog have positions accurate to better than $1''$ at 90% confidence. An empirical expression for the positional accuracy at 90% confidence is $(B^2 + \psi^2)^{1/2}(1/\mathrm{SNR} + 1/20)$ arcsec, where $B = 5''\!.4$ is the beam size, ψ is either the major or minor axis FWHM as given in the catalog, and SNR is the peak flux density signal-to-noise

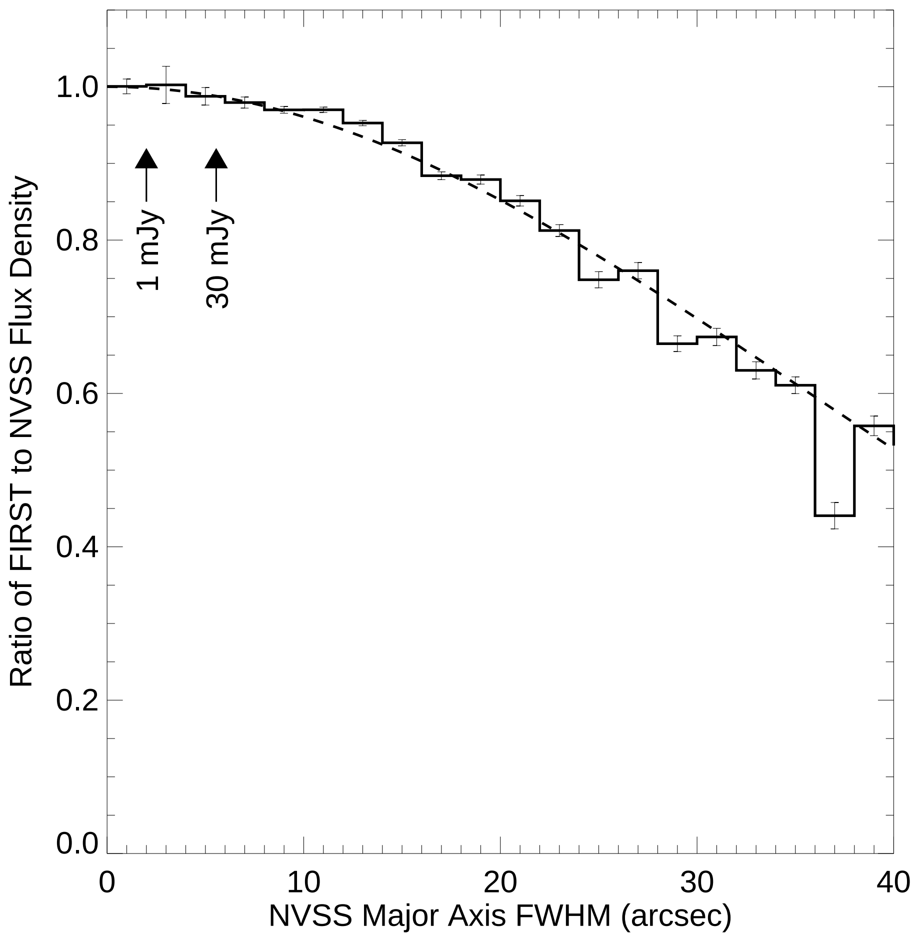

Fig. 7.— Mean ratio of the FIRST flux density to the NVSS flux density as a function of NVSS source size for 7,318 isolated sources in the FIRST catalog. The excellent agreement for small sizes indicates agreement in the FIRST and NVSS flux calibration scales. The arrows show the median sizes for 1 mJy and 30 mJy sources (Windhorst et al. 1990). The dashed line is a model fit: $\exp\left(-\psi^2/50^2\right)$, where the size ψ is measured in arcseconds. The great majority of sources in the FIRST catalog have lost little flux due to overresolution.

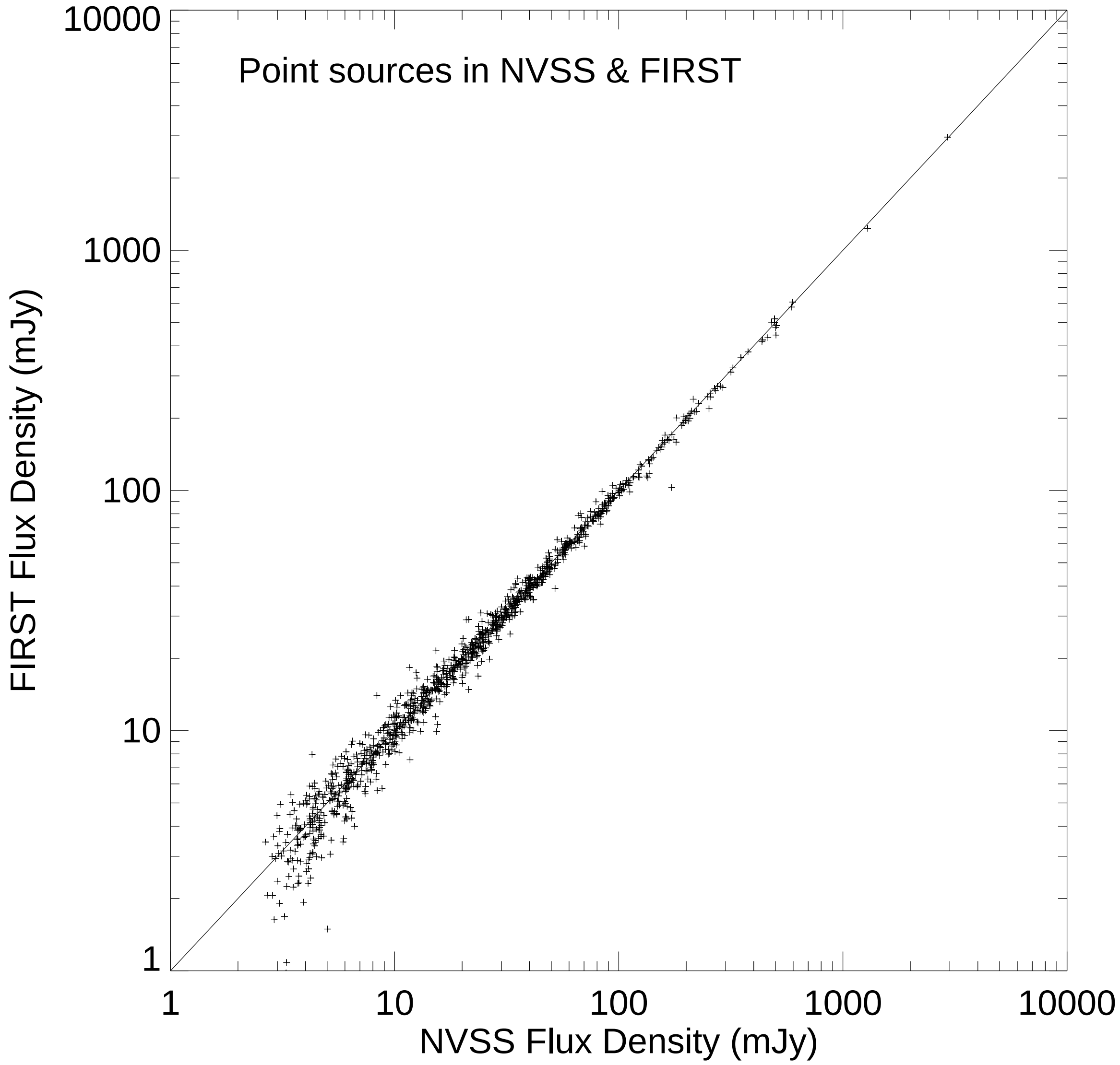

Fig. 8.— Comparison of FIRST flux densities with NVSS flux densities for point sources. Both flux densities have been corrected for CLEAN bias. The flux scales agree very well over 3 orders of magnitude in brightness.

ratio given by SNR $= (S_p - 0.25)/\sigma(S)$. The uncertainty is elliptical for elliptical sources. The limit on positional accuracy for bright sources is $\sim 0\farcs2$, which is $\sim 4\%$ of the synthesized beam FWHM B. Systematic astrometric errors are less than $0\farcs05$.

Source sizes — Uncertainties depend on both the brightness and the intrinsic source extent. Nearly unresolved objects at the catalog flux density limit have uncertainties of about $2''$ (i.e., faint objects with quoted major axes of $2''$ or less are consistent with point sources). Larger objects have somewhat more accurate sizes. An empirical estimate of the uncertainty as a function of flux density and size is

$$\sigma(\psi) = \left(\frac{B\sigma_F}{2}\right)^{1/2} \frac{1 + (\psi/B)^2 + (B/2\sigma_F)^{1/2}(\psi/B)^4}{1 + (B/2\sigma_F)(\psi/B)^4} \quad , \tag{1}$$

where $\sigma_F = \psi_F\,(1/\mathrm{SNR} + 1/75)$ is the RMS uncertainty in the fitted size $\psi_F = (B^2 + \psi^2)^{1/2}$. Flux on scales larger than $\sim 30''$ is largely resolved out by these B-configuration observations (Fig. 7).

6. The FIRST log N – log S

The number of sources as a function of flux density is of interest both as a measure of the evolution of cosmic radio sources and also as a powerful test of the completeness and accuracy of the FIRST radio catalog. We have computed the dN/dS distribution of FIRST sources from our catalog for comparison with previous observations of 1.4 GHz source counts.

A number of corrections must be applied to the raw histogram of source counts versus integrated flux density in order to derive the true dN/dS. We must first account for the varying sensitivity of the survey. Each source in the histogram is assigned a weight equal to the inverse of the area over which the source could have been detected (Katgert et al. 1973). We compute this area from the coverage map by summing the areas of all pixels with rms noise $\sigma(S) < S_p/5$ for a source with fitted peak flux density S_p before CLEAN bias correction. Note that sources with $S_p > 10$ mJy/beam can be detected anywhere in the survey area, so the effective area covered is constant above that peak flux density threshold. However, sources with large integrated flux densities can still have low peak flux densities, so the dN/dS histogram is affected by the areal coverage at integrated flux densities well above this threshold.

The next step is to combine close groups of sources, which usually are not independent objects but rather are components of a single object. To be consistent with previous studies (Oosterbaan 1978, Windhorst et al. 1985), we have merged sources that fall within $50''$ of their nearest neighbors. Nearly 30% of the FIRST catalog entries fall in such groups, which reduces the number of sources in the dN/dS histogram by $\sim 15\%$.

The appropriate weighting for merged sources is somewhat complicated. Since the sources in a pair have different fluxes, they can have different associated survey areas over which the sources could have been detected. Suppose we have two sources with integrated flux densities $S_{i1} \le S_{i2}$

and detection areas $A_1 \leq A_2$. In area A_1, both components of this pair are detected; in area $A_2 - A_1$ only the brighter component (source 2) would be detected; and in the remaining survey area neither component would be detected. Consequently, for this pair we not only have to add a contribution to the flux bin $S_{i1} + S_{i2}$, but we also have to *subtract* a contribution from bin S_{i2} to correct for source pairs where S_{i1} was not detected. The weights for the pair are $1/A_2 - 1/A_1$ in bin S_{i2} and $1/A_1$ in bin $S_{i1} + S_{i2}$. For groups with more than 2 sources, a contribution is subtracted from several bins representing partial sums of the brighter source flux densities and added to the bin that includes the sum of the integrated flux densities of all components.

Another complication introduced by merging sources is that some merged sources are not components of a single objects but are independent objects that happen to fall close together in the sky. The FIRST catalog has ~ 90 sources/deg^2, so $\sim 5\%$ of the sources are expected to have a random neighbor within $50''$. Most of the randomly merged sources fall at the faint end of the flux distribution, since most sources in the catalog are faint. If the true source distribution is $f(S) = dN/dS$, then the observed distribution after random mergers is

$$f_{obs}(S) = \sum_{n=1}^{\infty} f_n(S) \quad . \tag{2}$$

The function $f_n(S)$ is the flux distribution of randomly merged groups containing n sources:

$$f_n(S) = \frac{A}{n} \int_0^S dx\, f_{n-1}(x) f(S-x) \tag{3}$$

where

$$f_1(S) = f(S)\, e^{-\bar{N}} \tag{4}$$

is the distribution of single (unmerged) sources. Here $\bar{N}$ is the mean number of sources per merging area A:

$$\bar{N} = A \int_0^\infty f(S) dS \quad . \tag{5}$$

As long as $\bar{N} \ll 1$ (as is the case for the FIRST catalog with a $50''$ merging radius), a simple way to solve for $f(S)$ given $f_{obs}(S)$ is to use the method of successive substitutions, starting with $f = f_{obs}$ and iterating

$$f(S) = e^{\bar{N}} \left[f_{obs}(S) - \sum_{n=2}^{\infty} f_n(S) \right] . \tag{6}$$

Correcting the FIRST dN/dS distribution using this approach leads to an increase of $\sim 5\%$ in the source counts between 1 and 2 mJy, with the correction declining to a negligible level above ~ 6 mJy.

The resulting merged dN/dS distribution is compared with the distribution derived by Windhorst (Windhorst 1984, Windhorst et al. 1985, Windhorst et al. 1990) in Figure 9. The shape of our distribution is very similar to Windhorst's observations, but the FIRST counts are too low by $\sim 15\%$. This normalization difference is due to the fraction of sources that are partially resolved by

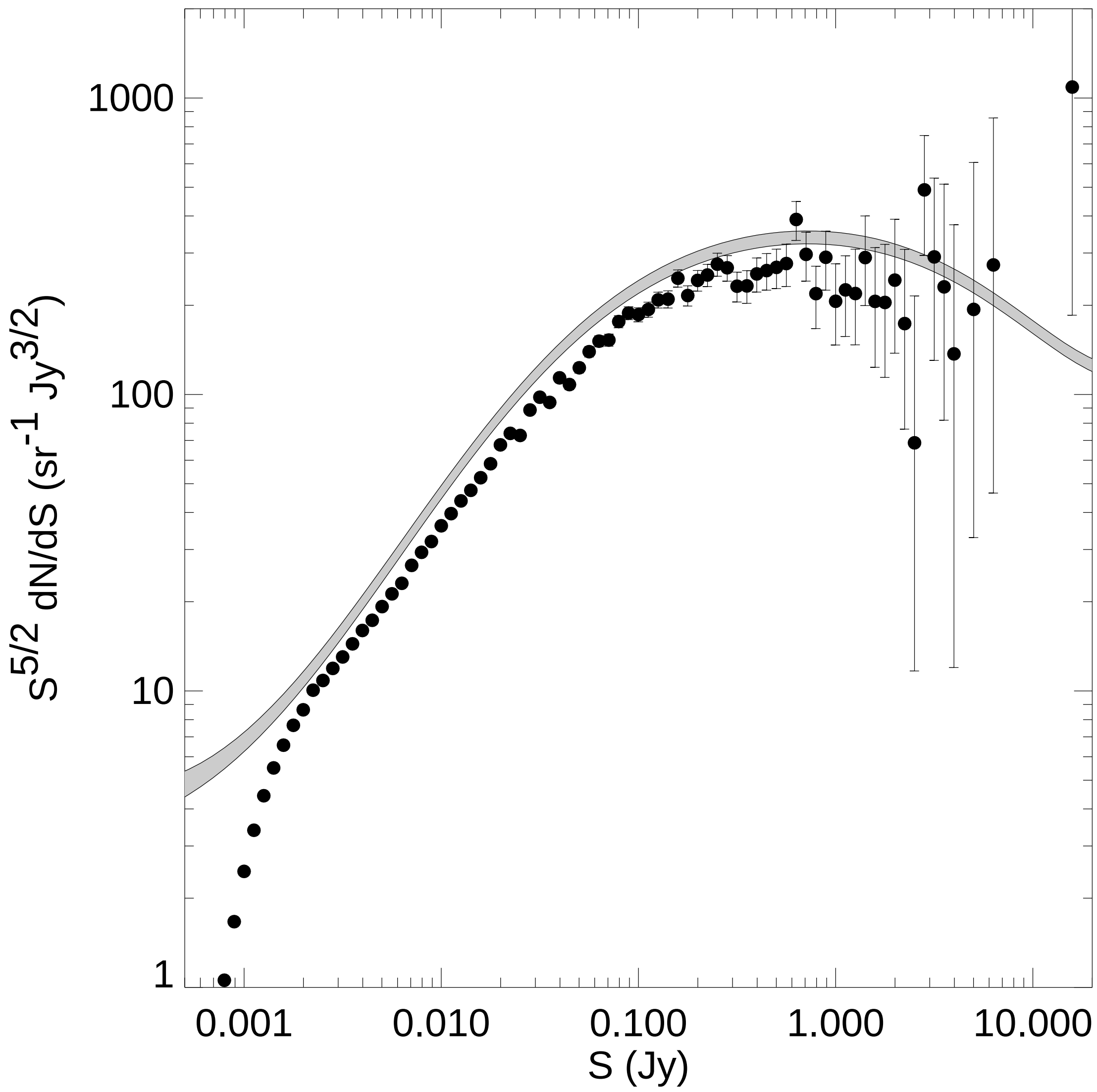

Fig. 9.— Flux density distribution of FIRST catalog sources before correction for resolution effects. The differential FIRST source counts, normalized to the distribution expected for a non-evolving population in flat space, is plotted versus the integrated flux density. Sources closer than $50''$ are assumed to be components of a single object and have been merged. Flux densities have been corrected for CLEAN bias, and a small correction has been applied to account for the random merging of independent sources. The shaded line shows the dN/dS distribution determined by Windhorst et al. (1984, 1985, 1990).

our B-configuration observations and so have integrated flux densities in the FIRST catalog that are too small. We can correct the dN/dS distribution for this resolution effect using the lost flux versus size distribution determined from comparison to the NVSS catalog (Fig. 7) along with the source size distribution. We show the dN/dS distribution before correction for resolution effects mainly to demonstrate the amplitude of these effects, which are modest and can be neglected for many purposes. In cases such as computing dN/dS where a correction must be applied, the corrections required are not dramatic.

Assume that the ratio R of the observed brightness of a source to its true brightness is a function only of the source's major axis FWHM ψ. We model the empirical $R(\psi)$ shown in Fig. 7 using the simple analytical function

$$R(\psi) = e^{-(\psi/50'')^2} \quad . \tag{7}$$

Using this function, we have computed the expected distribution of R for FIRST catalog sources by taking the size distribution of NVSS sources brighter than 20 mJy (fainter sources have very noisy sizes) and computing from it the distribution of R values. When normalized by the number of sources, this gives the probability that a source will have a given value of R. The results are shown in Figure 10.

The source size distribution (and so the R distribution) is known to vary with the source flux density (Oort 1988, Windhorst et al. 1990). Windhorst et al. 1990 propose two simple functions describing the size distribution. The integral size distribution is

$$h(>\psi) = \exp\left[-\ln 2\,(\psi/\psi_{med})^{0.62}\right] \tag{8}$$

where the median size is

$$\psi_{med}(S) = 2.0\, S^{0.30}\,\mathrm{arcsec} \quad , \tag{9}$$

with S in mJy. The line in Figure 10 is the distribution derived by combining this size distribution with the observed NVSS flux density distribution, accounting for the $S_p > 20$ mJy flux cut applied to the data. The model ratio distribution is practically identical to the observations, which gives us confidence that the Windhorst distribution can be applied to our data.

The effects of resolving out flux on dN/dS are most conveniently computed using the source count distribution in $z = \log S$. A flux ratio R corresponds to a difference between the true and observed log flux values $z_{obs} - z = \log R = Q$. The observed dN/dz distribution $g_{obs}(z)$ is a convolution of the true distribution $g(z)$ with the resolution function $P(Q)$:

$$g_{obs}(z) = \int_z^\infty dz' g(z') P(z - z') \quad . \tag{10}$$

Deriving $g(z)$ from $g_{obs}(z)$ is a deconvolution problem if the size distribution does not depend on the source flux density. In our case, the size distribution *does* depend on flux: $P(Q|S)$ is the probability

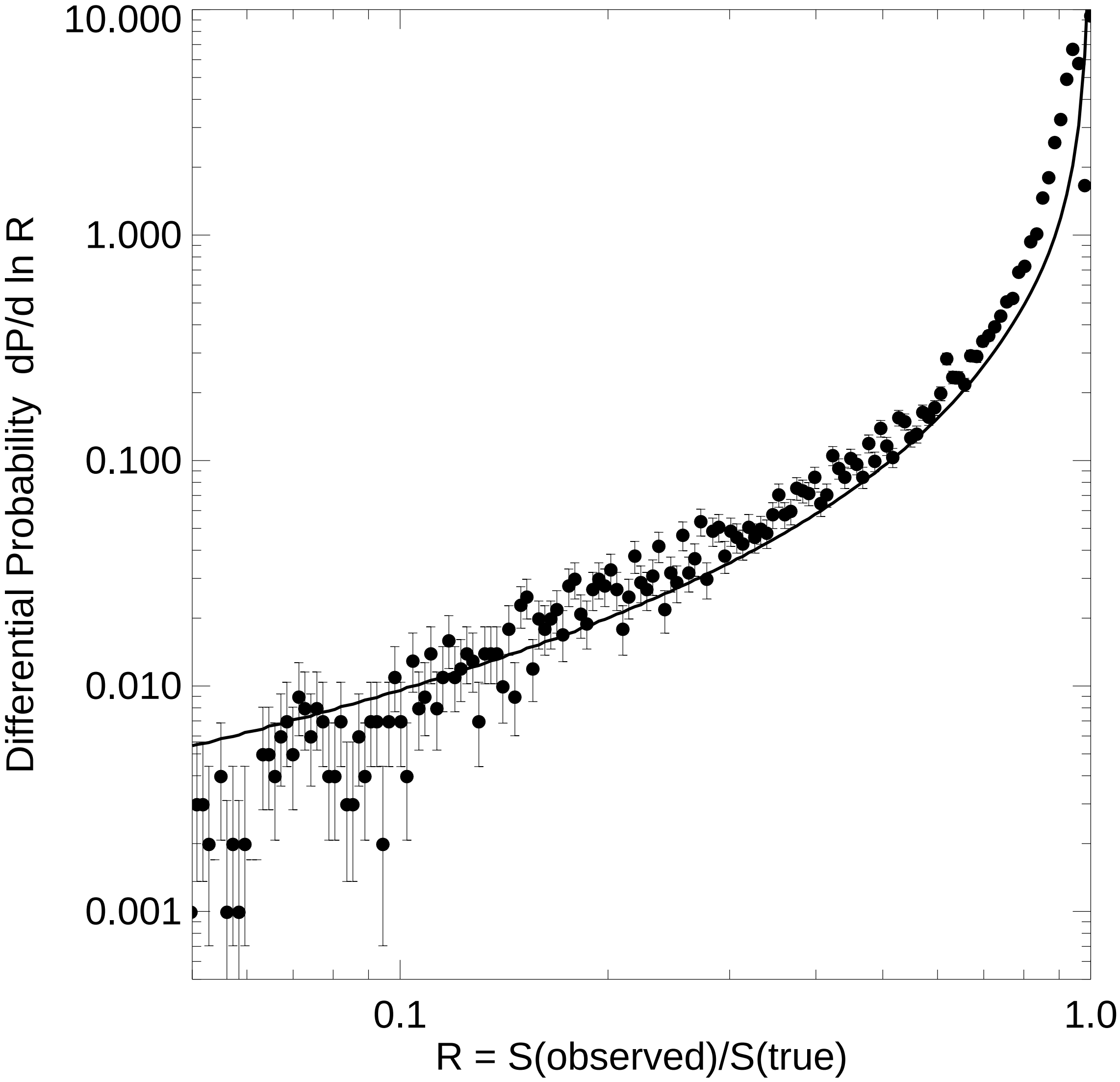

Fig. 10.— Distribution of ratio of observed flux density to true flux density for FIRST sources expected based on the NVSS size distribution and the FIRST/NVSS flux ratios (Fig. 7). The line shows the prediction based on the Windhorst et al. (1990) size distribution, which is in excellent agreement with the data. The excess of sources over the model near $R = 1$ is the result of noisy size estimates for sources that are small compared with the $45''$ NVSS beam.

of a log flux ratio Q for a source with true integrated flux density S. Then the equation we must solve is

$$g_{obs}(z) = \int_z^\infty dz' g(z') P(z - z'|z') \quad . \tag{11}$$

This is not simply a convolution, but it is still an inverse problem solvable by techniques similar to deconvolution.

We have used the Richardson-Lucy method (Lucy 1974) to perform the resolution correction for the merged FIRST dN/dS. The Richardson-Lucy iteration finds the maximum likelihood solution to the deconvolution problem in the presence of Poisson noise in the data (Shepp & Vardi 1982). Since the noise in our problem is mainly due to counting statistics (though it has been slightly modified by the weighting and merging schemes described above), this method is a good choice for the deconvolution. Forty iterations of the Lucy method were used; it converges rapidly because $P(Q|S)$ is strongly peaked near $Q = 0$.

The "point-spread function" $P(Q|S)$ was computed from the Windhorst et al. 1990 size distribution (Eq. 8) and the empirical flux loss function (Eq. 7). We have included in $P(Q|S)$ a Gaussian noise of 0.2 mJy/beam in the peak flux densities, which also acts to blur the true dN/dS. Note that the size distribution gets cut off at a flux-dependent maximum size where the peak flux density drops below the 1 mJy/beam survey threshold.

Figure 11 shows the FIRST dN/dS corrected for resolution effects. The corresponding counts are given in Table 2. It is in excellent agreement with previous determinations (Windhorst et al. 1985, Windhorst et al. 1990) over 4 orders of magnitude in integrated flux. Over the flux density range 2–30 mJy this is the most accurate determination ever made of the radio $\log N - \log S$ distribution. The statistical errors are very small, and we estimate that the systematic errors (from flux scale errors, uncertainties in the resolution correction, etc.) to be less than 3% except at faint flux densities. The roll-off at flux densities less than ~ 2 mJy is due to the low peak flux densities of many faint, extended sources, which make them undetectable in our survey. This is a useful indication of the incompleteness of the survey at the faint limit.

7. Comparison With Other Catalogs

The combined sensitivity and positional accuracy of the FIRST catalog are unprecedented compared with any previous wide-area radio catalog. (The NVSS [Condon et al. 1994] catalog has a sensitivity ~ 3 times lower and a resolution $\sim 1/8$ that of FIRST.) As such, our source locations have an accuracy that match or exceed those of all currently available large astronomical catalogs. As an illustration of the utility of the FIRST catalog, we describe here comparisons we have undertaken with three of the largest existing catalogs at other wavelengths: the Guide Star Catalog (Lasker et al. 1990), the IRAS Faint Source Catalog (Moshir et al. 1992), and the ROSAT WGACAT of X-ray sources (White, Giommi & Angelini 1994).

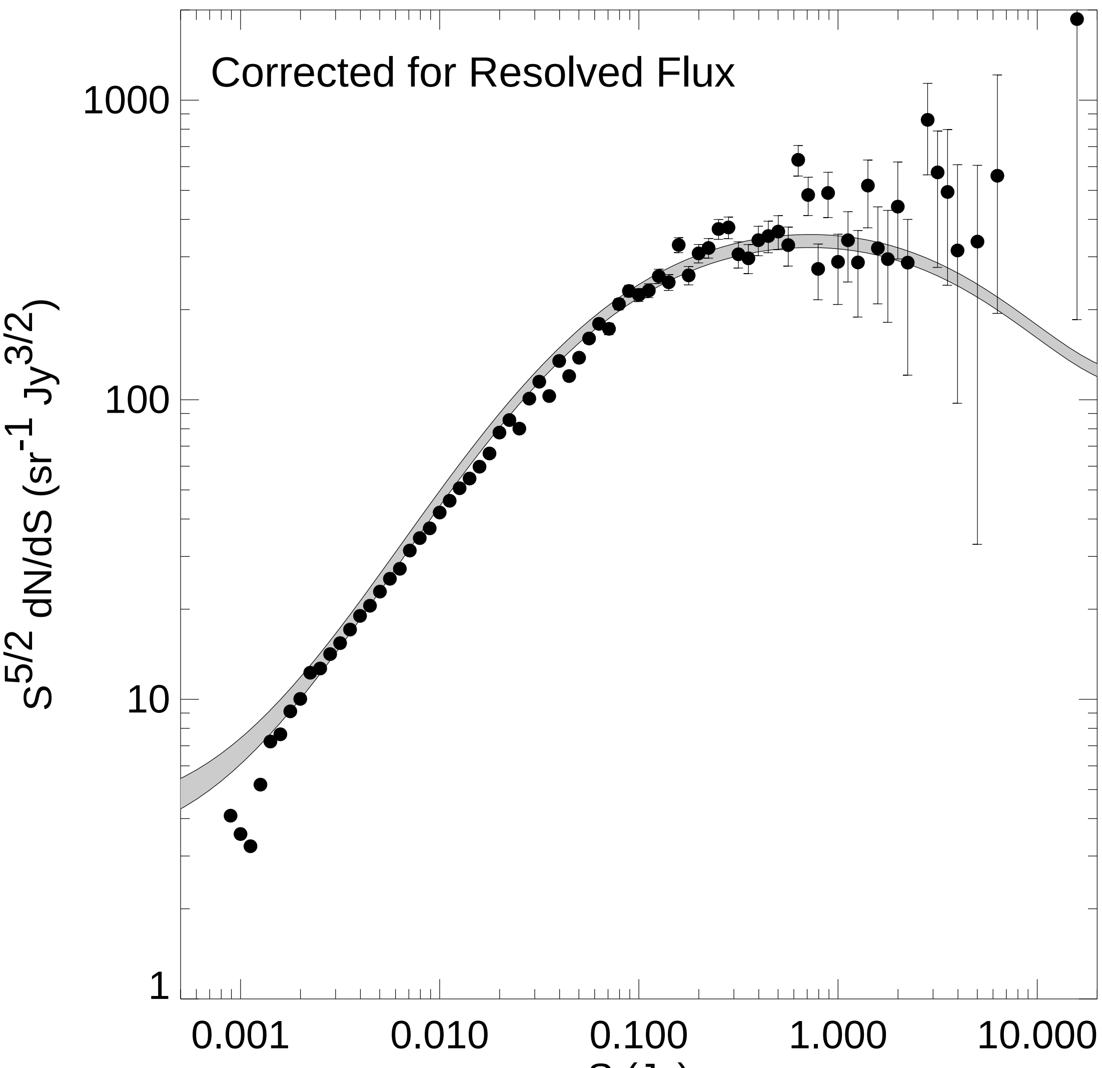

Fig. 11.— The FIRST source counts from Fig. 9 corrected to account for the loss of flux from some sources due to the resolution of extended objects in the VLA B-configuration (Fig. 10). The corrected distribution is in excellent agreement with the Windhorst distribution (shaded). The scatter at bright flux densities is the result of mild noise amplification in the deconvolution. Over the flux density range 2–30 mJy this is the most accurate determination ever made of the radio $\log N - \log S$ distribution.

TABLE 2

FIRST SOURCE COUNTS FROM 1 MJY TO 1 JY

S (Jy)	Counts[a]	RMS[b]	S (Jy)	Counts	RMS	S (Jy)	Counts	RMS
	(sr^{-1} Jy$^{3/2}$)			(sr^{-1} Jy$^{3/2}$)			(sr^{-1} Jy$^{3/2}$)	
0.00100	3.55	0.04	0.01000	42.04	0.85	0.10000	223.84	11.05
0.00112	3.23	0.05	0.01122	46.01	0.97	0.11220	231.56	12.26
0.00126	5.19	0.06	0.01259	50.67	1.11	0.12589	258.50	14.12
0.00141	7.23	0.08	0.01413	54.59	1.26	0.14125	246.72	15.03
0.00158	7.64	0.09	0.01585	59.78	1.43	0.15849	328.63	18.92
0.00178	9.12	0.11	0.01778	66.18	1.65	0.17783	260.16	18.35
0.00200	10.04	0.12	0.01995	77.71	1.94	0.19953	308.06	21.77
0.00224	12.27	0.15	0.02239	85.57	2.22	0.22387	321.10	24.23
0.00251	12.68	0.17	0.02512	80.12	2.35	0.25119	371.50	28.41
0.00282	14.16	0.19	0.02818	100.88	2.87	0.28184	375.95	31.16
0.00316	15.41	0.22	0.03162	114.91	3.34	0.31623	305.91	30.64
0.00355	17.10	0.25	0.03548	102.91	3.45	0.35481	296.63	32.89
0.00398	18.99	0.29	0.03981	134.73	4.30	0.39811	340.95	38.44
0.00447	20.54	0.33	0.04467	119.97	4.42	0.44668	351.93	42.58
0.00501	22.89	0.37	0.05012	138.13	5.17	0.50119	364.21	47.22
0.00562	25.26	0.43	0.05623	160.17	6.07	0.56234	328.27	48.88
0.00631	27.28	0.49	0.06310	179.24	7.00	0.63096	632.16	73.94
0.00708	31.38	0.57	0.07079	172.49	7.49	0.70795	482.54	70.43
0.00794	34.54	0.65	0.07943	208.60	8.98	0.79433	273.39	57.79
0.00891	37.24	0.74	0.08913	230.73	10.29	0.89125	490.07	84.35
						1.00000	288.70	74.38

[a]Normalized source counts, $S^{5/2}dN/dS$, corrected for resolution effects as described in text.

[b]RMS uncertainty from counting statistics in normalized source counts. Systematic effects are estimated to be less than 3% except at faint flux densities ($S < 2$ mJy).

7.1. The Guide Star Catalog

The Guide Star Catalog (GSC), produced by the Space Telescope Science Institute from scans of the "Quick-V" Palomar plates taken between 1982 and 1984, contains 18 million optical objects brighter than magnitude 16. The astrometric accuracy of the catalog is roughly $1''$, although the centers of large galaxies are uncertain by a larger amount. All objects are classified as either stellar or nonstellar although, as we discuss below, the former category contains a significant number of galaxies.

We performed separate matches between the 30,861 objects classified as nonstellar and the 407,473 stellar objects which are found within the boundaries of our survey region. A total of 475 nonstellar objects are coincident with FIRST radio sources within $1''$, where ~ 1 is expected by chance. A number of matches in excess of the false rate is seen out to $10''$, although the false rate becomes $> 50\%$ outside of $5''$. These matches represent a mix of normal, nearby galaxies, starburst galaxies and relatively low-redshift AGN. This complete, flux-limited list of radio sources associated with galaxies in the local Universe is a useful sample in which to identify such objects. Note that we have simply matched the individual radio source positions to each GSC entry; no effort has yet been made to match the counterparts of classical double radio sources by searching between the radio lobes where the parent galaxy is to be expected. We pursue this issue in some detail in an upcoming publication (McMahon et al. 1996).

The match results for stellar objects yielded a surprising number of coincidences given the rarity of luminous stellar radio sources: 98 (182) "stars" are found to be coincident with FIRST radio sources within $1''$ ($3''$), with an expected false rate of only ~ 4 (40) objects. Closer examination of the matches demonstrates that most of these objects are misclassified galaxies. A search of the NED and SIMBAD database revealed that 100 (55%) of the 182 "stellar" matches are, in fact, cataloged galaxies. Another 60 (33%) objects have not been identified previously, although the radio and optical properties of these objects are not statistically different from those of the galaxies in this sample. At most 12 of these objects are certifiably stars, 9 found in the SIMBAD database and 3 observed by us as part of a spectroscopic follow-up of bright optical counterparts to FIRST sources (Gregg et al. 1996), although even the 7 stellar candidates within $1''$ are not significantly greater than the total false rate expected for the GSC of ~ 4. While proper motions complicate the identification of FIRST radio stars, and some of the GSC sources which remain unidentified could be stellar radio sources, our basic results confirm the expected rarity of stars at flux densities above 1 mJy at 20 cm. A further discussion of the handful of true stellar radio sources discovered in our survey may be found in Becker et al. 1996.

7.2. The IRAS Faint Source Catalog

The IRAS Faint Source Catalog (FSC – Moshir et al. 1992) includes over 173,000 objects in the high latitude sky observed at 12, 25, 60, and 100 μm as part of the IRAS Sky Survey in 1983.

The sources represent a mix of stars and galaxies which are present in a ratio of 1:2 averaged over our range of Galactic latitudes. We have adopted the definition of Meurs & Harmon 1988 in using the far infrared colors to identify the probable galaxies; given the results of the previous section we do not expect a significant match rate with the stars. The FSC source position uncertainties are described by error ellipses with a mean major axis of 23″; these are putative 1σ errors, although we find a significant number of real matches out to four times these values and beyond.

Of the $\sim$ 7200 IRAS sources within the FIRST survey area, 1814 have radio counterparts within the 2σ error ellipse, with only 65 expected by chance. All but 11 (0.6%) of these sources have galaxy colors, confirming the nature of the identifications. An additional $\sim$ 320 matches in excess of the number expected by chance are found between 2 and 4 times the 1σ IRAS FSC error ellipse, although the false rate rises to 36%.

It is not surprising that 45% of all IRAS sources with galaxian far infrared colors appear in the FIRST survey. Helou, Soifer & Rowan-Robinson 1985 have shown that a tight correlation exists between the fluxes of radio and far infrared emission in normal and starbursting galaxies. Given our peak flux density threshold of 1 mJy/beam and the median $S(60\mu\text{m})/S(20\text{cm})$ flux density ratio of 125 ± 30 for a radio flux-limited sample (Hacking et al. 1989), we would expect to detect $\sim$ 50% ($\sim$ 90%) of galaxies with 60 μm flux densities above $S(60\mu\text{m})$ thresholds of $\sim$ 125 mJy ($\sim$ 250 mJy). We fail to include many of these nearby, extended objects in our catalog because their radio surface brightnesses fall below our detection limit of 1 mJy/beam (including 0.25 mJy/beam CLEAN bias). We also suffer some secondary loss in sensitivity both as a consequence of resolving out the diffuse flux and as a result of the CLEAN bias (discussed above), which is more severe for extended sources. As we show in BWH95, however, a large fraction of faint extended sources with peak flux densities below our threshold *are* present in our images at $> 5\sigma$ significance, implying that direct queries of the image database will allow us to detect a much larger fraction of FSC sources. Further exploration of the IRAS source component of the FIRST survey is left to a subsequent publication.

7.3. The ROSAT Catalog

The Position Sensitive Proportional Counter (PSPC) on ROSAT conducted nearly 5000 pointings covering π deg^2 each in the 0.1–2.4 keV X-ray band during its four-year lifetime. The High Energy Science Archive Research Center (HEASARC) has produced a catalog of 68,907 sources derived from these images (the WGACAT; White, Giommi & Angelini 1994). The sources include a wide mix of astronomical objects in which coronally emitting stars and AGN dominate. Positional uncertainties are quoted as being 13″ (1σ) if the source was observed within the central 18′ of the field and 50″ if it fell outside this region. A total of 188 (164, 56) PSPC pointings with exposure times exceeding 500 s (2000 s, 10,000 s), covering a total of 520 deg^2, fall within the FIRST survey region.

Of the 2212 sources with the more accurate positions which fall within our radio coverage area, 311 are coincident to within 2 times the quoted X-ray uncertainty (nominally, 95% confidence error circles) with one or more radio sources; only 31 matches are expected by chance. Thirty-eight of these are double radio sources with separations of $\lesssim 50''$, while an additional 5 are radio triples (only ~ 2 double or triple radio matches are expected by chance). The 3312 sources with larger X-ray positional uncertainties provide an additional ~ 600 real matches, meaning that 17% of the ROSAT sources in our survey area have FIRST radio counterparts. This sample represents one of the largest lists of radio/X-ray coincidences available to date (cf. Brinkmann et al. 1995 and Moran et al. 1996), and extends the exploration of X-ray-selected radio sources a factor of 30 fainter in radio flux density than previous large-area surveys. In addition, unlike comparisons with previous radio catalogs, the radio data here provide a substantial improvement in positional accuracy, greatly simplifying the task of X-ray source optical identification.

In addition to the deep, but limited, coverage of the X-ray sky offered by the WGACAT, there exists a uniform, complete survey of the soft X-ray sky produced in the ROSAT All-Sky Survey (RASS) completed over five years ago. The scientific utility of a match between the RASS and the FIRST survey catalog is manifest, but the absence of a public version of this catalog has prevented us from carrying out the match. Perhaps this situation will be rectified in the near future.

8. Public Accessibility and Future FIRST Releases

The FIRST survey is a long-term project whose completion is many years in the future. The area of sky covered by the FIRST survey will increase in increments of ~ 1300 deg^2 every 16 months. At the very least, then, the FIRST catalog of sources will grow with time. More confusing for the potential user, however, the catalog of sources for some particular region of the sky may change with a new release. This will be generally true for regions currently at the edge of the survey for which a previously constructed coadded image will be enhanced as new fields become available. Likewise, sources in the vicinity of internal gaps will change when the gaps are subsequently filled. In addition, source positions and flux densities can be expected to change at the one-sigma level as new data are added. At worst, some sources will disappear and others appear with new releases of the catalog as coverage gaps are filled and as we develop more sophisticated algorithms to exclude spurious sources. It is our intention to archive each release of the FIRST catalog so that researchers using a particular version of the catalog will have continuing access to that version. Instructions for accessing all FIRST data products are available via the FIRST homepage (`http://sundog.stsci.edu`).

The success of the FIRST survey is in large measure due to the generous support of a number of organizations. In particular, we acknowledge support from the NRAO, the NSF (grants AST-94-19906 and AST-94-21178), IGPP/LLNL, the STScI, NATO, the National Geographic Society (grant NGS No. 5393-094), Columbia University, and Sun Microsystems. We also thank Columbia undergraduates Ali Kinkhabwala, Beth Willman, and Gabe Perez-Giz for assistance in performing the comparisons of the FIRST catalog with the data from other wavelength regimes and with the

Lehar high-resolution survey. This paper is Contribution Number 601 of the Columbia Astrophysics Laboratory.

REFERENCES

Argue, A. N. et al. 1984, A&A, 130, 191

Becker, R. H., White, R. L., Helfand, D. J., Gregg, M. D., & McMahon, R. G. 1996, in *Radio Emission from Stars and the Sun*, eds. J. M. Paredes & R. Taylor, in press

Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559 (BWH95)

Brinkmann, W., Seebert, J., Reich, W., Furst, E., et al. 1995, A&AS, 109, 147

Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., & Broderick, J. J. 1994, preprint

Gregg, M., D., Becker, R. H., White, R. L., Helfand, D. J., McMahon, R. G., & Hook, I. M. 1996, AJ, 112, 407

Gunn, J. E., & Knapp, G. R. 1992, PASP, 43, 267

Hacking, P., Condon, J. J., Houck, J. R., and Beichman, C. A. 1989, ApJ, 339, 12

Helfand, D. J., Das, S. R., Becker, R. H., & White, R. L. 1996, in *Proceedings of the Workshop on Blazar Variability*, eds. H. R. Miller & J. R. Webb, in press

Helou, G., Soifer, B. T., & Rowan-Robinson, M. 1985, ApJ, 298, L7

Johnston, K. J. et al. 1995, AJ, 110, 880

Katgert, P., Katgert-Merkelijn, J. K., Le Poole, R. S., Van der Laan, H. 1973, A&A, 23, 171

Lasker, B. M., Sturch, C. R., McLean, B. J., Russell, J. L., Jenkner, H., & Shara, M. M. 1990, AJ, 99, 2019

Lehar, J. 1991, Ph. D. Thesis, MIT (L91)

Lindegren, L., & Kovalevsky, J. 1995, A&A, 304, 189

Lucy, L. B. 1974, ApJ, 79, 745

McMahon, R. G., et al. 1996., in preparation

Meurs, E. J. A., & Harmon, R. T. 1988, A&A, 206, 53

Moran, E. C., Helfand, D. J., Becker, R. H., & White, R. L. 1996, ApJ, 461, 127

Moshir, M. et al. 1992, *Explanatory Supplement to the IRAS Faint Source Survey, Version 2* (Pasadena: JPL)

Murthy, S. K., Kasif, S., & Salzberg, S. 1994, Journal of Artificial Intelligence Research, 2, 1

Oort, M. J. A. 1988, A&A, 193, 5

Oosterbaan, C. E. 1978, A&A, 69, 235

Salzberg, S., Chandar, R., Ford, H., Murthy, S. K., & White, R. L. 1995, PASP, 107, 279.

Shepp, L. A., & Vardi, Y. 1982, IEEE Trans. Medical Imaging, MI-1, 113

White, N. E., Giommi, P., & Angelini, L. 1994, IAU Circ. No. 6100

Windhorst, R. 1984, Ph. D. Thesis, University of Leiden

Windhorst, R. A., Miley, G. K., Owen, F. N., Kron, R. G., & Koo, D. C. 1985, ApJ, 289, 494

Windhorst, R., Mathis, D., & Neuschaefer, L. 1990, in *The Evolution of the Universe of Galaxies: Hubble Centennial Symp.*, ed. R. G. Kron (ASP Conf. Ser., 10), 389
