

RADIO SOURCES IN THE NCP REGION OBSERVED WITH THE 21 CENTIMETER ARRAY

QIAN ZHENG^{1,2,3}, XIANG-PING WU^{2,3,4}, MELANIE JOHNSTON-HOLLITT¹, JUN-HUA GU^{2,3}, AND HAIGUANG XU⁵*Received xxx; accepted xxx; published 2016 xxx*

ABSTRACT

We present a catalog of 624 radio sources detected around the North Celestial Pole (NCP) with the 21 Centimeter Array (21CMA), a radio interferometer dedicated to the statistical measurement of the epoch of reionization (EoR). The data are taken from a 12 h observation made on 2013 April 13, with a frequency coverage from 75 to 175 MHz and an angular resolution of $\sim 4'$. The catalog includes flux densities at eight sub-bands across the 21CMA bandwidth and provides the in-band spectral indices for the detected sources. To reduce the complexity of interferometric imaging from the so-called “w” term and ionospheric effects, the present analysis are restricted to the east-west baselines within 1500 m only. The 624 radio sources are found within 5 degrees around the NCP down to ~ 0.1 Jy. Our source counts are compared, and also exhibit a good agreement, with deep low-frequency observations made recently with the GMRT and MWA. In particular, for fainter radio sources below ~ 1 Jy, we find a flattening trend of source counts towards lower frequencies. While the thermal noise (~ 0.4 mJy) is well controlled to below the confusion limit, the dynamical range ($\sim 10^4$) and sensitivity of current 21CMA imaging is largely limited by calibration and deconvolution errors, especially the grating lobes of very bright sources, such as 3C061.1, in the NCP field which result from the regular spacings of the 21CMA. We note that particular attention should be paid to the extended sources, and their modeling and removals may constitute a large technical challenge for current EoR experiments. Our analysis may serve as a useful guide to design of next generation low-frequency interferometers like the Square Kilometre Array.

Subject headings: radio continuum: general—instrumentation: interferometric—radio continuum: galaxies—methods: observational

1. INTRODUCTION

Low-frequency observations are very important for the study of the statistical signatures of extragalactic sources, in particular source counts in the low-frequency sky, which until recently have been only characterized at the very highest end of the flux scale. The number of steep-spectrum sources increases rapidly as the sample selection frequency is lowered (Massaro et al. 2014), because these sources become too faint to be detected at high frequencies leading to an expectation that low-frequency sky surveys will uncover this population in greater numbers than current cm-wavelength studies. As it is not trivial to infer the low-frequency ($\nu \leq 300$ MHz) sky from higher frequency data, deep source counts at the same frequency as EoR observations are required for accurate foreground modeling and subtraction. Better understanding of the statistical properties and behavior of radio sources is required and motivated by a number

of science goals, such as the foreground modeling and subtraction in the detection of the EoR.

Over the last 30 years, a large and ever increasing number of surveys and targeted observations have been performed at low frequencies commencing with the 6th Cambridge (6C) Survey of Radio Sources (Hales et al. 2007), which covered the entire northern sky above 30 degrees at 151 MHz. Since then a number of other surveys have been completed such as the 7th and 8th Cambridge Survey of Radio Sources (7C, and 8C: Rees 1990), the targeted survey on the Culgoora Circular Array (Slee 1995) in which the flux densities of 1800 high-frequency-selected radio sources were measured at 80 and 160 MHz, the 74 MHz VLA Low-Frequency Sky Survey (VLSS: Cohen 2007) above -30 degrees, the Murchison Widefield Array Commissioning Survey (MWACS: Hurley-Walker et al. 2014) covering $6,100 \text{ deg}^2$ between 104 and 196 MHz, and the Precision Array for Probing the Epoch of Reionization (PAPER; Jacobs et al. 2011) low resolution survey of the Southern sky at 145 MHz. Current low frequency radio interferometers, such as LOFAR (van Haarlem et al. 2013), GMRT (Paciga et al. 2013), LWA (Taylor et al. 2012), MWA (Bowman et al. 2013; Tingay et al. 2013) and PAPER (Jacobs et al. 2011) have made a considerable advance in our understanding of the population of extragalactic radio sources at frequencies below 300 MHz with all-sky low frequency surveys (Wayth et al. 2015; Heald et al. 2015; Intema et al. 2016). The upcoming Square Kilometre Array (SKA) is expected to continue to push the boundaries of low-frequency astronomy, observing the radio sky at much higher sensitivity and resolution.

wpx@bao.ac.cn

¹ School of Chemical and Physical Sciences, PO Box 600, Victoria University of Wellington, Wellington 6140, New Zealand² National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Beijing 100012, China³ Center for Astronomical Mega-Science, Chinese Academy of Sciences, 20A Datun Road, Beijing 100012, China⁴ Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China⁵ Department of Physics and Astronomy, Shanghai Jiao Tong University, 800 Dongchuan Road, Minhang, Shanghai, 200240, China

Recently, Franzen et al. (2016) determined the 154MHz source counts with MWA observation of the so-called EoR0 field (which is centered at J2000 $\alpha = 00^h00^m00^s, \delta = -27^\circ00'00''$) and compared these with the 7C counts at 151MHz (Hales et al. 2007), as well as the GMRT source counts at 153MHz (Intema et al. 2011; Ghosh et al. 2012; Williams, Intema & Röttgering 2013) finding good agreement to a limiting flux density of ~ 40 mJy. Using LOFAR, van Weeren (2014) examined source counts at 34, 46 and 62 MHz, a wavelength regime not explored since the 8C observations at 38 MHz 25 years earlier. The LOFAR observations are the deepest images ever obtained at these frequencies and reach noise levels of the order of a few mJy, being nearly 2 orders of magnitude deeper than the 8C survey. van Weeren (2014) confirmed the previously known result that the average spectral index of radio galaxies flattens towards lower frequencies, as is expected from synchrotron losses (Franzen et al. 2016). In addition to understanding the properties of the low frequency radio population to study the population itself, it is of great importance to properly characterize the population if we are to fully exploit these instruments to detect the EoR as a precise sky model is vital for subtracting foreground sources (Jelić et al. 2008; Ghosh et al. 2012; Moore et al. 2013; Jelić et al. 2014; Thyagarajan et al. 2015; Offringa et al. 2016). Thus, source count studies of the low frequency sky are strongly motivated by a number of complementary science goals.

The North Celestial Pole (NCP) region is covered at low-frequencies by the WSRT (Bernardi et al. 2010) and LOFAR (Yatawatta et al. 2013), it is also the sky region targeted by the 21 Centimeter Array (21CMA). The key science goal for the 21CMA is to statistically measure the redshifted 21cm signal of neutral hydrogen from the EoR. Determining the radio source counts and understanding the spectral properties of the radio sources are important for removing the foreground sources and extracting the faint EoR signal. In this paper, we present targeted observations centered on the NCP and extending 5 degrees in radius using the 21CMA. We determine the source counts using eight sub-band images between 75 MHz and 175 MHz with flux densities down to ~ 0.1 Jy. We also investigate the spectral properties of catalog sources and compare our source counts with the recent surveys.

The outline of the paper is as follows. Section 2 describes the 21CMA telescope, the observation and data reduction pipeline. Section 3 presents the source extraction process and resultant source catalog and properties including number counts, spectral indices, completeness and errors. Discussion and concluding remarks are given in Section 4.

2. OBSERVATIONS AND DATA REDUCTION

2.1. The 21CMA

The 21CMA is a ground based radio interferometer dedicated to detection of the EoR. The array, sited in the Ulatai valley of western China, consists of 81 pods or stations, with a total of 10287 log-periodic antennas are deployed in two perpendicular arms along an east-west (6.1 km) (Figure 1) and north-south (4 km) direction, respectively. Spacing of these 81 pods is chosen such



Figure 1. Photograph of the east-west baseline of the 21CMA, in which the building to the left is the control center.

that a sufficiently large number of redundant baselines and a good uniform uv coverage can both be guaranteed. Each antenna element has 16 pairs of dipoles with lengths varying from 0.242 m to 0.829 m, optimized to cover a frequency range of 50-200 MHz, which gives rise to an angular resolution of $3'$ at 200 MHz. All the antennas are fixed on the ground and point at the NCP for the sake of simplicity and economy, which allows us to observe the same patch of sky for 24 hours a day through the whole year. Therefore, we can reach a higher sensitivity in a region of a few ten square degrees around the NCP in a relatively short time. Phase delayed Coaxial cables are used to combine the signal from each antenna for each pod, and the output signal is digitized at a sampling rate of 400 MHz with 8 bit precision to accommodate the bandwidth of 200 MHz, although signal below 50 MHz is actually filtered out. Fast Fourier Transforms (FFT) and correlations are performed in software by a cluster of 83, dual-core Intel(R) Xeon(TM) servers, equipped with gigabit ethernet networks (PCI-E 10-Gbps 4X) and a high speed network (Infiniband) switch (CISCO SFS-7008P). The 200 MHz bandwidth is divided into 8192 channels, giving rise to a frequency resolution of 24.4 kHz. Data in each channel are integrated for about 3 seconds in memory before they are output to a disk array of 32 terabytes. Hard disks are transported to the headquarters of the National Astronomical Observatories of China in Beijing for analysis.

Construction of the 21CMA was completed in 2006, and a significant upgrade was made in 2009 by installing a new type of low noise amplifier (40 K) on each antenna and 40 sets of GPUs in a new data acquisition cluster. The former considerably suppresses thermal noise from the phased delayed cables, while the later provides a very efficient way to speed up correlation computation processing. We have used software instead of hardware or FPGAs to perform the FFT and correlations. After this upgrade we also decided to run the E-W baseline only to maintain a higher efficiency (50%) of data acquisition by reducing the computing demands for cross-correlations. Furthermore, this allows us to work with a relatively simple interferometric imaging algorithm without the con-

cern of the so-called “w” term.

In the early commissioning observations between 2005–2009, many efforts were made to understand the performance of the 21CMA system including antennas, receivers, data acquisition process, and even the power supply. We have developed all software packages and techniques such as identification and mitigation of radio frequency interference (RFI), calibration, deconvolution, wide-field imaging, etc. necessary for full data processing of 21CMA data. Since July of 2010 after its upgrade, the 21CMA has routinely observed the low frequency sky round the NCP, delivering about 2 terabits of visibility data per day for offline analysis.

While the primary goal of the 21CMA is to accumulate a deep observation for EoR detection, even with relatively short integrations we can characterize the extragalactic source population. On one hand, these bright extragalactic sources, a byproduct of our EoR experiment, can be of great interest for exploration of their astrophysical properties. On the other hand, these radio sources constitute one of the major contamination components for our EoR detection, and therefore should be perfectly imaged and subsequently subtracted. In the current work we present the point radio sources observed with the 40 pods of the 21CMA E-W baselines for an integration of 12 hours made on 2013 April 13. An extra deep sample with a higher sensitivity from a longer integration time of up to years will be published later.

2.2. RFI Removal

Although the radio environment of the 21CMA site, Ulaistai, is exceptionally quiet at low frequency and comparable to sites such as the Murchison Radio Observatory (Offringa et al. 2015), RFI is still a prominent concern for doing the EoR experiment. Figure 2 displays the average spectrum of cross-correlation between two pods over 12 hours with a frequency resolution of $\Delta\nu = 24.4$ kHz, which reveals the RFI at the site more clearly than the self-correlation of the same pod because thermal noises from different pods are not correlated. The overall shape reflects the decreasing sky noise at high frequencies and the decreasing efficiency of the antennas at low frequencies with a turnover at roughly 85 MHz. The strongest RFI sources are identified as low orbiting satellites at 137 MHz, local train communications at 150 MHz, civil aviation aircraft around 130 MHz and at 119 MHz, FM radio broadcasting at 88–108 MHz scattered by meteor tails and aircraft, and AM radio broadcasting around 70 MHz, and several other contaminators (Huang et al. 2016). All these sources of RFI are highly time variable, and can be easily flagged for blanking during the editing process.

After the mitigation of strong, time variable RFI, we apply a second statistical algorithm for further removal of other possible RFI using the fact that the sampling of visibilities should exhibit a Gaussian distribution for a given frequency, a scheme often adopted in data analysis of recent low frequency experiments (e.g. Bowman et al. 2007; Ghosh et al. 2011). In the second stage of RFI mitigation we excluded all the data points beyond 3σ deviation from Gaussian statistics. This operation is applied to the real and imaginary parts separately, and the visibility data are edited out if either of the parts fail to reach the 3σ restrictions. This still allows us to use

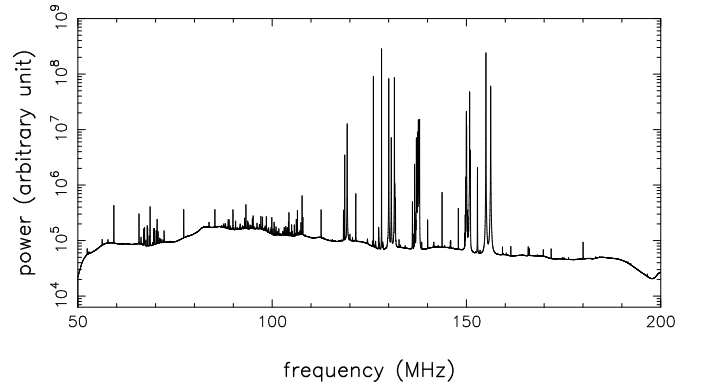


Figure 2. A typical example of the average spectrum of visibility between two pods (E20 and W20, separated by a baseline of 2740 m) integrated over 12 hours. The frequency resolution is 24.4 kHz.

the bulk (about 97%) of the data except for some of the most noisy channels around 137 MHz and 150 MHz. The reader is referred to Huang et al. (2016) for a detail description of our RFI identification and mitigation processing.

2.3. Self-calibration

Low-frequency radio interferometric observations are often affected by ionospheric turbulence, resulting in fluctuations of both source positions and flux densities. We run two algorithms to correct the phase error in the visibilities due primarily to ionospheric effects: self-calibration and the closure relation. Software developed in-house is used for self-calibration, in which a time interval of 1 hour is used for the present study. We have tested shorter intervals of down to 10 minutes, and found that the changes are only minor when longer baselines of $L > 1500$ m are excluded (see below). We have already improved our self-calibration algorithm by combining conventional self-calibration and redundant calibration, we will apply it for all baselines using a short time interval in the future data reduction (Zheng et al. in preparation). The NCP sky at 50–200 MHz is actually dominated by two bright sources: a central bright quasar (NVSS J011732+892848) and a galaxy 3C61.1. 3C61.1 is a very bright giant galaxy with a flux density of 33 Jy at 150 MHz, which also shows a complex structure. We select 10 radio sources including these two bright sources (Table 1) to correct the phase errors with self-calibration. Following this, we use the closure phase relation to check, and modify if necessary, the results.

2.4. Imaging

To reduce further the ionospheric effect on our interferometric imaging and improve image quality, we exclude both longer baselines of $L > 1500$ m and the shorter ones of $L < 100$ m in the present analysis. We perform a fast Fourier transform of the uv map using a uniform weighting with 4096^2 grids, sampled by different baselines and snapshots for each frequency channel, to generate the dirty maps. This yields a total of 6144 images over 50–200 MHz frequency bands with a frequency resolution of 24.4 kHz. Then, each dirty image is deconvolved using the conventional Högbom CLEAN algorithm (Högbom 1974) with a loop-gain of 0.05. We terminate the CLEAN

Table 1
10 sources used in the sky model for self-calibration
(Data are taken from NASA/IPAC EXTRAGALACTIC DATABASE (NED))

Object Name	R_0^a (degree)	$S_{150\text{MHz}}^b$ (Jy)	α^c	Source Type
NVSS J011732+892848	0.52	5.36 ± 0.91	0.01 ± 0.02	point source
NVSS J062205+871948	2.67	4.38 ± 0.49	-0.86 ± 0.02	point source
NVSS J133218+865005	3.17	2.94 ± 0.33	-0.98 ± 0.02	point source
NVSS J092016+862845	3.52	4.04 ± 0.44	-0.84 ± 0.02	point source
3C061.1	3.71	33.23 ± 7.82	-0.80 ± 0.05	resolved
NVSS J122518+860839	3.86	2.51 ± 0.29	-0.69 ± 0.02	point source
NVSS J1013.7+8553	4.12	3.01 ± 0.34	-0.77 ± 0.02	point source
NVSS J190350+853648	4.39	4.60 ± 0.52	-0.72 ± 0.02	point source
NVSS J194136+850138	4.97	4.53 ± 0.71	-0.86 ± 0.03	point source
NVSS J212926+845326	5.11	5.68 ± 0.74	-0.70 ± 0.03	point source

^a The angular distance from the NCP.

^b Flux density measured at 150 MHz

^c The best-fit spectral index in terms of available data in NED.

process when the fractional change in the total CLEANed flux density is less than 10^{-4} in the iteration.

Employment of a smaller threshold does not alter the result significantly. After the deconvolution we adopt a larger frequency bin of 1.56 MHz to combine the 6144 images into 96 mosaics, and concentrate on only the central field of view of 1024×1024 pixels, corresponding to a 14^2 degree square area. Note that this gives rise to an angular resolution of ~ 1 arcmin per pixel while the best resolution of the array is only 3 arcmin at 200 MHz.

2.5. Calibrations of Primary Beam and Gains

Directivity and gain of the 21CMA antenna element, the log-periodic antenna, has been precisely measured in laboratory, which indeed allows us to compute the primary beam and gain of the 21CMA pod according to the pattern multiplication theorem. However, in situ calibration has to be made to account for the possible effect of crosstalk between the antennas themselves and the antennas and the ground. For a 12 hour (or more) observation of the 21CMA centered on the NCP, the primary beam of a pod has circular symmetry. Moreover, in terms of our laboratory measurement of the spatial response of the 21CMA log-periodic antenna and theoretical prediction based on the pattern multiplication theorem, the primary beam pattern of a 21CMA pod is nicely approximated by a Gaussian profile characterized by the standard deviation $\theta_b = 3^\circ.62(\nu/100\text{MHz})^{-1}$. Guided by our knowledge and experience, we adopt a modulated Gaussian function parametrized by

$$F(\theta, \nu) = G(\nu)(1 + A\theta + B\theta^2)\exp(-\theta^2/2\theta_b^2) \quad (1)$$

for the primary beam and gain of the 21CMA pod, where G is the system gain which varies with frequency ν , θ measures the radial distance from the NCP, θ_b denotes the standard deviation of the Gaussian distribution, and A and B are the coefficients in the polynomial of degree 2.

Unfortunately, there are no standard radio sources in the NCP region that can be used for calibration purposes. Thus we use the same sources selected in the self-calibration (Table 1) as our flux calibrators. We collect the flux density measurements of these sources made at different frequencies ranging from a few ten MHz to a few GHz from the literature and obtain the spectral index,

α , for each source by fitting the observed flux densities to a power law, $S_\nu = S_0\nu^\alpha$. Except for the flat-spectrum sources, the fitting errors in amplitude S_0 and spectral index α are typically $\sim 10\%$ and $\sim 3\%$, respectively. Errors in these calibrators will be included in fitting of gain and spatial response function of the telescope.

We then use the total flux density within a frequency bin of 1.56 MHz to accommodate our CLEANed images over the same frequency band. However, the primary beam is frequency-dependent, and varies roughly as an exponential profile characterized by θ_b for the 21CMA. This reduces further the number of the bright radio sources that can be used for calibration in the higher frequency bands. As a result, only the inner 4 sources within $3^\circ.52$ can be used for the 48 higher frequency bands with $\nu \geq 150$ MHz, for which uncertainties in the determination of gain and spatial response might be large.

We start with a Gaussian profile of $F(\theta, \nu) = G(\nu)\exp(-\theta^2/2\theta_b^2)$, and find the best-fit parameters G and θ_b in each frequency channel. By fixing G and θ_b , we then use the same calibrators to find the modification parameters A and B to the Gaussian distribution. It turns out that the Gaussian profile provides rather a good description for the 12-hour averaged primary beam, and the modification parameters A and B are actually around zero (see Figures 3 and 4). Moreover, the best-fit Gaussian beam of

$$\theta_b = 3^\circ.58 \left(\frac{\nu}{100\text{MHz}} \right)^{-1.21} \quad (2)$$

is in good agreement with our theoretical prediction. Figure 5 shows the resultant system gain against frequency. The design of the 21CMA antenna is such that the effective area remains roughly constant over a wide frequency range, which yields a frequency dependent gain of $G \propto \nu^2$. The steeper profile of the fitted gain factor also reflects the compensation for signal attenuation of the phased-delayed coaxial cable used for signal combination.

2.6. Dynamical Range

The CLEANed image in each of the 96 frequency channels is corrected for beam response and flux gain. Then the images are stacked into 12 sky maps with a band-

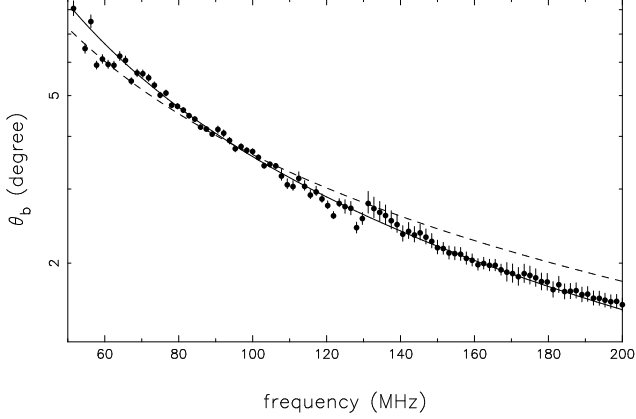


Figure 3. The primary beam of the 21CMA pod, characterized by the standard deviation of the Gaussian distribution θ_b . The dots with error bars are the fitting results from all the sub-band images with a frequency width of 1.56 MHz. The solid curve presents the best-fit, $\theta_b = 3^\circ.58(\nu/100\text{MHz})^{-1.21}$, compared with the theoretically predicted variation of $\theta_b = 3^\circ.62(\nu/100\text{MHz})^{-1}$ (dashed).

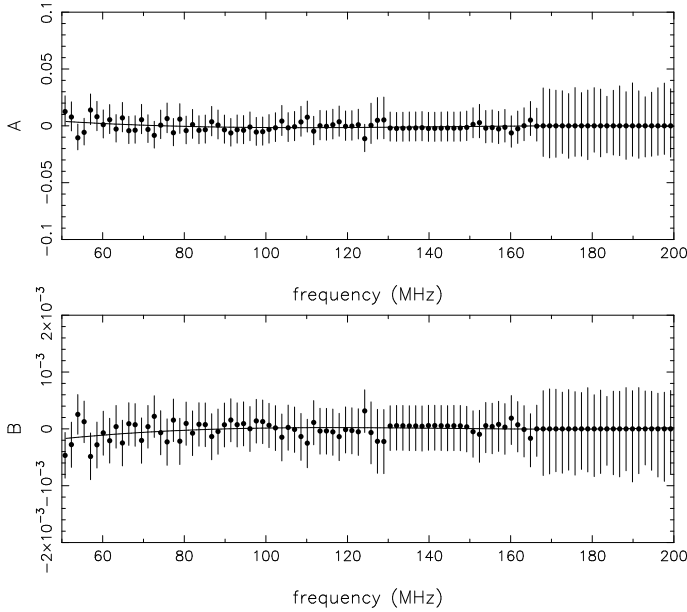


Figure 4. The fitting results of the coefficients A (top panel) and B (bottom panel) in the polynomial of degree 2 to modify the Gaussian primary beam. The solid curve in each panel presents the best-fit, which is very close to zero.

width of 12.5 MHz. To reduce further the possible roll-off effect of the bandpass filters used in the 21CMA receivers, we will not use the data at the two frequency ends, 50-75 MHz and 175-200 MHz, in the present paper. This yields a total of 8 sky maps centered at 81.25, 93.75, 106.25, 118.75, 131.25, 143.75, 156.25 and 168.75 MHz, respectively, with an image noise of $\sim 0.7-0.4$ mJy. The noise and beam size in each 12.5 MHz subband image are summarized in Table 2. We calculate the dynamical range of each sky map to provide a quantitative estimate of the image quality, which we defined here as the ratio

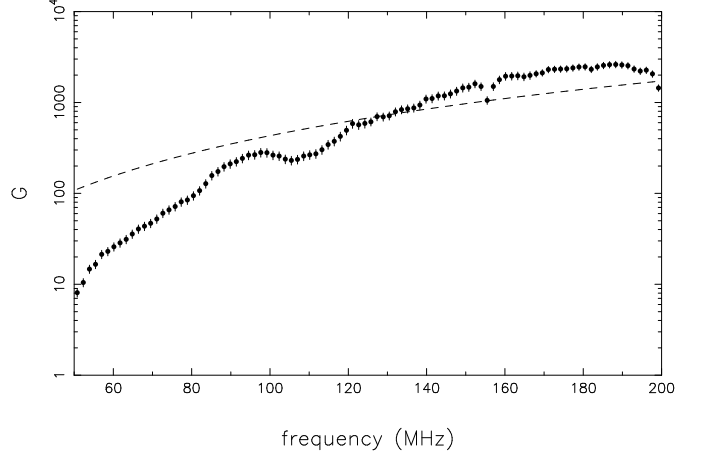


Figure 5. Gain dependence of a 21CMA pod on frequency. Dots represent the best-fit result from the sub-band images with a frequency width of 1.56 MHz. The dashed line shows the gain variation, $G \sim \nu^2$, for a single antenna element in terms of our design in order to guarantee a constant effective area over 50-200 MHz.

Table 2
Image noise and beam size in the 12.5 MHz subband

central frequency (MHz)	image noise (mJy)	beam size (arcmin)
81.25	0.70	6.78
93.75	0.59	5.93
106.25	0.58	5.28
118.75	0.50	4.75
131.25	0.49	4.32
143.75	0.45	3.96
156.25	0.42	3.65
168.75	0.41	3.39

of the peak brightness S_{max} of the image to the rms noise S_{noise} , and the result is shown in Figure 6 for the 8 sky maps. It appears that the ratio of $S_{\text{max}}/S_{\text{noise}}$ reaches a value of $10^4 - 6 \times 10^4$ and shows an increase trend with frequency. Such a variation is simply the combined effect of the system sensitivity, which is dominated by the Milky Way (e.g. Holder 2012) $S_{\text{noise}} \propto T_{\text{sky}} \propto \nu^{-2.6}$, and the power-law behavior of the background radio source flux density, $S_\nu \propto \nu^\alpha$ with $\alpha \approx -0.8$. As a consequence, the dynamical range of the sky map varies with frequency roughly as $S_{\text{max}}/S_{\text{noise}} \propto \nu^{1.8}$. Yet, the current dynamical range is still one order of magnitude below the EoR detection requirement (see Barkana & Loeb 2001; Furlanetto, Oh & Briggs 2006; Pritchard & Loeb 2010 for review), and we will address the possible reasons in the sub-section on noise analysis (Section 3.5) below. We have also made a full band 75-175 MHz image by mosaicing the 8 sub-band images together after convolving to a common resolution of ~ 6 arcmin (Figure 7 and Figure 8). This combined image will be used to remove spurious sources (see below).

3. SOURCE CATALOG AND PROPERTIES

3.1. Source Selection

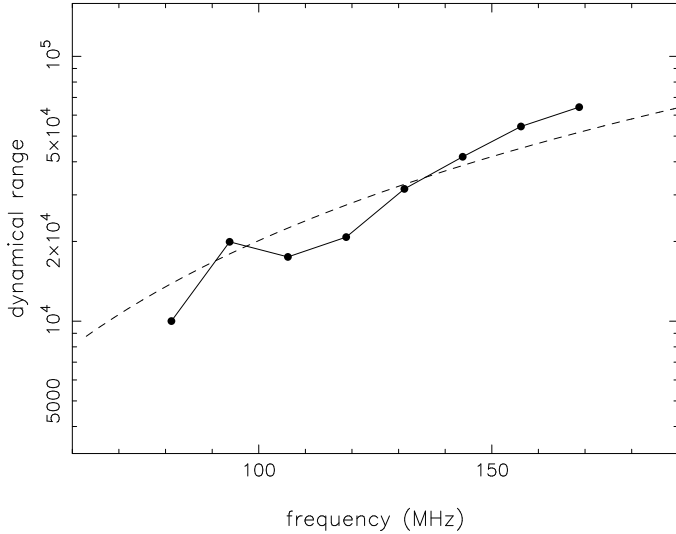


Figure 6. Variation of dynamical ranges with frequencies. The dots denote the estimated ratio of $S_{\max}/S_{\text{noise}}$ from the 8 sky maps centered at 81.25, 93.75, 106.25, 118.75, 131.25, 143.75, 156.25 and 168.75 MHz, respectively. The dashed line represents the theoretically expected variation given by the flux density ratio of the cosmic radio sources ($S_{\nu} \propto \nu^{-0.8}$) to the Milky Way as the noise ($S_{\text{noise}} \propto T_{\text{sky}} \propto \nu^{-2.6}$), for which we have used an arbitrary amplitude.

The field of view (FoV) depends on the observing frequency, such that $\text{FoV} \propto \nu^{-2}$. As a result, we can only extract full spectral information for radio sources within a radius of $\sim 3^\circ$ around the NCP. Nonetheless, we also count the sources between $3^\circ - 5^\circ$ to expand our radio source catalog, for which spectral measurements are based on only the four lower frequency channels, 81.25, 93.75, 106.25 and 118.75 MHz. We note that, many more sources have been found beyond 5° at the lower frequency end, but we will not include them in the present catalog.

Source candidates are first selected in the full-band image covering 75-175 MHz, using a method similar to DUCHAMP (Whiting 2011), in which a threshold of 3σ is adopted to identify source candidates. The position of the peak flux of each candidate in the full-band image is tagged and then used for selection of source candidate in other sub-bands in a procedure similar to that described in Hurley-Walker et al. (2014). The source candidates will be removed from the list if their counterparts are missing in one of the frequency bands. This allows us to identify and thus remove all the spurious sources that result from the sidelobes of the brightest radio source, 3C061.1 (see Figure 7 and Figure 8). Indeed, we are still unable to perfectly CLEAN the sidelobes of the brightest radio galaxy, 3C061.1, located at $3^\circ.71$ degrees from NCP. The complex structure of this source (Lawrence et al. 1996) is marginally resolved with 21CMA. Even with a better angular resolution and a multi-directional calibration using SAGE-CAL (Yatawatta et al. 2009; Kazemi et al. 2011), LO-FAR imaging of the same source also contains significant errors (Yatawatta et al. 2013). Our current experiment towards the detection of the EoR is thus largely limited by the presence of this very bright, complex radio galaxy

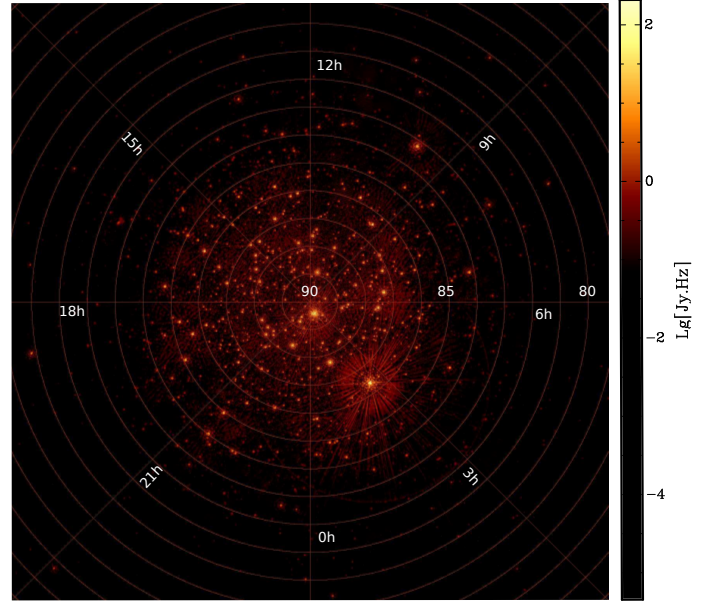


Figure 7. The dirty map centered on the NCP in the frequency range of 75-175 MHz. A larger field around the NCP is shown, with an annulus of 1 degree in radial distance.

in the NCP field.

A Gaussian profile is used to fit the “surface brightness” of a radio source in each of the sub-bands and the corresponding flux density can be computed straightforwardly. When two sources show an overlap due to poor resolutions, a two-component Gaussian profile is adopted to fit the images simultaneously. Yet, even this double Gaussian models may fail to fit two sources with significant overlap especially at low frequencies. In this case, we will not provide the flux density information for the two sources but still treat them as separate sources. We calculate the spectral index for each source using the measured flux densities in all the available bands, up to 8 bands for sources within 3° and up to 4 bands for the ones in the annulus of $3^\circ - 5^\circ$.

Our method of source identification is indeed reliable for mitigation of artificial sources. However, the fainter sources below our threshold at least in one of the eight sub-bands can not be included with the current criterion of source selection. In addition, the threshold varies with frequency channels, and therefore, we may have missed a significant fraction of faint sources. We have thus tested two other methods by relaxing the above criterion for source selection. The first is to work with an extreme case: all the candidates with peak fluxes above 3σ in any sub-band are selected. This provides an overestimate of source population because image artifacts, due to imperfect CLEANing of sidelobes of bright sources, especially 3C061.1, have contaminated our observing field. We demonstrate in Section 3.3 an example of the completeness estimate from such a source selection method for the highest angular resolution sub-band at 168.75 MHz. It turns out that the contamination of spurious sources is indeed very serious. The second test is to select sources based on adjacent frequency channels instead

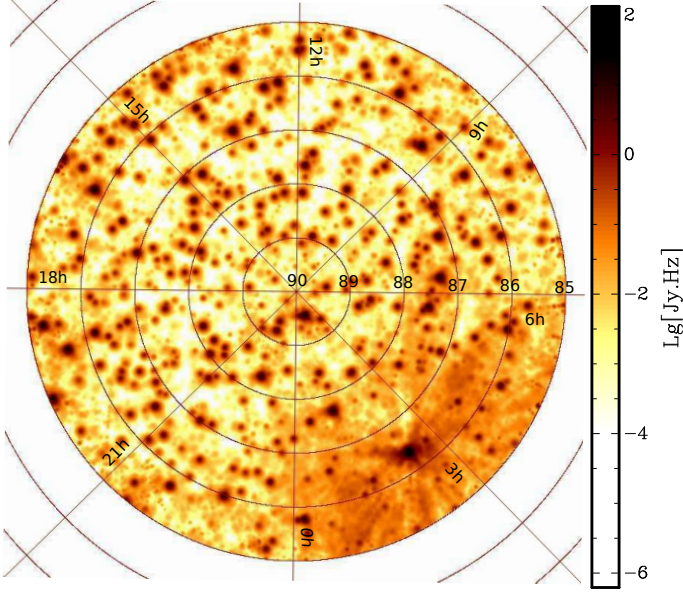


Figure 8. The restored sky map in the frequency range of 75-175 MHz. A central area of $5^\circ.5 \times 5^\circ.5$ is shown, with an annulus of 1 degree in radial distance. The deconvolution is performed through the Högbom CLEAN algorithm with a loop gain of 0.05. The brightest source (lower right), 3C061.1, is located at a distance of $3^\circ46'$ from the NCP, and deconvolution noise around 3C061.1, which arises mainly from grating lobes and inaccurate calibration, is clearly visible. Note that we have enhanced the image contrast to display the noise field which is actually four orders of magnitude fainter than the typical bright sources in the field.

of full bands. To deal with the sidelobes from the bright sources, source candidates should be removed from the list once their counterparts are missing only in one of the adjacent channels, using the frequency-dependence properties of sidelobes (Figure 9). This allows us to identify more faint sources yet leaves the following two drawbacks: (1) the position variations of the sidelobes-induced “sources” become indistinguishable at low frequencies because of poor angular resolution, and (2) the spectral index cannot be estimated reliably if a source shows up only in two frequency bands. While we adopt our conservative way to select sources and construct our source catalog in this work, we will compare the completeness and source counts among the three selection methods below and will improve our source selection algorithm in our future work.

A total of 624 radio sources are found around the NCP region, and their positions, integrated fluxes and spectral indices are listed in Table 3. Note that for sources within 3° we fit integrated fluxes in eight subbands to a power-law to estimate their spectral indices. While for sources in annulus of $3^\circ - 5^\circ$ we only use the integrated fluxes measured in 4 lower frequency subbands. Moreover, data points in some of the subbands have not been taken if the integrated fluxes of images cannot be precisely measured due to poor resolution and/or significant overlap. Of the 624 sources, 322 sources lie between 3 and 5 deg from the field center NCP which are selected using the four lower frequency subbands (81.25, 93.75, 106.25 and 118.75 MHz). An immediate cross-check with the exist-

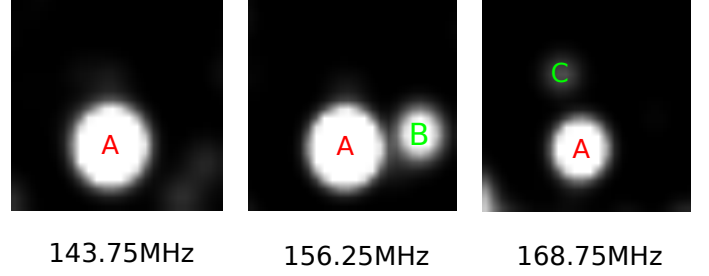


Figure 9. An example of source selection based on adjacent channels: ‘A’ is selected as a cosmic source in the subband 156.25 MHz because it occurs at the same position in at least one of the adjacent channels (143.75 MHz and 168.75 MHz). On the contrary, neither ‘B’ nor ‘C’ will be classified as source candidate because they show up in a single frequency channel only.

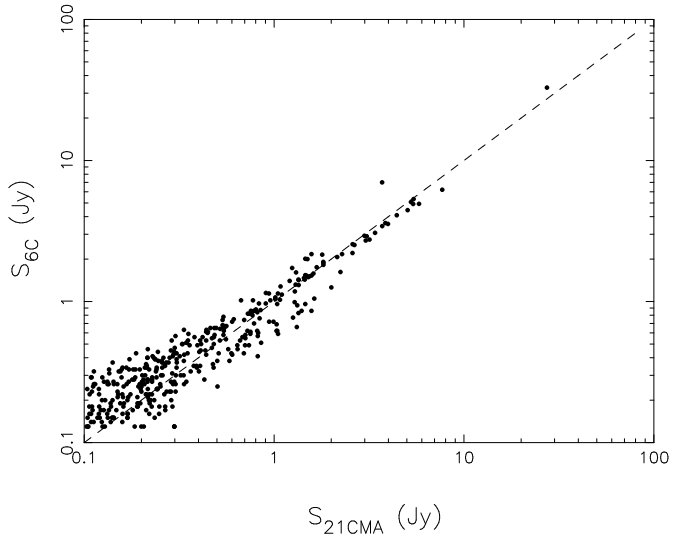


Figure 10. A comparison of flux densities measured at 151 MHz in the 6C survey and the 21CMA observation above 0.1 Jy. Errors on the points are omitted for clarity. The dashed line indicates equal flux density values.

ing the 6C catalog at 151 MHz reveals 490 counterparts. The missing sources are the fainter ones below the detection limit of 0.1 Jy in the 6C survey. Namely, about 134 new sources are detected around and below 0.1 Jy at 151 MHz. Figure 10 compares the flux densities of these 490 sources at 151 MHz listed in both 6C catalog and our new catalog, and two measurements above 0.1 Jy show rather a good agreement.

3.2. Spectral index

Figure 11 illustrates the histogram of the spectral indices of the 624 sources presented here. While the distribution peaks at approximately $\alpha \approx -0.8$, consistent with what is expected for the overall population of cosmic radio sources, the asymmetrical behavior indicates that more sources tend to have a steeper spectral index. This could arise as the radio sources with steeper indices are more easily detected at low frequencies. In

Table 3

Source Catalog. The coordinates, total fluxes at eight frequency bands (if available) and spectral indices are provided for 624 sources detected within 3° (8 sub-bands) and in annulus of $3^\circ - 5^\circ$ (4 lower sub bands) around the NCP. The positional accuracy of the central coordinates is ~ 5 arcmin. Table 3 is published in its entirety in the electronic edition of MRT-sourcecatalog.txt, and a portion is shown here for guidance regarding its form and content.

RA (J2000) (deg)	DEC (J2000) (deg)	$F_{81.25}$ (Jy)	$F_{93.75}$ (Jy)	$F_{106.25}$ (Jy)	$F_{118.75}$ (Jy)	$F_{131.25}$ (Jy)	$F_{143.75}$ (Jy)	$F_{156.25}$ (Jy)	$F_{168.75}$ (Jy)	α
9.49	89.68	$0.033^{+0.005}_{-0.005}$	$0.031^{+0.003}_{-0.003}$	$0.026^{+0.003}_{-0.003}$	$0.025^{+0.003}_{-0.003}$	$0.013^{+0.002}_{-0.002}$	$0.007^{+0.001}_{-0.001}$	$0.015^{+0.003}_{-0.003}$	$0.021^{+0.004}_{-0.004}$	$-1.37^{+0.20}_{-0.20}$
5.23	89.65	$0.391^{+0.055}_{-0.055}$	$0.287^{+0.029}_{-0.029}$	$0.303^{+0.032}_{-0.032}$	$0.339^{+0.037}_{-0.037}$	$0.302^{+0.047}_{-0.047}$	$0.218^{+0.025}_{-0.025}$	$0.195^{+0.036}_{-0.036}$	$0.184^{+0.038}_{-0.038}$	$-0.93^{+0.19}_{-0.19}$
7.59	89.63	...	$0.044^{+0.005}_{-0.005}$	$0.033^{+0.004}_{-0.004}$	$0.026^{+0.003}_{-0.003}$	$0.030^{+0.005}_{-0.005}$	$0.027^{+0.003}_{-0.003}$	$0.038^{+0.007}_{-0.007}$	$0.053^{+0.011}_{-0.011}$	$0.22^{+0.19}_{-0.19}$
0.29	89.58	...	$0.222^{+0.023}_{-0.023}$	$0.245^{+0.026}_{-0.026}$	$0.230^{+0.025}_{-0.025}$	$0.226^{+0.035}_{-0.035}$	$0.271^{+0.031}_{-0.031}$	$0.247^{+0.046}_{-0.046}$	$0.185^{+0.038}_{-0.038}$	$-0.11^{+0.20}_{-0.20}$
12.30	89.57	$0.054^{+0.006}_{-0.006}$	$0.052^{+0.008}_{-0.008}$	$0.047^{+0.005}_{-0.005}$	$0.049^{+0.009}_{-0.009}$	$0.054^{+0.011}_{-0.011}$	$-0.09^{+0.49}_{-0.49}$
16.55	89.55	$0.512^{+0.072}_{-0.072}$	$0.443^{+0.045}_{-0.045}$	$0.548^{+0.058}_{-0.058}$	$0.584^{+0.063}_{-0.063}$	$0.451^{+0.070}_{-0.070}$	$0.417^{+0.047}_{-0.047}$	$0.322^{+0.060}_{-0.060}$	$0.273^{+0.056}_{-0.056}$	$-0.75^{+0.16}_{-0.16}$
1.29	89.47	$7.272^{+1.125}_{-1.125}$	$6.613^{+0.750}_{-0.750}$	$6.450^{+1.230}_{-1.230}$	$7.498^{+1.542}_{-1.542}$	$0.09^{+0.56}_{-0.56}$
9.52	89.47	$0.164^{+0.017}_{-0.017}$	$0.127^{+0.014}_{-0.014}$	$0.150^{+0.023}_{-0.023}$	$0.143^{+0.016}_{-0.016}$	$0.113^{+0.021}_{-0.021}$	$0.113^{+0.023}_{-0.023}$	$-0.69^{+0.29}_{-0.29}$
6.31	89.45	$1.101^{+0.155}_{-0.155}$	$0.896^{+0.092}_{-0.092}$	$0.800^{+0.085}_{-0.085}$	$0.585^{+0.063}_{-0.063}$	$0.586^{+0.091}_{-0.091}$	$0.466^{+0.053}_{-0.053}$	$0.431^{+0.080}_{-0.080}$	$0.374^{+0.077}_{-0.077}$	$-1.48^{+0.15}_{-0.15}$
10.86	89.40	$0.072^{+0.014}_{-0.014}$	$0.066^{+0.007}_{-0.007}$	$0.071^{+0.008}_{-0.008}$	$0.048^{+0.005}_{-0.005}$	$0.070^{+0.011}_{-0.011}$	$0.059^{+0.007}_{-0.007}$	$0.073^{+0.014}_{-0.014}$	$0.057^{+0.012}_{-0.012}$	$-0.16^{+0.16}_{-0.16}$

other words, low-frequency observations would probably miss some of the flat-spectrum sources if they are not bright enough. Yet, recent observations have argued a spectral flattening towards both lower frequencies and fluxes (e.g. Intema et al. 2011; Heywood et al. 2013; van Weeren 2014). To investigate if similar behavior is evident in our radio source catalog, we restrict the sources within 3 degrees and compare in Figure 12 the spectral indices between the 4 lower frequency subbands and the 4 higher frequency subbands. It turns out that there is a significant excess of steep-spectrum sources in the higher frequency subbands. This can be further quantified by considering the means, medians and their ratios (means/medians) of the two distributions, yielding $(-1.25, -1.15, 1.09)$ and $(-0.86, -0.87, 0.99)$ for the higher and lower frequency subbands, respectively. We have also compared the low-frequency sources in the inner radius of 3 degrees and the ones in the outer annulus of $3^\circ - 5^\circ$, and found that their spectral indices are essentially similar (Figure 13). While the absence of bias in spectral index between the sources within the central region ($\leq 3^\circ$) and the ones in the outer region ($3^\circ - 5^\circ$) simply reflects the uniformity of cosmic radio sources, it also indicates that our beam model is correct. It turns out that our analysis does support a flattening trend toward lower frequencies. If confirmed, this may add a further difficulty for low-frequency observations to detect faint, flat-spectrum sources.

3.3. Completeness

Before we perform Monte-Carlo simulations to estimate the completeness of our source selections, we compare our source counts with the deep low-frequency observations of the NOAO Boötes field down to 6 mJy made with GMRT at 153 MHz (Intema et al. 2011). This allows us to make a quick yet rough judgment on the incompleteness of our source counts at different fluxes. Of course, an extrapolation based on a spectral index conversion should be made for source counts at other frequencies for the “control” field. We notice that the reliability of such an extrapolation is probably questionable, as argued recently by Franzen et al. (2016). Taking the differential source count by Intema et al. (2011) at 153

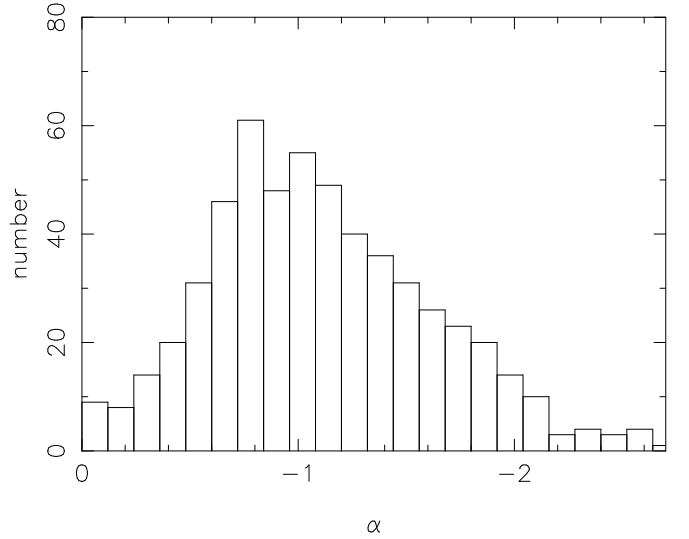


Figure 11. Histogram of spectral indices α ($S \sim \nu^\alpha$) for the 21CMA source catalog.

MHz, $dN/dS = 5370S^{-1.59}$, we calculate the ratios of our observed sources $N(> S)$ to what are expected in terms of this single power law, and plot the results for the 168.75 MHz frequency channel in Figure 14. It appears that the incompleteness for our standard method of source selection occurs around 0.1 Jy, where the ratio drops to $\sim 80\%$. Namely, a significant fraction of the fainter sources may be missed below 0.1 Jy when sources are required to be visible in all subbands to guarantee that the artificial sources from sidelobes be clearly discriminated and perfectly removed. We have also shown the completeness estimates for the other two selection methods: The extreme approach that treats all the peaks above 3σ as “sources” gives rise to an apparent overestimate of true sources even down to ~ 0.01 Jy, indicating that the contamination is rather significant. Sources that are selected in terms of their locations at adjacent frequency channels seem to be complete until ~ 0.1 Jy, and their completeness drops to $\sim 50\%$ at 0.01 Jy. While

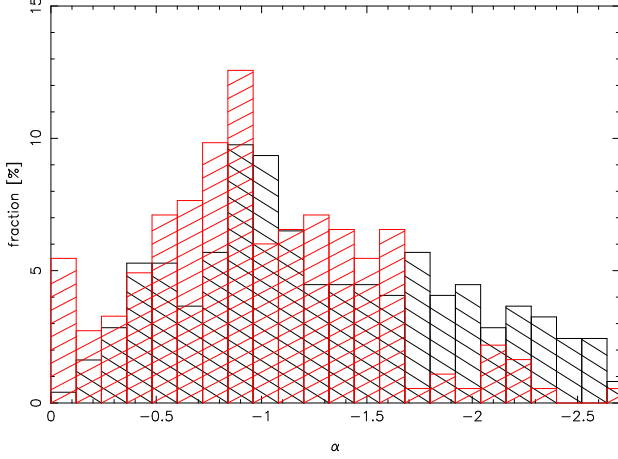


Figure 12. Comparison of spectral indices derived between the high-frequency subbands (black) and lower-frequency ones (red) within 3 degrees. Note that the vertical axis denotes the source fraction

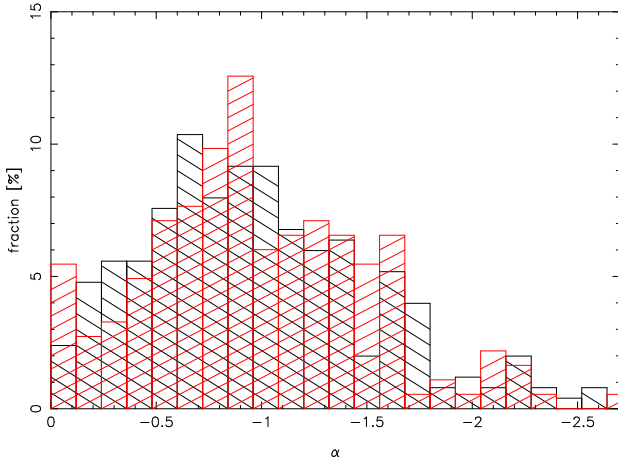


Figure 13. Comparison of spectral indices for sources in the central region of 3° (red) around the NCP and in the outer annulus of $3^\circ - 5^\circ$ (black). The vertical axis denotes the source fraction.

the latter provides a “better” completeness of source selection, the poor angular resolution at lower frequencies may be inefficient for discrimination of position shifts of the artificial sources due to sidelobes, as was pointed out above.

We now perform a Monte-Carlo simulation to quantify the completeness of the source sample. We generate 20 fields with a radius of 3° , each of which contains 1600 point sources with positions randomly distributed. The fluxes of these artificial sources are assumed to follow a power-law of $dN/dS \propto S^{-1.59}$ and are further restricted within $5 \text{ mJy} \leq S \leq 30 \text{ Jy}$, for which a spectral index of -0.8 is assumed. We include the beam smearing effect by convolving the image with a Gaussian function characterized by the true beam at each frequency band. We add noise to the simulated map at each channel using the true

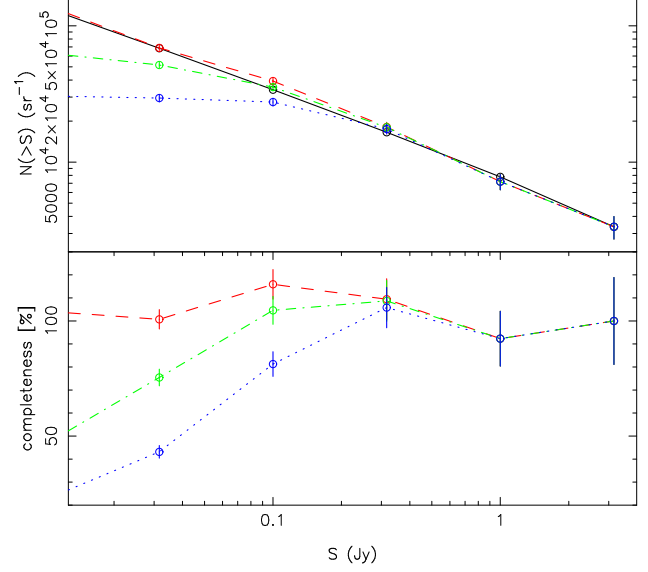


Figure 14. The integrated source counts $N(> S)$ (top) and detection fractions (bottom) estimated using the Intema et al. (2011) deep survey as control field, $dN/dS = 5370S^{-1.59}$ (The integrated form is plotted as a black line). Flux densities are converted to 168.75 MHz by assuming a power law with spectral index of -0.8 , and only a central field with 3 degrees is chosen. (1) Red points: peaks above 3σ in any subbands are taken to be “sources”; (2) Green points: source selection is based on their positions at two adjacent frequency bands; (3) Blue points: source should be visible at all eight sub-bands.

residual map from our deconvolution after point sources are peeled out. Furthermore, the radial variation in sensitivity around the NCP is taken into account by the spatial response of the telescope described by Eq.(2). Now, we adopt the same mechanism, the so-called standard method, to extract sources at the same eight frequency channels, using the image flux and position information over the frequency bands. Figure 15 shows the differential completeness against total flux density for each of the eight channels, respectively, in which the error bars represent the standard statistical deviations. As expected, the turning points of differential completeness exhibit a weak increasing trend towards higher frequencies. Unfortunately, the differential completeness decreases rapidly to zero beyond $S \sim 0.1 \text{ Jy}$, although we can still detect some sources down to $S \sim 0.01 \text{ Jy}$ for all the 8 frequency bands. Our estimated differential completeness will be used for correcting the incompleteness of source counts below.

3.4. Source Counts

We plot in Figure 16 the differential source counts dN/dS of our sample over the central field of radius 3° for each of the 8 frequency bands, in which both the raw data and completeness-corrected data in terms of the above Monte-Carlo simulation are displayed. For data points with completeness of 50% or above, which corresponds to the brighter sources with fluxes greater than $\sim 0.1 \text{ Jy}$, a single power law fits nicely the completeness-corrected data:

$$\frac{dN}{dS} = kS^{-\gamma} \text{Jy}^{-1} \text{sr}^{-1}. \quad (3)$$

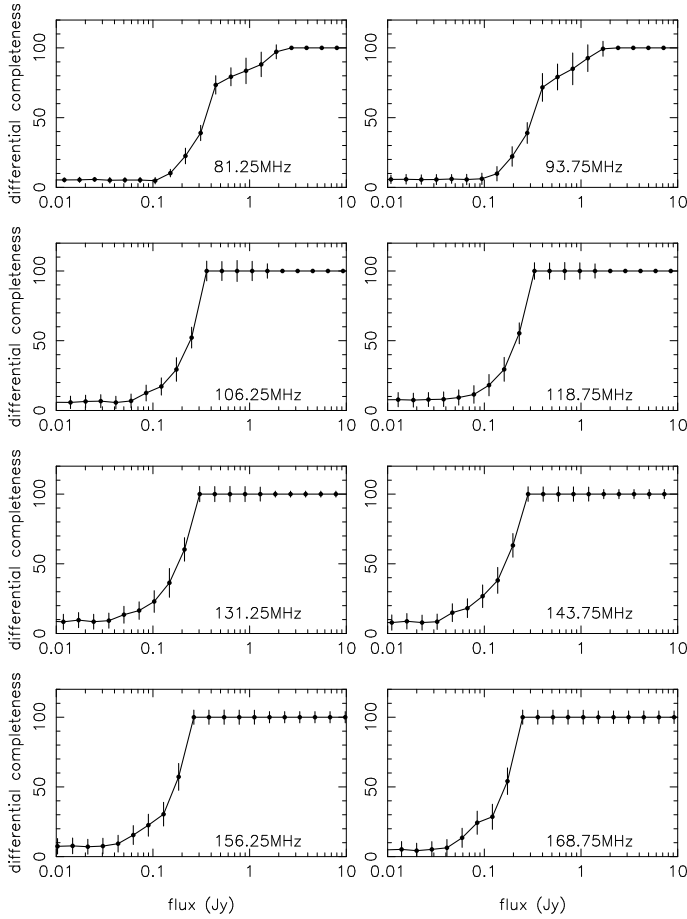


Figure 15. The estimated differential completeness against total flux density for the eight frequency subbands. A total of 1600 artificial sources following $dN/dS \propto S^{-1.59}$ are randomly distributed over a circular field with a radius of 3 degrees, and noise and beam smearing are added in terms of the true measurements on the NCP field. The same source selection method as we adopted for the NCP field is applied to the simulation. Error bars at each frequency represent the standard deviation from 20 realizations.

If we adopt a lower completeness of 20% to include more fainter sources, though the data points below 50% already become unreliable, dN/dS would follow the well-known two power laws for cosmic radio population at low frequencies (Hales, Baldwin & Warner 1988; Cohen 2007; Moore et al. 2013).

$$\frac{dN}{dS} = \begin{cases} k_1 S^{-\gamma_1} \text{Jy}^{-1} \text{sr}^{-1}, & S < S_0; \\ k_2 S^{-\gamma_2} \text{Jy}^{-1} \text{sr}^{-1}, & S \geq S_0 \end{cases} \quad (4)$$

where S_0 indicates the turnover flux which will be fixed to $S_0 = 0.8$ Jy in the fitting. The best-fit parameters for the 8 frequency subbands are summarized in Tables 4 and 5 for the 50% and 20% completeness data, respectively. It appears that for the bright population with $S > S_0$, our best-fit indices vary between 2.10 and 2.55, which are both consistent with each other within error bars, and are also close to 2.5 which is expected for a Euclidean cosmology. For the fainter radio sources below S_0 , we find a weak dependence of γ_1 on frequency, with a flattening source count towards lower frequencies. We note

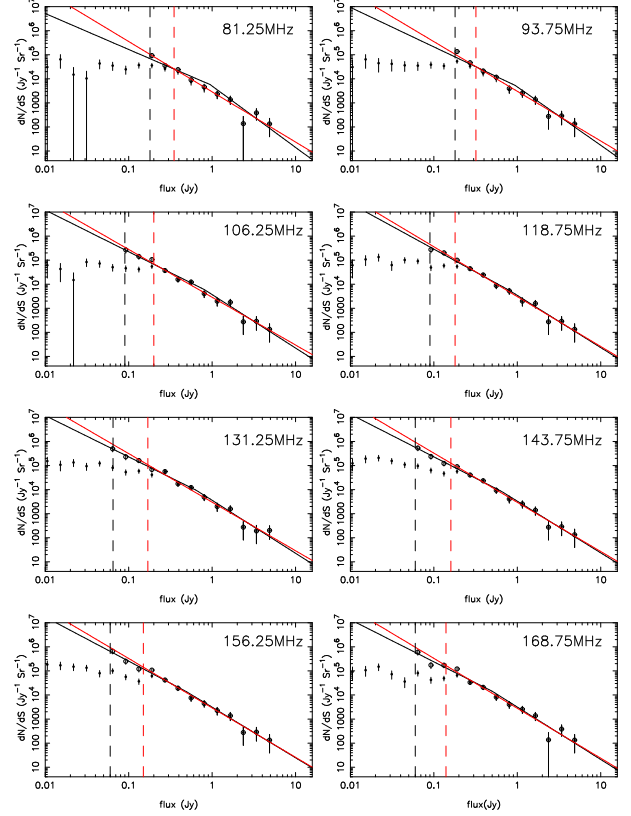


Figure 16. Differential source counts over the central field within a radius 3° for each of the eight frequency subbands. Filled and open circles represent the raw data and completeness-corrected data, respectively. The two vertical dashed lines indicate the 20% (black) completeness and 50% (red) completeness, respectively. The best-fit two power laws for the 20% completeness limit are shown by the black solid lines, while the results for the 50% completeness data are illustrated by the red solid lines.

that a power-law fit at $S < 0.4$ Jy to the GMRT counts from Williams, Intema & Rottgering (2013), Intema et al. (2011) and Ghosh et al. (2012) gives $\gamma_1 = 1.54$ at 154 MHz (Franzen et al. 2016) and the GMRT targeted survey at 153 MHz finds essentially similar results, $\gamma_1 = 1.59$ (Intema et al. 2011, see also Ghosh et al. 2012). Our best-fit power indices at the two closest bands, 143.75 and 156.25 MHz, are 1.77 ± 0.06 and 1.83 ± 0.10 , respectively, which are slightly higher than these results due to the lack of fainter sources in our samples.

To compare our results with other deep surveys at low frequencies, we derive the Euclidean-normalized differential source counts, $S^{5/2} dN/dS$. We take the source counts from two recent surveys at frequencies in the range of 151-154 MHz with GMRT (Intema et al. 2011; Ghosh et al. 2012) and MWA (Franzen et al. 2016). We choose our results in the 156.25 MHz band to compare with $S^{5/2} dN/dS$ given at 151-154 MHz with the GMRT and MWA by scaling the observing frequencies with a single power-law of $S \propto \nu^{-0.8}$ to 154 MHz (Figure 17). Within the error bars, our source counts are in good agreement with these recent surveys. In Table 6 and Table 7 we have provided the tabulated normalized differential source counts for the lower (81.25, 93.75, 106.25

Table 4
The best-fitted parameters for differential source counts (50% completeness).

ν_c (MHz)	k	γ
81.25	2896 ± 898	2.09 ± 0.22
93.75	3044 ± 737	2.06 ± 0.21
106.25	2879 ± 635	2.01 ± 0.24
118.75	3062 ± 979	2.07 ± 0.32
131.25	3007 ± 679	2.02 ± 0.28
143.75	2974 ± 789	2.05 ± 0.30
156.25	2911 ± 757	2.05 ± 0.33
168.75	2812 ± 464	2.07 ± 0.28

Table 5
The best-fitted parameters for differential source counts (20% completeness).

ν_c (MHz)	S_0 (Jy)	k_1	γ_1	k_2	γ_2
81.25	0.95	5318 ± 1384	1.49 ± 0.15	5038 ± 1180	2.55 ± 0.19
93.75	0.95	4850 ± 928	1.64 ± 0.08	4658 ± 615	2.43 ± 0.13
106.25	0.8	4136 ± 520	1.73 ± 0.07	3708 ± 355	2.22 ± 0.13
118.75	0.8	3664 ± 441	1.89 ± 0.08	3447 ± 327	2.17 ± 0.12
131.25	0.65	4432 ± 444	1.71 ± 0.06	3594 ± 276	2.20 ± 0.12
143.75	0.65	3917 ± 455	1.77 ± 0.06	3327 ± 240	2.15 ± 0.16
156.25	0.5	3747 ± 613	1.83 ± 0.10	3112 ± 341	2.10 ± 0.12
168.75	0.5	4269 ± 1167	1.75 ± 0.13	3170 ± 344	2.18 ± 0.18

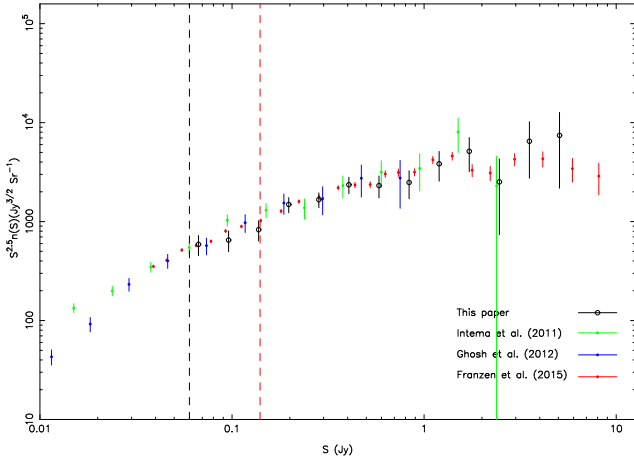


Figure 17. Euclidean normalized differential counts from this paper at 154 MHz (black points). Green and blue points are the GMRT observations at 153 MHz and 150 MHz, respectively (Intema et al. 2011; Ghosh et al. 2012). Red points are recent MWA results at 154 MHz (Franzen et al. 2016). The two vertical dashed lines represent 20% (black) and 50% (red) completeness limits, respectively.

and 118.75 MHz) and higher (131.25, 143.75, 156.25 and 168.75 MHz) bands, respectively, and the results are also displayed in Figure 18 and Figure 19, respectively.

3.5. Noise Analysis

We now analyze the noise in our interferometric imaging towards the NCP region, which comprises primarily four contributors: thermal noise, confusion noise, cali-

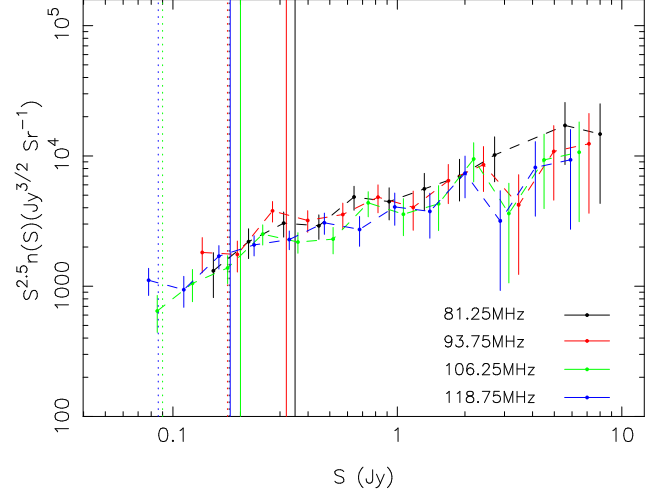


Figure 18. Euclidean normalized differential counts for the lower frequency bands (81.25, 93.75, 106.25 and 118.75 MHz). The vertical lines represent 20% (dotted) and 50% (solid) completeness limits.

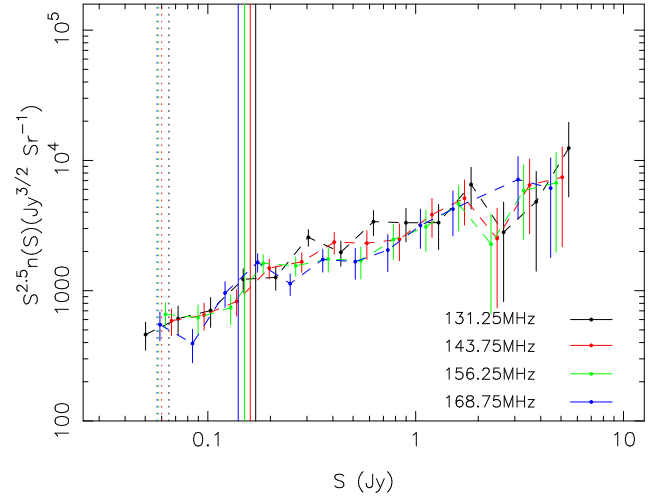


Figure 19. Euclidean normalized differential counts for the higher frequency bands (131.25, 143.75, 156.25 and 168.75 MHz). The vertical lines represent 20% (dotted) and 50% (solid) completeness limits.

bration error and deconvolution error.

The thermal noise for a radio interferometer like 21CMA with an effective area A_{eff} and system temperature T_{sky} can be estimated simply from

$$S_{\text{thermal}} = \frac{\sqrt{2} k_B T_{\text{sys}}}{A_{\text{eff}} \eta \sqrt{\Delta \nu \Delta t}}, \quad (5)$$

where k_B is Boltzmann's constant, $\Delta \nu$, Δt and η are the frequency bandwidth, observation duration, and efficiency factor, respectively. T_{sys} is composed of the telescope temperature (50K for the 21CMA) and the sky (the Milky Way) temperature described approximately by $60(\nu/300\text{MHz})^{-2.55}$ (e.g. Chapman et al. 2012). For the observation made in this work, we have only used 559

Table 6

Normalized differential counts in the lower frequency bands (81.25, 93.75, 106.25 and 118.75 MHz) from the 21CMA NCP field. N_{raw} is the raw number counts in each frequency band.

Bin _{start} (Jy)	Bin _{end} (Jy)	N_{raw} correction factor	Normalized counts (Jy ^{3/2} sr ⁻¹)	Completeness
81.25 MHz				
0.367	0.527	22	2910±620	1.3
0.527	0.755	23	4841±1009	1.2
0.755	1.083	13	4462±1237	1.1
1.083	1.554	10	5587±1767	1.0
1.554	2.228	8	6964±2462	1.0
2.228	3.196	7	10172±3844	1.0
4.584	6.575	4	17150±8575	1.0
6.575	9.430	2	14730±10415	1.0
93.75 MHz				
0.327	0.470	28	3198±604	1.3
0.470	0.674	20	3553±794	1.2
0.674	0.966	17	4830±1171	1.1
0.966	1.386	9	4030±1343	1.0
1.386	1.987	9	6453±2151	1.0
1.987	2.850	7	8567±3238	1.0
2.850	4.088	2	4204±2973	1.0
4.088	5.863	3	10833±6254	1.0
5.863	8.410	2	12405±8772	1.0
106.25 MHz				
0.207	0.296	32	2513±444	1.9
0.296	0.425	31	2183±392	1.0
0.425	0.609	19	2299±527	1.0
0.609	0.874	21	4364±952	1.0
0.874	1.254	10	3570±1129	1.0
1.254	1.798	7	4292±1622	1.0
1.798	2.579	9	9479±3160	1.0
2.579	3.699	2	3618±2558	1.0
3.699	5.305	3	9322±5382	1.0
5.305	7.608	2	10675±7459	1.0
118.75 MHz				
0.189	0.271	32	2075±367	1.8
0.271	0.389	37	2280±375	1.0
0.389	0.557	29	3070±570	1.0
0.557	0.800	15	2728±704	1.0
0.800	1.147	13	4061±1126	1.0
1.147	1.645	7	3756±1420	1.0
1.645	2.359	8	7373±2606	1.0
2.359	3.384	2	3166±2239	1.0
3.384	4.853	3	8158±4710	1.0
4.853	6.961	2	9341±6605	1.0

baselines among 780 along the E-W arm after the shorter and longer baselines are taken out. This reduces the total effective area to $A_{\text{eff}} = \sqrt{559}a_{\text{eff}}$, in which $a_{\text{eff}} \approx 216m^2$ is the effective area of a single pod which remains roughly constant over our observing frequency range in terms of our antenna design. We also adopt the current efficiency of $\eta = 50\%$ for 21CMA to provide the numerical estimation.

Confusion noise σ_c from the position uncertainties of unresolved, faint radio sources below a flux threshold S_{lim} with finite synthesis beam Ω_b can be estimated through (e.g. Scheuer 1957; Condon 1974)

$$\sigma_c^2 = \Omega_b \int_0^{S_{\text{lim}}} \frac{dN}{dS} S^2 dS. \quad (6)$$

In principle, we could take a simple extrapolation of our current differential source counts dN/dS to lower fluxes beyond S_{lim} by considering a lower completeness of our catalog especially at lower frequencies. Here we would rather choose the deeper GMRT observations at

153 MHz (Intema et al. 2011) and extrapolate their differential number counts to $S_{\text{lim}} = 0.1$ mJy, two orders of magnitude fainter than their flux limit. Yet, uncertainty in this extrapolation to very low flux remains unclear. Moreover, it seems that the steep power law for the number count may flatten towards lower flux below a few mJy (see Franzen et al. 2016 for a recent compilation).

Figure 20 plots the theoretically predicted thermal noise and confusion noise, together with the image noise measured across all eight bands. We have set $S_{\text{lim}} = 3\sigma$ in the estimate of confusion limit in each of the subbands. Thermal noise for 12 hour observations and a bandwidth of 12.5 MHz has already become smaller than the image noise, indicating that the key factor influencing the sensitivity of low-frequency interferometric imaging is no longer the system noise. Indeed, the confusion limit arising mainly from the unsolved faint sources especially at lower frequencies has actually made a significant contribution to the noise level. Unlike the thermal noise which decreases as $t^{-1/2}$, the confusion noise is very insensitive to observing time t . Instead, the upper limit of integral in

Table 7

Normalized differential counts in the higher frequency bands (131.25, 143.75, 156.25 and 168.75 MHz) from the 21CMA NCP field. N_{raw} is the raw number counts in each frequency band.

Bin _{start} (Jy)	Bin _{end} (Jy)	N_{raw} correction factor	Normalized counts (Jy ^{3/2} sr ⁻¹)	Completeness factor
131.25 MHz				
0.174	0.250	24	1266±258	1.6
0.250	0.359	47	2569±375	1.0
0.359	0.515	21	1972±430	1.0
0.515	0.738	21	3387±739	1.0
0.738	1.059	12	3324±960	1.0
1.059	1.518	7	3331±1259	1.0
1.518	2.178	8	6538±2312	1.0
2.178	3.123	2	2807±1985	1.0
3.123	4.480	2	4823±3410	1.0
4.480	6.425	3	12427±7174	1.0
143.75 MHz				
0.162	0.233	33	1489±259	1.5
0.233	0.334	34	1666±286	1.0
0.334	0.478	28	2357±445	1.0
0.478	0.686	16	2313±578	1.0
0.686	0.984	10	2484±785	1.0
0.984	1.411	9	3839±1279	1.0
1.412	2.025	7	5129±1939	1.0
2.025	2.904	2	2517±1780	1.0
2.904	4.165	3	6486±3745	1.0
4.165	5.974	2	7428±5252	1.0
156.25 MHz				
0.152	0.218	36	1624±271	1.7
0.218	0.312	35	1551±262	1.0
0.312	0.448	23	1752±365	1.0
0.448	0.642	13	1701±472	1.0
0.642	0.921	11	2472±745	1.0
0.921	1.321	8	3088±1092	1.0
1.321	1.894	7	4641±1754	1.0
1.894	2.717	2	2278±1611	1.0
2.717	3.896	3	5869±3388	1.0
3.896	5.589	2	6720±4752	1.0
168.75 MHz				
0.142	0.205	38	1653±268	1.8
0.205	0.293	28	1132±214	1.0
0.293	0.421	25	1736±347	1.0
0.421	0.604	14	1670±446	1.0
0.603	0.866	10	2049±648	1.0
0.866	1.242	9	3167±1056	1.0
1.242	1.781	7	4232±1599	1.0
2.554	3.664	4	7135±3567	1.0
3.664	5.255	2	6128±4333	1.0

Eq.(6), S_{lim} , now determines the variation trend in confusion limit. One of our primary tasks in low-frequency interferometric imaging is to suppress the low flux limit S_{lim} to allow more fainter sources to be identified and resolved.

In fact, S_{lim} is controlled by both calibration error and deconvolution error, in addition to telescope resolution. Here the current imaging algorithm has not taken ionospheric effects into account, though self-calibration may partially correct the effect if a timestep is chosen to be within a few minutes or only short baselines are used. Performing calibration on a timescale of a few minutes turns out to be very time consuming since the correction should also be made at each frequency channel. Moreover, the grating lobes due to numerous redundant baselines of the 21CMA generate very prominent, equally spaced rings round bright sources such as B004713+891245 and 3C061.1. Complete removals of these ring-like structures in deconvolution processing is

still difficult at present. First, an accurate sky model containing both bright point and extended sources should be constructed in the calibration processing. Significant uncertainties arise from the inaccurate models of (marginally) resolved extended sources, in which surface brightness, spectral index and position of each component of the sources should be determined by other radio observations or left as free parameters in the sky model. The brightest radio galaxy, 3C061.1, in the NCP field is a typical example of this kind, and its residuals after deconvolution are very prominent. Actually, 3C061.1 is the dominant source of both calibration and deconvolution errors for our imaging towards the NCP region. Note that calibration error is also present even if a more sophisticated model for 3C061.1 is adopted for LOFAR data (Yatawatta et al. 2013). Second, the deconvolution should be made over a sky map with the field-of-view much larger than the primary beam so that the leakage of sidelobes from off-field strong radio sources can

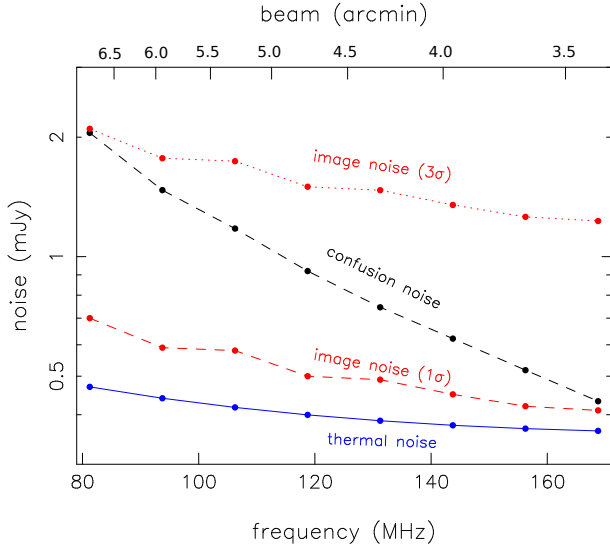


Figure 20. Sensitivity and noise as a function of observing frequencies for a bandwidth of 12.5 MHz. The blue and solid line denotes the expected thermal noise of the 21CMA for a 12 hour integration. Red points, together with dashed and dotted lines, are the 1σ and 3σ noise levels, respectively, measured at eight sub-bands over the NCP field. The black dashed line represents the theoretically estimated confusion limit due to the unresolved fainter sources below S_{lim} , for which we have taken $S_{\text{lim}} = 3\sigma$.

be corrected for. For the 21CMA field, the two brightest radio sources, Cas A and Cygnus A, are visible on the sky map of $60^\circ \times 60^\circ$ (Huang et al. 2016), and their grating lobes have already entered into the central NCP field at certain frequencies. At present the imaging has not included and corrected the sidelobes from these two sources, although they have already contributed to the deconvolution noise.

In summary, it appears that with time integration the thermal noise dominated by the Milky Way can be well controlled to below the confusion noise in current low-frequency interferometric imaging. However, to beat down the confusion noise, one needs to deal with not only the unresolved, faint radio sources, but also the calibration and deconvolution errors. In particular, calibration errors from both ionospheric effects and bright, extended sources are difficult to handle at present. These errors can propagate to the deconvolution processing and affect the lower threshold of identifying the faint radio sources through S_{lim} . As a result, the conventional confusion limit, which is determined by unresolved sources and synthesized beam, is actually affected by the extension of accuracy and reliability of calibration and deconvolution. In fact, the major source of errors in our current imaging arises from the calibration and deconvolution processing. The former is related to both ionospheric effects and modeling of extended, bright radio sources in the field, and the latter is involved with the removals of grating lobes and sidelobes of strong radio sources, both within and outside of the field.

4. DISCUSSION AND CONCLUSIONS

We have analyzed 12 hours of data taken from the 21CMA observations centered on the NCP. To reduce the

complexity of interferometric image processing and the influence of ionospheric perturbation, we have used all 40 pods along the east-west arm but restricted our analysis to within a maximum baseline of 1500 m. We have calibrated the gain and primary beam of our telescope using the bright radio sources in the NCP field. Following conventional self-calibration and deconvolution methods, we have detected a total of 624 radio sources over the central field within 3° in frequency range of 75 - 175 MHz and the outer annulus of $3^\circ - 5^\circ$ in the 75 - 125 MHz bands. By performing a Monte-Carlo simulation we have estimated a completeness of 50% at $S \sim 0.2$ Jy. We have compared our source counts with the deep low-frequency observations made recently with GMRT and MWA, and the completeness-corrected source counts show a good agreement with these recent surveys at the corresponding frequency bands.

While we are able to detect the fainter sources down to 10 mJy, the detection fraction has dropped to $\sim 10\%$, making a ultra deep observation rather difficult and even impossible with current imaging algorithms. This is primarily caused by errors in sidelobes, calibration, and deconvolution as well as the confusion limit of the array. Spacings between the 21CMA pods are the integral multiples of 20 m. The original design of such configuration is for the purpose of redundant calibration and statistical measurement of the EoR power spectrum (e.g. Noordam & de Bruyn 1982; Tegmark & Zaldarriaga 2010). However, this layout also generates the grating lobes evident in our images (Bracewell & Thompson 1973; Amy & Large 1990). These equally spaced rings around the bright sources in the field have their brightness almost comparable to the central sources, and therefore should be accurately removed in the deconvolution processing. Otherwise, their residuals may be the major sources of contaminations for our imaging. Indeed, our image quality is largely limited by the presence of imperfectly subtracted grating lobes of the brightest source in the NCP field, 3C061.1. Although this arises partially from the inaccurate calibration, the main reason is the poor modeling of 3C061.1 in our self-calibration processing. 3C061.1 is marginally resolved with 21CMA, and its modeling should take the structured components with different surface brightness and spectral index into account (e.g. Yatawatta et al. 2013). Our next step is to improve our sky model by using a sophisticated model for 3C061.1 and even for very bright, far field sources such as Cas A and Cygnus A. Finally, the theoretically expected thermal noise (~ 0.4 mJy) is well below the confusion limit due to the unresolved, fainter sources, indicating that the dominated factor of influencing the sensitivity for low-frequency interferometric imaging is no longer the system noise but the cosmic sources in the sky.

Several methods to improve our imaging algorithm including an accurate calibration (e.g. sky modeling and redundant calibration), application of the w-term correction, and “peeling” sources (Noordam 2004) are currently under investigation, with our ultimate goal towards detecting the EoR power spectrum. Our experience with the 21CMA operation and data analysis may provide a useful guide to the design of next generation low-frequency radio array such as the SKA. Indeed, employment of numerous redundant baselines helps to im-

prove the precision of calibration. However, it brings about very prominent grating lobes from bright sources both in-beam and in the far-field. This not only requires a more careful and accurate calibration of visibilities but also adds an extra difficulty to deconvolution processing. In particular, the grating lobes of the bright, structured sources like giant radio galaxies with jets are hard to model and thus subtract. In this sense, both random antenna element and station layouts should preferentially be chosen for low-frequency radio interferometers to suppress the grating lobes and sidelobes of bright sources. We note that an extensive study on this topic has recently been carried out by Mort et al. (2016).

To summarize, with current layout of the 21CMA and conventional imaging algorithm, we have reached a sensitivity of a few mJy and a dynamical range of 10^4 . We may have to improve our imaging quality by an order of magnitude in order to see statistically the EoR signature.

5. ACKNOWLEDGMENTS

We thank Judd Bowman for useful discussions. We also thank Heinz Andernach for comments. We gratefully acknowledge the constructive suggestions by an anonymous referee that greatly improved the presentation of this work. Technical support was provided by the 21CMA collaboration. This work was partially supported by the National Science Foundation of China under Grant No. 11433002. QZ and MJ-H are supported in this work through a Marsden Fund grant to MJ-H in New Zealand. The 21CMA is jointly operated and administrated by National Astronomical Observatories of China and Center for Astronomical Mega-Science, Chinese Academy of Sciences. This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

REFERENCES

- Amy, S. W., & Large, M. I. 1990, *Astronomical Society of Australia*, 8, 308
- Barkana, R., & Loeb, A. 2011, *PhR*, 349, 125
- Bernardi, G., de Bruyn, A. G., Harker, G., et al. 2010, *A&A*, 522, A67
- Bowman, J. D., Barnes, D. G., Briggs, F. H., et al. 2007, *AJ*, 133, 1505
- Bowman, J. D., Cairns, I., Kaplan, D., et al. 2013, *PASA*, 30, 31
- Bracewell, R. N., & Thompson, A. R. 1973, *ApJ*, 182, 77
- Chapman, E., Abdalla, F. B., Harker, G., et al. 2012, *MNRAS*, 423, 2518
- Cohen, A. S., Lane, W. M., Cotton, W. D., et al. 2007, *AJ*, 134, 1245
- Condon, J. J. 1974, *ApJ*, 188, 279
- Franzen, T. M. O., Jackson, C. A., Offringa, A. R., et al. *MNRAS*, 459, 3314
- Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, *PhR*, 433, 181
- Ghosh, A., Bharadwaj, S., Ali, S. S., & Chengalur, J. N. 2011, *MNRAS*, 418, 2584
- Ghosh, A., Prasad, J., Bharadwaj, S., Ali, S. S., & Chengalur, J. N. 2012, *MNRAS*, 426, 3295
- Hales, S. E., Baldwin, J. E., & Warner, P. J. 1988, *MNRAS*, 234, 919
- Hales, S. E. G., Riley, J. M., Waldram, E. M., Warner, P. J., Baldwin, J. E. 2007, *MNRAS*, 382, 1639
- Heald, G. H., Pizzo, R. F., Orrù, E., et al. 2015, *arXiv:1509.01257*
- Heywood, I., Jarvis, M. J., & Condon, J. J. 2013, *MNRAS*, 432, 2625
- Högbom, J. A. 1974, *A&AS*, 15, 417
- Holder, J. 2012, *Astroparticle Physics*, 39, 61
- Huang, Y., Wu, X. -P., Zheng, Q., Gu, J. -H., & Xu, H. 2016, *RAA*, 16, 16
- Hurley-Walker, N., Morgan, J., Wayth, R. B., et al. 2014, *PASA*, 31, 45
- Intema, H. T., van Weeren, R. J., Röttgering, H. J. A., & Lal, D. V. 2011, *A&A*, 535, A38
- Intema, H. T., Jagannathan, P., Mooley, K. P., Frail, D. A. 2016, *arXiv:1603.04368*
- Jacobs, D. C., Aguirre, J. E., Parsons, A. R., et al. 2011, *ApJ*, 734, L34
- Jelić, V., Zaroubi, S., Labropoulos, P., et al. 2008, *MNRAS*, 389, 1319
- Jelić, V., de Bruyn, A. G., Mevius, M., et al. 2014, *A&A*, 568, A101
- Kazemi, S., Yatawatta, S., Zaroubi, S., et al. 2011, *MNRAS*, 414, 1656
- Lawrence, C. R., Zucker, J. R., Readhead, A. C. S., et al. 1996, *ApJS*, 107, 541
- Massaro, F., Giroletti, M., D’Abrusco, R., et al. 2014, *ApJS*, 213, 3
- Moore, D. F., Aguirre, J. E., Parsons, A. R., Jacobs, D. C., & Pober, J. C. 2013, *ApJ*, 769, 154
- Mort, B., Dulwich, F., Razavi-Ghods, N., de Lera Acedo, E., & Grainge, K. 2016, *arXiv:1602.01805*
- Noordam, J. E., & de Bruyn, A. G. 1982, *Nature*, 299, 597
- Noordam, J. E. 2004, *Ground-based Telescopes*. Edited by Oschmann, Jacobus M., Jr. *Proceedings of the SPIE*, 5489, 817
- Offringa, A. R., Wayth, R. B., Hurley-Walker, N., et al. 2015, *PASA*, 32, 8
- Offringa, A. R., Trott, C. M., Hurley-Walker, N., et al. 2016, *MNRAS*, 459, 3314
- Paciga, G., Albert, J. G., Bandura, K., et al. 2013, *MNRAS*, 433, 639
- Pritchard, J. R., & Loeb, A. 2010, *PhRvD*, 82, 3006
- Rees, N. 1990, *MNRAS*, 244, 233
- Scheuer, P. A. G. 1957, *PCPS*, 53, 764
- Slee, O. B. 1995, *Australian Journal of Physics*, 48, 143
- Taylor, G. B., Ellingson, S. W., Kassim, N. E., et al. 2012, *Journal of Astronomical Instrumentation*, 1, 50004
- Tegmark, M., & Zaldarriaga, M. 2010, *PhRvD*, 82, 103501
- Thyagarajan, N., Jacobs, D. C., Bowman, J. D., et al. 2015, *ApJ*, 804, 14
- Tingay, S. J., Geoke, R., Bowman, J. D., et al. 2013, *PASA*, 30, 7
- van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, *A&A*, 556, A2
- van Weeren, R. J., Williams, W. L., Tasse, C., et al. 2014, *ApJ*, 793, 82
- Wayth, R. B., Lenc, E., Bell, M. E., et al. 2015, *PASA*, 32, 25
- Whiting, M. T. 2011, *MNRAS*, 421, 3242
- Williams, W. L., Intema, H. T., & Röttgering, H. J. A. 2013, *A&A*, 549, A55
- Yatawatta, S., Zaroubi, S., de Bruyn, G., Koopmans, L., Noordam, J. 2009, *arXiv:0810.5751*

Yatawatta, S., de Bruyn, A. G., Brentjens, M. A., et al. 2013, A&A, 550, A136