CrossMark

North Celestial Region Observed with 21 CentiMeter Array

Bi-Xuan Zhao[1], Qian Zheng[1], Huan-Yuan Shan[1], Quan Guo[1], and Kuan-Jun Li[2]

[1] Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China; zhaobixuan@shao.ac.cn

[2] National Astronomical Observatory, Chinese Academy of Sciences, 20 Datun Road, Beijing 100101, China

Abstract

The 21 CentiMeter Array (21CMA) is a radio interferometer in western China. The key science objective of 21CMA is statistically measuring the redshifted H I 21 cm signal of the Epoch of Reionization (EoR). In order to achieve this, a catalog of radio sources is necessary for both calibration and foreground removal. In this work, we present a catalog of 832 radio sources within a radius of 5° around the North Celestial Pole region observed with 21CMA. Three days of data taken in 2013 are used in the analysis, with a frequency coverage from 75 to 200 MHz and an angular resolution of about two arcminutes at 200 MHz. Flux densities and estimated spectral indices of the radio sources are provided in the source catalog. For simplicity, only the east–west baseline is used for the observations to avoid the so-called $w-$term in the data analysis. The longest baseline along the east–west direction is 2780 m. Comparing our source counts with other radio observations, we find a good agreement with the observations made with the Giant Metrewave Radio Telescope and Murchison Widefield Array. Our data reduction is still limited by grating lobes of very bright sources in the field due to the redundant configuration of 21CMA. We note that understanding the properties of radio sources, and modeling and removing the radio sources are challenges for the EoR experiments.

Key words: radio continuum: general – instrumentation: interferometers – radio continuum: galaxies – methods: observational

1. Introduction

Understanding the statistical signatures of extragalactic sources at low frequency radio band is required and motivated by many science goals including the detection of the Epoch of Reionization (EoR). Radio sources observed below 300 MHz are one of the dominant contaminations for detecting EoR signals. Modeling and subtracting the sources accurately are required. Furthermore, the incompleteness of the detected radio sources leads to inaccuracy in the sky model, and introduces errors in calibration and deconvolution.

A number of steep-spectrum sources are too faint to be detected at high frequencies, therefore low-frequency observations become important for detecting the faint steep-spectrum sources (Massaro et al. 2014). Current low frequency surveys (below 300 MHz) are limited to the confusion effects caused by unresolved sources, but this limitation is expected to be improved with high sensitivity and long-baseline radio interferometers such as the upcoming Square Kilometre Array (SKA). Along with the fainter sources that can be resolved, foreground radio sources can be more accurately modeled and subtracted in the EoR detection.

Over the past few years, there have been an increasing number of surveys and targeted observations being performed at low frequencies (e.g., Rees 1990; Hales et al. 2007; Cohen et al. 2007; Jacobs et al. 2011; Hurley-Walker et al. 2014).

Low-frequency radio interferometers, such as Murchison Widefield Array (MWA, Bowman et al. 2013; Tingay et al. 2013), Low Frequency Array (LOFAR, van Haarlem et al. 2013), Giant Metrewave Radio Telescope (GMRT, Paciga et al. 2013), Long Wavelength Array (LWA, Taylor et al. 2012), 21CMA (Zheng et al. 2016, hereafter PAPER 1) and Precision Array to Probe Epoch of Reionization (PAPER, Jacobs et al. 2011) have done a great deal of work to help us understand the population of extragalactic radio sources at frequencies below 300 MHz (e.g., Heald et al. 2015; Wayth et al. 2015; Intema et al. 2016). In this paper, we focus on 21CMA, which is designed for the measurement of the redshifted 21 cm emission signatures of neutral hydrogen against the cosmic microwave background (CMB) in the EoR. The antennas of 21CMA are fixed on the ground and point at the North Celestial Pole (NCP) region, which allows 21CMA to observe the same sky region 24 hr a day. The Westerbork Radio Telescope (WSRT, Bernardi et al. 2010) and LOFAR (Yatawatta et al. 2013) telescopes also cover the NCP region and were employed to report corresponding results. PAPER 1 published a source catalog around NCP with 21CMA using the data taken from 12 hr observation made on 2013 April 13. The analysis was restricted to the east−west (E−W) baselines within 1500 meters; 624 radio sources were found within a radius of 5° around the NCP down to about 0.1 Jy. In this work, we improve the data reduction pipeline of 21CMA and increase the

integration time to three days to generate sky maps around the NCP region. The longest baseline used is 2780 m and the angular resolution reaches about 2′ at 200 MHz; 832 radio sources within a radius of 5° around the NCP are detected. The number of sources detected are increased compared with the source catalog published in PAPER 1. The new source catalog allows us to have a better understanding of the statistical properties of radio sources and, further, the sky model for calibration and foreground subtraction can be improved.

This paper is organized as follows. We describe the observation and data reduction pipeline in Section 2. Section 3 presents the source extraction process and the source catalog. Number counts and spectral indices of the detected radio sources are also presented in Section 3. Discussion and conclusion are given in Section 4.

2. Observations and Data Reduction

2.1. Observations

The 21CMA is a ground-based radio interferometer sited in the Ulastai Valley of western China. A total of 10,287 log-periodic antennas are deployed in two arms along the E−W and north−south (N−S) directions. The 21CMA covers a frequency range of 50–200 MHz. All antennas point at the NCP region so that the same patch of sky can be observed 24 hr a day. The frequency resolution is 24.4 kHz. Data in each frequency channel are integrated for 3.56 s then output to the disk array. The baseline has components (u, v, w) in a right-handed coordinate system. u and v are measured in a plane normal to the direction of the phase reference position: v is measured toward the north as defined by the plane through the origin and u toward the east. The component w is measured in the direction which is the phase reference position. If the baseline is an E-W line on the Earth's surface, then the baseline vector has no component parallel to the rotation axis, and the $w −$ term can be neglected when the synthesized field is not too large, so choosing E−W baselines allows us to work with a relatively simple imaging algorithm without being concerned about the so-called $w −$ term (Perley 1999). The longest baseline used in this work is 2780 m, which provides resolution for resolved structures of about 2′ at 200 MHz.

The data considered in this work were collected for three days in May 2013 (May 10 2013, May 13 2013, May 22 2013). We integrated data for more than 60 hr. A full-band 50–200 MHz image by mosaicking the sub-band images together after convolution is displayed in Figure 1. Details about the data processing are presented in the following subsections.

There are two bright sources with complex structures in the NCP: 3C061.1 and B004713+891245, which can be resolved with the 21CMA. Sidelobes generated by them are hard to be removed in the following data reduction process (Yatawatta et al. 2013). The sidelobes caused by 3C061.1 reach a few Jy at 150 MHz in our observations. Beating down the noise caused by these two sources is a challenge for the 21CMA EoR observation targeted at the NCP region.

2.2. RFI Removal

Radio Frequency Interference (RFI) is one of the prominent concerns for low frequency radio data reduction. Although the radio environment of the 21CMA site is quiet, there are still strong RFI sources identified in the data, which mainly include AM radio broadcasting at about 70 MHz, FM radio broadcasting scattered by meteor tails and aircraft which covers 88–108 MHz, civil aviation aircraft at about 119 and 130 MHz, the low orbiting satellites at about 137 MHz and local train communications at about 150 MHz (Huang et al. 2016). The listed RFIs are time variable and can be identified and flagged in the first step of RFI removal. These can be easily flagged and removed in the visibilities.

Then we perform a statistical algorithm for further RFI removal. All of the visibility data points beyond 3σ deviation from Gaussian statistics are marked and flagged. Real parts of the data refer to the flux and imaginary parts refer to the position of the radio sources. Real parts and imaginary parts of the data are dealt with separately. For either the real part or the imaginary part beyond 3σ deviation, the corresponding visibility data point is removed. Specific details about the statistical treatment method for 21CMA RFI removal can be found in Section 3 of Huang et al. (2016).

2.3. Self-Calibration

Low-frequency radio interferometric observations are often affected by ionospheric turbulence, resulting in fluctuations on both source positions and flux densities. The spatial distribution of total free electron content in the ionosphere causes changes in the path delays measured by radio interferometers every few minutes. Therefore, correction of the undesired phase fluctuations caused by the ionosphere at low frequencies is an important consideration in the low frequency radio data reductions.

The spatial distribution of total free electron content in the ionosphere causes changes in the path delays measured by radio interferometers. The positional offsets of radio sources caused by ionospheric turbulence are direction-dependent and increase toward lower frequencies. The positional offsets can be corrected through self-calibration. For a given radio source, the distance between two interferometer pods (pods here mean a group of antennas. For the 21CMA case, there are 81 tiles or pods in 21CMA. In each pod, there are 127 dipoles.) is D, and the line of sight distances between the source and the interferometer pods are ℓ_1 and ℓ_2, respectively. $n(x) \sim 1 - \omega_p^2/2\omega^2$ is the refractive index. ω_p is the frequency of local electron plasma, and ω is the angular frequency of radio waves. Then the optical path difference observed by the two interferometer pods can be written as $\int_{\ell_1} n(x)dx - \int_{\ell_2} n(x)dx \sim TEC(\ell_1) - TEC(\ell_2)$. The
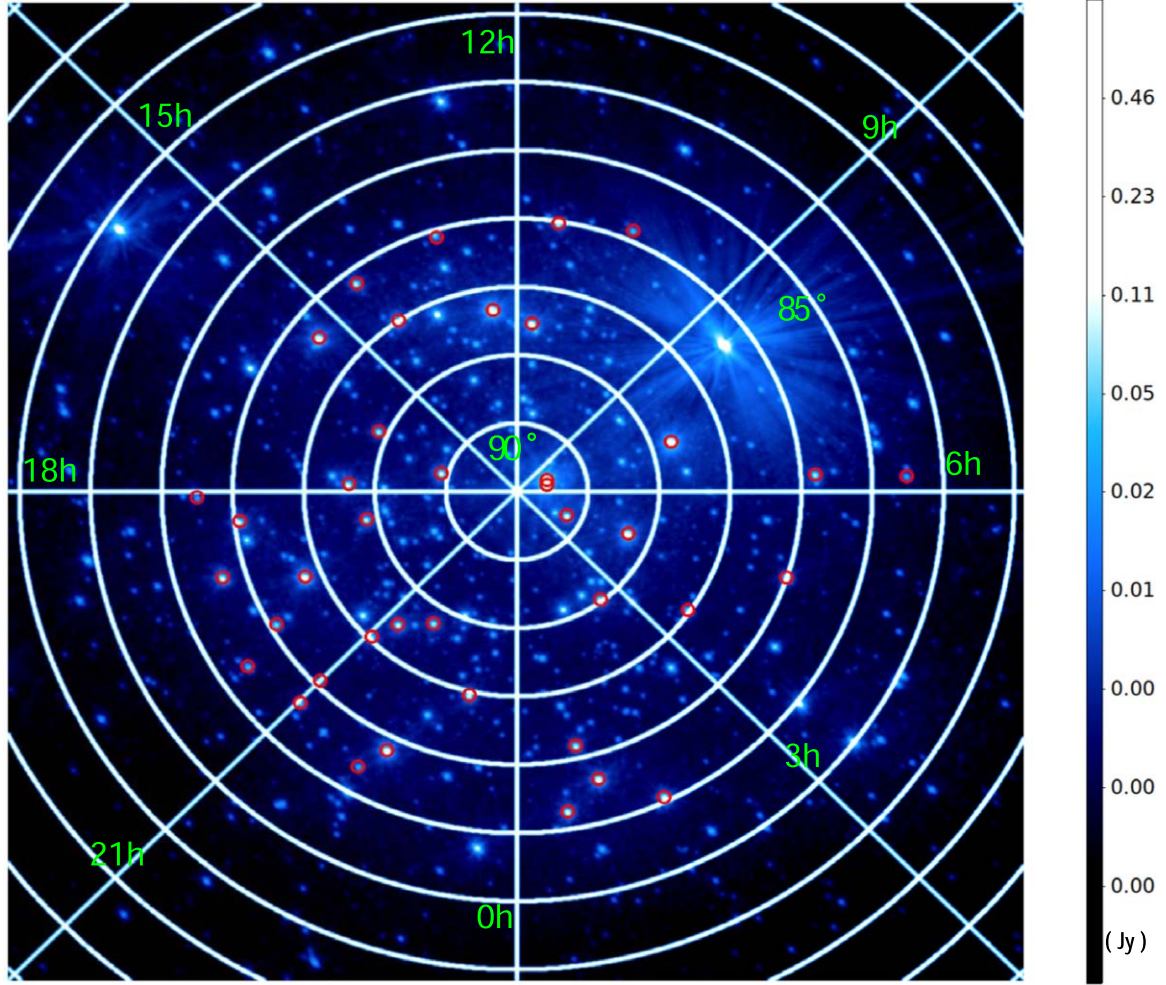
Figure 1. Sky map centered on the NCP in the frequency range of 50–200 MHz. A larger field around the NCP is shown, with an annulus of $1°$ in radial distance. Red circles represent the 40 calibrators used in calibration.

optical path difference causes the position shift of the observed radio source. In the small angle approximation, the angular shift in source position $\delta\theta$ is

$$\delta\theta \sim \frac{\int_{\ell_1} n(x)\,dx - \int_{\ell_2} n(x)\,dx}{D} \sim -\frac{c\Delta TEC}{\omega D} \quad (1)$$

where c is the speed of light. Normally, the angular shifts of observed radio sources are several arcsec at about $100 -200$ MHz, and can be measured by position offsets of radio sources in radio images.

In our data analysis, a closure phase relation is adopted to calibrate the telescope. There are 6144 frequency sub-bands between $50-200$ MHz with a frequency resolution of 24.4 kHz. In order to get better uv coverage, we combine the data to 1.56 MHz frequency bandwidth. We perform flux and phase corrections for each combined image with 1.56 MHz frequency bandwidth every 3 minutes. The corrections cannot

be applied to 24.4 kHz data, because the uv coverage does not allow us to obtain the required image quality to perform the corrections. The visibilities for each 3 minutes and 1.56 MHz snapshot are gridded to 4096 times 4096 uv maps, while uniform weighting is used to minimize sidelobe levels. The ionospheric calibration in the 21CMA data analysis is conducted by applying a model of the brightest source in the field to correct the data in the visibility domain. The source position offsets caused by ionospheric effects are a few arcmins at frequencies below 200 MHz, so data collected by long baselines, especially longer than 1 km, are more affected. In PAPER 1, baselines longer than 1500 m are avoided to reduce the ionospheric effect on the interferometric imaging to improve the image quality. In this work, all baselines along the E−W direction are considered, with the longest baseline being 2780 m. We perform corrections for baselines longer than 500 m every 3 minutes to correct the position shifts of the

radio sources. Then we combine the corrected 1.56 MHz images to obtain 25 MHz images. We used the 25 MHz images to subtract the radio sources.

2.4. Imaging

Fast Fourier Transform (FFT) is performed on the *uv* maps to generate the dirty map for each snapshot. Each of the 96 images with a frequency bandwidth of 1.5625 MHz is deconvolved utilizing the conventional Högbom CLEAN algorithm (Högbom 1974), with a loop gain of 0.05 to reduce sidelobe-induced artifacts. The CLEAN process is terminated when the fractional change in the total CLEANed flux density is less than 10^{-4} in the iteration, or a when negative value appears.

2.5. Primary Beam Calibration and Flux Calibration

After we use self-calibration to calibrate the *uv* data for each single baseline, we perform the primary beam correction on the images. For a 24 hr (or more) observation, the primary beam of a 21CMA pod has circular symmetry. In terms of the laboratory measurement of the response of the 21CMA antenna and also based on the pattern multiplication theorem, a Gaussian profile can be adopted. Based on PAPER 1, the Gaussian profile provides a good description for the 12 hr averaged primary. We start with a Gaussian profile of $F(\theta, \nu) = G(\nu)exp(-\theta^2/2\theta_b^2)$ to estimate the primary beam pattern of the 21CMA pod. Then the best-fit parameters θ_b and G can be obtained in each frequency channel. θ_b denotes the standard deviation of the Gaussian distribution and $G(\nu)$ is the gain parameter at frequency ν.

By fixing θ_b, we rely on 40 calibrators within a radius of 5° around NCP to build a model for beam correction. We collect the available flux of the 40 sources from the NASA/IPAC Extragalactic Database, fit the spectral indices of the sources based on the observation data points at different frequencies, then build a model for beam calibration. These sources are the top 40 brightest sources in the NCP region we observe with available multi-frequency observations. The fluxes at 150 MHz of the 40 sources are in the range of 1.13–5.54 Jy. The fitted spectral indices of the 40 sources are in the range of $\alpha = -0.31$ to $\alpha = -1.10$. The 40 calibrators show a random distribution within the region 5° around NCP. The distribution of the 40 calibrators is displayed in Figure 1. We used the 40 sources as the sky model to fit the primary beam and correct the flux. Then the best-fit parameters can be found in each frequency channel.

The data of the frequency channels $65-200$ MHz were used in fitting the primary beam. The antenna elements of 21CMA are designed to have the best response from $75-200$ MHz, but can be extended to 50 MHz. Avoiding the lower frequencies in the fitting can reduce the errors induced by the possible roll-off effect of the bandpass filters incorporated in the receivers. The integrated fluxes of the observed sources are used for the
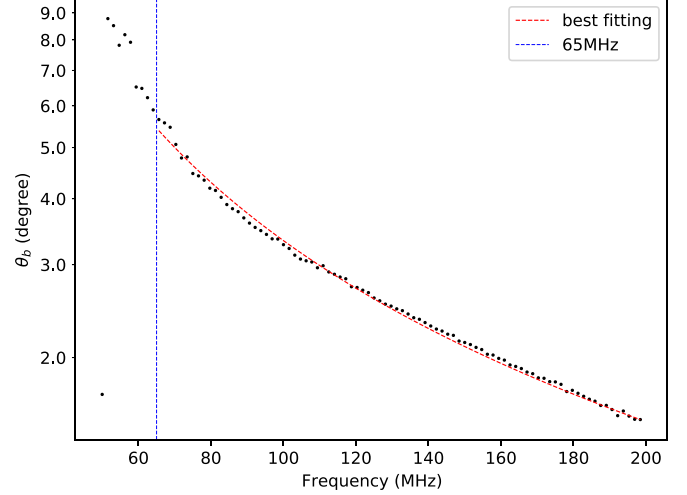

Figure 2. The primary beam of the 21CMA pod, characterized by the standard deviation of the Gaussian distribution θ_b. The dots are the fitting results from the sub-band images with a frequency width of 1.56 MHz. The data points of the frequency channels $65-200$ MHz were included in fitting the primary beam. The red dashed curve presents the best-fit, $\theta_b = 3°.33(\pm0.01)(\nu/100$ MHz$)^{-1.14(\pm0.01)}$ (dashed), and χ^2 of the fitting is 1.01. The blue dashed line shows that the data points of the frequency channels < 65 MHz were not used in fitting the primary beam.

calculation. It turns out that the Gaussian profile provides rather a good description for the 24 hr (or more) averaged primary beam (Figure 2). The best-fit standard deviation of a Gaussian beam is

$$\theta_b = 3°.33^{+0.01}_{-0.01}\left(\frac{\nu}{100\text{MHz}}\right)^{-1.14^{+0.01}_{-0.01}}. \quad (2)$$

We perform beam response correction and flux correction for 96 CLEANed sub-band image. Then we stack the 96 images to form six stacked images with the frequency bandwidth of 25 MHz. The center frequencies for the six sub-band images are 62.5, 87.5, 112.5, 137.5, 162.5 and 187.5 MHz. In order to avoid the lower frequencies in the fitting, which can reduce the errors induced by the possible roll-off effect of the bandpass filters used in the receivers, we do not use the image of the lowest frequency band (62.5 MHz) for the following source extraction, flux and position measurements. In PAPER 1, because of the limitation of the angular resolutions and errors induced by the roll-off effect and side lobes, including the lowest and highest bands did not improve the results of the following statistical studies such as calculating the differential number counts, so we chose to avoid those two bands. In this work, the angular resolution is improved, and side lobes can be better reduced through CLEAN. Although the roll-off effect cannot be avoided, the image quality allows us to do the analysis for the highest frequency band.
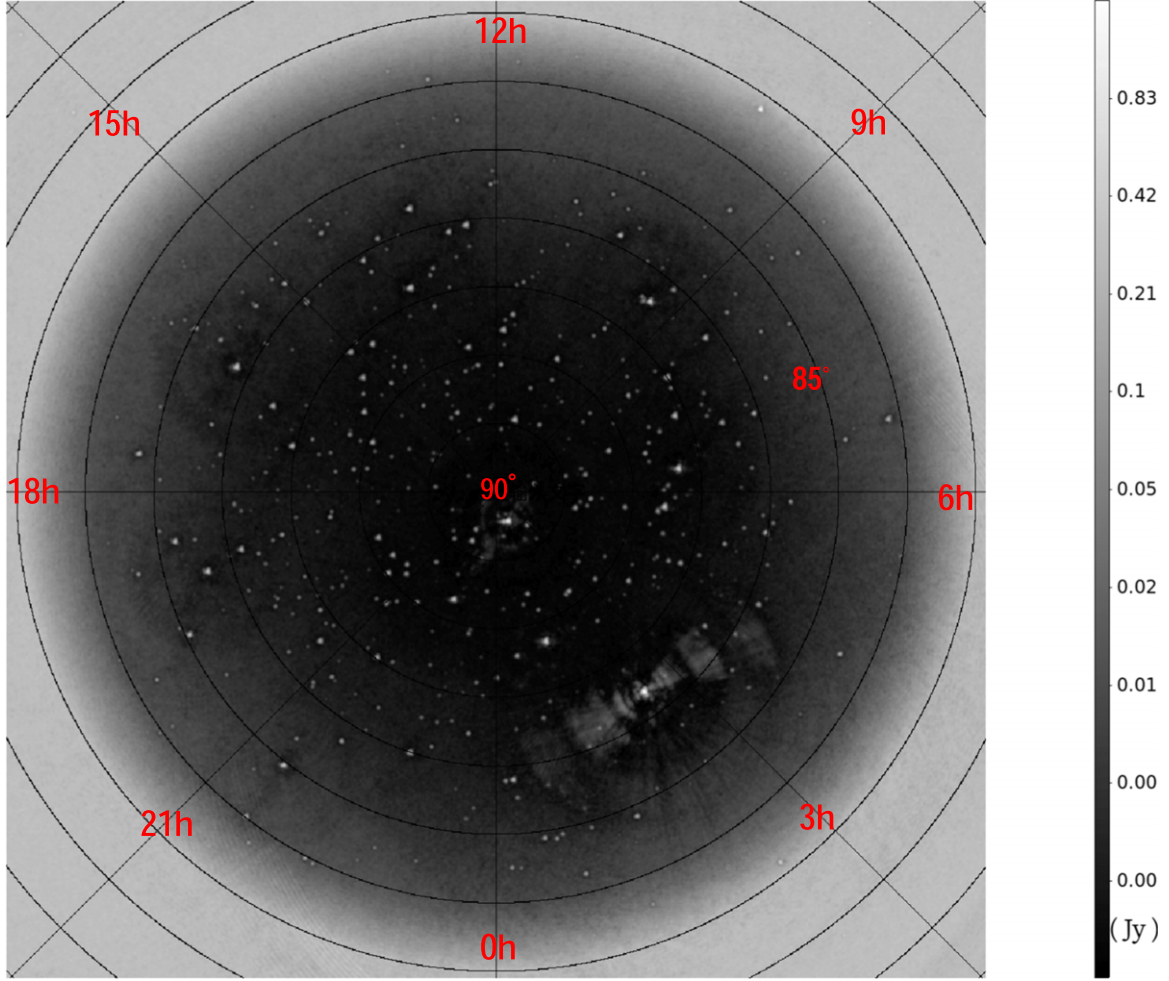
Figure 3. The final restored sub-band image at 137.5 MHZ.

3. Source Catalog and Properties

3.1. Source Identification

Radio sources are identified in each sub-band image. We present one of the sub-band images in Figure 3 as an example. The field-of-view (FoV) depends on the observing frequency. For 21CMA, the sources within a radius of $\sim 3°$ around the NCP can be extracted in the images from 75 to 200 MHz. We also count the sources between $3°$ and $5°$ to expand the source catalog as we did in PAPER 1.

We use the $75-200$ MHz full-band image to select the source candidates. A threshold of 3σ root mean square (rms) noise is adopted to identify source candidates (e.g., Whiting 2011). The positions of the peak flux of the candidates in the full-band image are tagged and then used in the sub-band images for the selection of source candidates. As we discussed in Section 2.1, the sidelobes of the brightest source (3C061.1) in the field cannot be perfectly CLEANed. The spurious sources generated by the sidelobes can be removed utilizing the sub-band images.

Flux density of each source candidate is fitted by a Gaussian profile in each sub-band image. A double Gaussian model is adopted when two sources are close to each other and partially overlap, but the models may fail due to poor resolutions. In this case, the flux density information for the sources is not provided in the source catalog when the model fails. For very faint sources which cover less then ten pixels in the images, the flux density information is not provided when a good fit of the model cannot be obtained.

The source candidates are removed from the list if their counterparts are missing in one of the five frequency bands. A total of 832 radio sources are found around the NCP region, and their positions, integrated fluxes and spectral indices are listed in Table 1. We measure the sidelobes of the two brightest sources in the field and estimated errors; 20% error is adopted in the flux measurements. This error is given from the error

Table 1
Source Catalog

	R.A. (J2000) (h:m:s)	Decl. (J2000) (d:m:s)	$F_{87.5}$ (Jy)	$F_{112.5}$ (Jy)	$F_{137.5}$ (Jy)	$F_{162.5}$ (Jy)	$F_{187.5}$ (Jy)	α
1	11:5:31.6	85.0:1.0:20.6	$0.5817^{+0.12}_{-0.12}$	$0.3668^{+0.07}_{-0.07}$	$0.319^{+0.06}_{-0.06}$	$0.339^{+0.07}_{-0.07}$	$0.3507^{+0.07}_{-0.07}$	$-0.63\ ^{+0.44}_{-0.44}$
2	12:2:21.3	85.0:1.0:27.9	$0.7931^{+0.16}_{-0.16}$	$0.6021^{+0.12}_{-0.12}$	$0.5603^{+0.11}_{-0.11}$	$0.4204^{+0.08}_{-0.08}$	$0.395^{+0.08}_{-0.08}$	$-0.93\ ^{+0.81}_{-0.81}$
3	13:38:44.1	85.0:3.0:3.6	$0.5292^{+0.11}_{-0.11}$	$0.3795^{+0.08}_{-0.08}$	$0.3614^{+0.07}_{-0.07}$	$0.2377^{+0.05}_{-0.05}$	$0.288^{++0.06}_{-+0.06}$	$-0.91\ ^{+0.70}_{-0.70}$
4	23:40:36.3	85.0:3.0:33.1	$0.5614^{+0.11}_{-0.11}$	$0.4049^{+0.08}_{-0.08}$	$0.3221^{+0.06}_{-0.06}$	$0.2756^{+0.06}_{-0.06}$	$0.2978^{+0.06}_{-0.06}$	$-0.91\ ^{+0.71}_{-0.71}$
5	17:30:43.5	85.0:3.0:35.7	$0.4417^{+0.09}_{-0.09}$	$0.2841^{+0.06}_{-0.06}$	$0.2233^{+0.04}_{-0.04}$	$0.2564^{+0.05}_{-0.05}$	$0.224^{+0.04}_{-0.04}$	$-0.81\ ^{+0.55}_{-0.55}$
6	16:11:23.8	85.0:7.0:3.5	$0.7417^{+0.15}_{-0.15}$	$0.4041^{+0.08}_{-0.08}$	$0.3419^{+0.07}_{-0.07}$	$0.329^{+0.07}_{-0.07}$	$0.3412^{+0.07}_{-0.07}$	$-0.99\ ^{+0.82}_{-0.82}$
7	9:24:9.0	85.0:7.0:29.3	$0.5026^{+0.10}_{-0.10}$	$0.3507^{+0.07}_{-0.07}$	$0.2593^{+0.05}_{-0.05}$	$0.2423^{+0.05}_{-0.05}$	$0.2543^{+0.05}_{-0.05}$	$-0.96\ ^{+0.73}_{-0.73}$
8	1:2:21.4	85.0:7.0:42.2	$1.1422^{+0.23}_{-0.23}$	$1.2448^{+0.25}_{-0.25}$	$0.7254^{+0.15}_{-0.15}$	$1.0829^{+0.22}_{-0.22}$	$0.6646^{+0.13}_{-0.13}$	$-0.63\ ^{+0.62}_{-0.62}$
9	6:59:57.1	85.0:8.0:11.6	$3.4223^{+0.68}_{-0.68}$	$1.9566^{+0.39}_{-0.39}$	$1.6942^{+0.34}_{-0.34}$	$1.5458^{+0.31}_{-0.31}$	$1.8536^{+0.37}_{-0.37}$	$-0.84\ ^{+0.98}_{-0.98}$
10	13:13:24.3	85.0:8.0:14.3	$2.3216^{+0.46}_{-0.46}$	$1.2942^{+0.26}_{-0.26}$	$1.1836^{+0.24}_{-0.24}$	$1.0343^{+0.21}_{-0.21}$	$1.1821^{+0.24}_{-0.24}$	$-0.89\ ^{+0.95}_{-0.95}$

Note. The coordinates, total fluxes at five frequency bands (if available) and spectral indices are provided for 832 sources detected within a radius of 5° around the NCP. The positional accuracy of the central coordinates is ~2'. This table 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.

estimated from the calibrators. The average error of the fitted spectral indices for the 40 sources is 20.3 percent. Sources found outside the radius of 5° around NCP are not included in the source catalog.

We perform a cross-check with the 6C catalog at 151 MHz here. (The flux densities are on the scale of Roger et al. 1973.) We extrapolate the flux of sources in our catalog to 151 MHz to make the comparison. Because the FoV of 21CMA above 150 MHz is about 3° in radius, Figure 4 presents the comparison of the flux densities of the sources within the radius of 3° around NCP in the 6C catalog and this work, and two measurements above 0.1 Jy show a good agreement.

Obviously, more fainter sources (<1 Jy) are detected in this work. More fainter sources can be resolved due to the improvement in angular resolution and the reduction of noise via long time integration. The noise of the images is reduced from 10.3 to 4.4 mJy at 150 MHz. We detected 832 radio sources in the region within a radius of 5° around NCP. There are 208 sources more than we detected in PAPER 1. We extrapolate the flux measured in sub-band images to 150 MHz for the source catalog given in both PAPER 1 and this work. The are 110 more sources below 0.5 Jy detected in this work than in PAPER 1. In the region within a radius of 3° around the field center, there are 86 more sources detected below 0.5 Jy in this work.

The reasons for the increase in number of detected radio sources can be summarized as follows: a. longer baselines are included and the angular resolution is improved; b. three days of data are used in this work, the integration time is increased from 12 hr (PAPER 1) to 3 days (this work) and the noise of the image is reduced, which allows faint radio sources be resolved; c. position offsets caused by the ionosphere are corrected, and CLEAN can perform better to reduce the sidelobes.
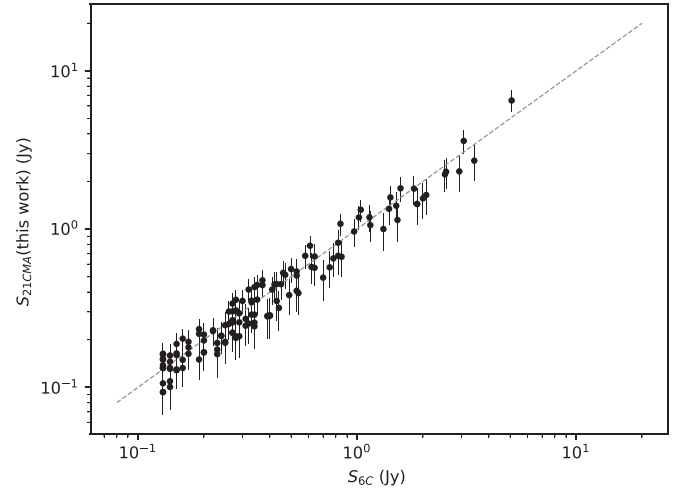

Figure 4. Comparison of flux densities measured in the 6C catalog and this work above 0.1 Jy. The dashed line indicates equal flux density values.

3.2. Spectral Index

Figure 5 illustrates the histogram of the spectral indices of the 832 sources presented in this paper. The spectral indices of most of the sources are in the range of $\alpha \approx -0.5$ to $\alpha \approx -1.0$. For the spectral indices of all 832 sources, the mean value is approximately equal to $\alpha = -0.73$ and median value is $\alpha = -0.75$. Radio sources with steeper spectra are more easily detected at low frequencies, so some sources in the 21CMA source catalog have steeper spectral indices of $\alpha \approx -1.0$. We only perform the spectral fitting within a very narrow frequency band, hence uncertainties are inevitable.

We have also compared the spectral indices in the inner radius of 3° and the ones in the outer annulus of 3°–5° (Figure 6) as we did in PAPER 1. The FoV of 21CMA at high
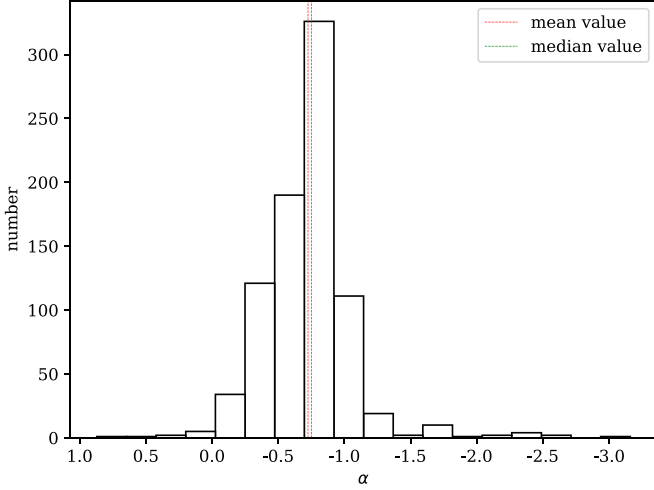
Figure 5. Histogram of spectral indices α ($S \sim \nu^{\alpha}$) for the 21CMA source catalog. The mean value (red dotted vertical line) is approximately equal to $\alpha = -0.73$ and median value (green dotted vertical line) is $\alpha = -0.75$.
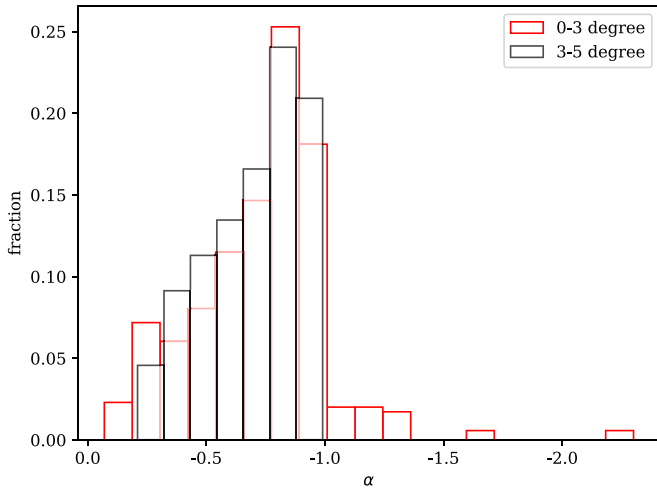

Figure 6. Comparison of spectral indices of sources in the central region of 3° (red) around the NCP and in the outer annulus of 3°–5° (black). The vertical axis signifies the source fraction.

frequencies (above 150 MHz) is limited to the radius of 3° around NCP. The histogram of spectral indices in the inner radius of 3° and the outer annulus displays a similar distribution, and the maximum value is about -0.6 to -0.9. It reflects that our primary beam correction is sufficiently unbiased.

3.3. Completeness

The deep low frequency observation made with GMRT (Intema et al. 2011) at 153 MHz provides the source number count down to the flux of $\sim$6 mJy, which allows us to do a quick check of the completeness of our source catalog.
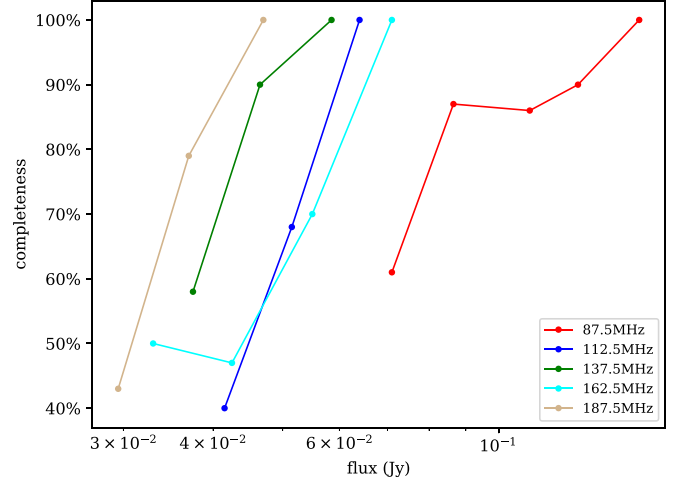

Figure 7. Completeness of sources within a radius of 3° around NCP at different frequency bands.

Completeness can be estimated through a Monte-Carlo simulation. The differential source count given in Intema et al. (2011), $dN/dS = 5370S^{-1.59}$, is used to generate radio sources randomly distributed in a sky field of radius 5°. Then the smearing effects of the beam at different frequencies are included and noises of bright sources are considered. (Side-lobes/grating lobes of bright sources generate noises in the images which are hard to deal with, which are caused by the redundant configuration of 21CMA.) We perform the source extraction procedure. (The same as what was used in the catalog. Actually our simulation still has limitations, since the artifacts such as the grading lobes and noise induced by CLEAN are impossible to be perfectly simulated.) The completeness is assumed to be constant in the field. We generated 20 simulated fields with sources randomly distributed. The flux of the sources is restricted from 5 mJy to 30 Jy, and the spectral index is assumed to be -0.7. The effect of beam smearing is included by convolving the image with a Gaussian beam. In Figure 7, the completeness of sources within a 5° radius around NCP in each flux bin of the five frequency bands is presented. The error here can be estimated by how many times the simulation is run for the estimation. The result suggests that in the lowest frequency band (87.5 MHz) the incompleteness for our standard method of source selection occurs around 0.1 Jy, where the ratio drops to $\sim$80%. We reach 50% completeness at $\sim$0.05 Jy at the lower two frequency bands, and we reach 50% completeness at $\sim$0.03 Jy at the higher three frequency bands.

3.4. Source Counts

The Euclidean normalized differential number counts dN/dS would follow the two well-known power laws for cosmic radio population at low frequencies (Hales et al. 1988; Cohen et al.
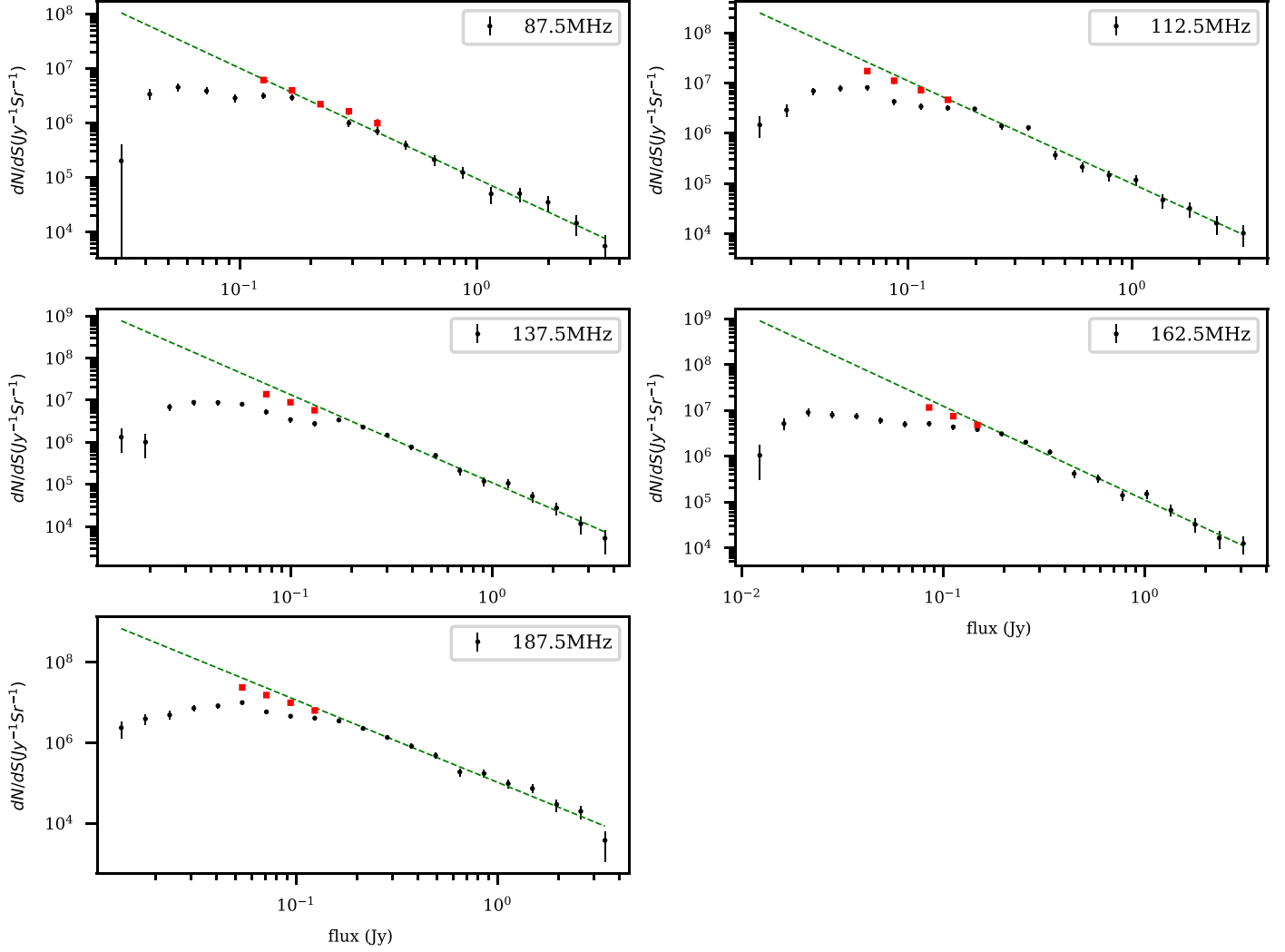
Figure 8. Differential source counts over the central field within a radius of 5° for each of the five frequency sub-bands. The best-fit power law is signified by green dotted lines. Black and red points represent raw data and completeness-corrected data. The data points for completeness above 50% are adopted.

2007; Moore et al. 2013).

$$\frac{dN}{dS} = \begin{cases} k_1 S^{-\gamma_1} Jy^{-1}sr^{-1}, & S < S_0 \\ k_2 S^{-\gamma_2} Jy^{-1}sr^{-1}, & S \geqslant S_0 \end{cases}. \qquad (3)$$

Here S_0 indicates the turnover flux. For sources with fluxes greater than S_0, a single power law fit can be adopted

$$\frac{dN}{dS} = k S^{-\gamma} Jy^{-1}sr^{-1}. \qquad (4)$$

We calculate the Euclidean normalized differential number counts dN/dS of our sample over the central field of radius 5° for each of the five frequency bands. For simplicity, we adopt the flux corresponding to 50% completeness as S_0 and fit the single power law form for the sources with the fluxes brighter than S_0. Completeness-corrected data are used, and a single

Table 2
The Best-fitted Parameters for Differential Source Counts (50% Completeness)

ν_c (MHz)	k	γ
87.5	6675.74 ± 1.19	1.79 ± 0.08
112.5	6428.85 ± 1.18	1.75 ± 0.07
137.5	6147.13 ± 1.20	1.74 ± 0.08
162.5	5706.97 ± 1.23	1.76 ± 0.09
187.5	4829.59 ± 1.23	1.80 ± 0.09

power law fits the data nicely. The fitting result is displayed in Figure 8. Our fitting parameters k and γ for the five frequency bands are listed in Table 2. We present the differential number counts for sources in the region within 5° in radius around NCP. But since the simulation used for corrections has limitations, the artifacts caused by grating lobes and side lobes
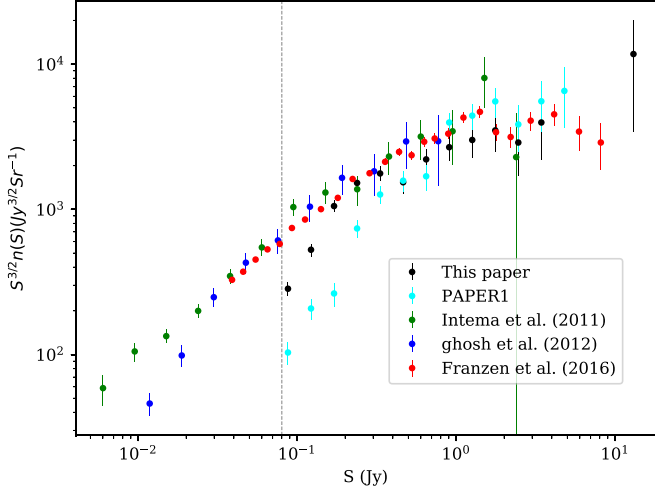
Figure 9. Euclidean-normalized differential counts based on the flux densities of radio sources within radius 5° around NCP measured in the 168.5 MHz frequency band in this work (black points). The PAPER 1 results are presented as cyan points. Green and blue points are the GMRT observations at 153 MHz and 150 MHz, respectively (Intema et al. 2011; Ghosh et al. 2012). The MWA result at 154 MHz (Franzen et al. 2016) is signified by red points. The vertical dotted line at ~ 0.08 Jy corresponds to the completeness of 50%.

cannot be perfectly involved in the simulation, so it is still possible for the correction to be biased for real data, especially for faint sources. Improving the data analysis technique is one of the most important works. But due to the current 21CMA configuration, artifacts like grating lobes of bright sources are hard to deal with. It is also a good lesson for us, and although redundant configuration can improve the instrument calibration, grating lobes should be carefully dealt with. That is why the future SKA1-low will use a random configuration.

It is shown that a power-law fit at $S < 0.4$ Jy to the GMRT counts from Williams & Intema et al. (2013), Intema et al. (2011) is acceptable. Another research work gives $\gamma_1 = 1.54$ (Ghosh et al. 2012) at 154 MHz (Franzen et al. 2016). The GMRT targeted survey at 153 MHz finds essentially similar results of $\gamma_1 = 1.59$ (Intema et al. 2011), see also (Ghosh et al. 2012). We take the source counts from two 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). The comparison of our results with other deep surveys at low frequencies is depicted in Figure 9. In this plot, we used all sources observed within the field with radius 5° in the 21CMA observation and extrapolate the data points to 150 MHz to compare with the GMRT and MWA observations. The data points above 50% completeness are presented in the figure. We made comparisons in Figure 9, together with other observations. Instead of relying on the sources in the region of 3° around the NCP, we use the sources in the region of 5° in radius around the NCP for both PAPER 1 and this work. Instead of utilizing the corrected data points, we plot the

original data points for PAPER 1 and this work. It is clear to see that the catalog of this work is improved over that in PAPER 1, but our result is still limited by resolution of the 21CMA and the grating lobes generated by bright sources.

3.5. Noise Analysis

The key science goal for the 21CMA is detecting the redshifted 21 cm signals from the EoR. Noise analysis is important in both the statistical measurements and imaging of the EoR, which primarily comprises system noise, confusion noise, calibration error and deconvolution error.

For a radio interferometer like 21CMA, the thermal noise can be estimated simply from

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

where A_{eff} is the effective area and T_{sys} is the system temperature. k_B is Boltzmann's constant, and $\Delta\nu$, Δt and η are the frequency bandwidth, observation duration and efficiency factor, respectively. We used 40 stations along the E-W arm in this work, so we have 780 baselines. The effective area of a single station is about $a_{\text{eff}} = 216$ m^2, so the total effective area can be written as $A_{\text{eff}} = \eta\sqrt{780}\,a_{\text{eff}}$, where η is the efficiency which is 50% for the 21CMA. T_{sys} is made up by the telescope temperature (50 K for the 21CMA) and the sky temperature can be described approximately by $60(\nu/300\,\text{MHz})^{-2.55}$ (e.g., Chapman et al. 2012). When the frequency bandwidth is 30 MHz and integration time is 3 days, the estimated thermal noise for 21CMA is about 0.1 mJy at 150 MHz.

When surface density of faint radio sources is high, the sources cannot be resolved and detected individually below the fluctuation of the background sky brightness due to the confusion. Classical confusion noise depends on the source counts and the synthesized beam of the radio telescope. Confusion noise σ_c from the positional uncertainties of unresolved, faint radio sources below a flux threshold S_{lim} with finite synthesis beam Ω_b can be estimated by (e.g., Scheuer 1957; Condon 1974)

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

The differential number counts of GMRT observations at 153 MHz is used to calculate the confusion noise. The flux threshold S_{lim} is 0.1 mJy. Then the confusion noise is about 2 mJy at 150 MHz.

There are two bright sources in the NCP region, B004713 +891245 and 3C061.1. Due to the redundant baselines of 21CMA, grating lobes generated by the bright sources are hard to remove, and the spaced ring-like structures generate errors in the data reduction. The resolution of 21CMA is limited to a few arcmins. A more accurate sky model of the bright sources, especially the sources with complex structures, is required to
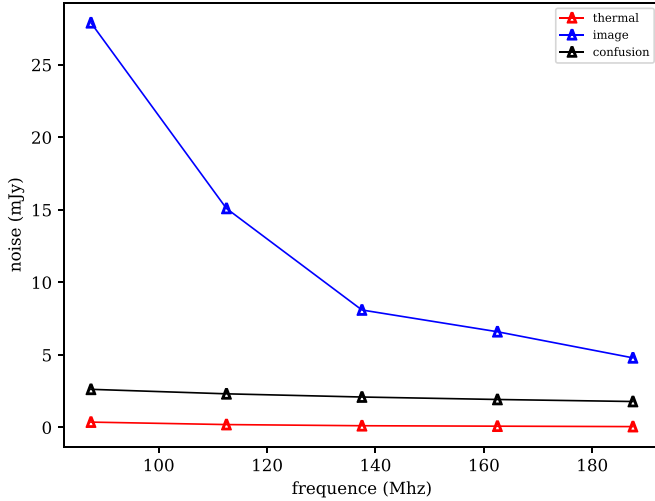
Figure 10. Thermal noise, image noise and confusion noise of sub-band images.

improve the data reduction procedures such as calibration and deconvolution.

We measure the noise of the sub-band images, which is about 4.4 mJy at 150 MHz. The image noise, confusion noise and thermal noise of each sub-band image are plotted in Figure 10. In conclusion, thermal noise can be controlled and reduced by increasing the observation time. Reducing the confusion noise is limited by the unresolved radio sources and very faint sources. Errors generated from calibration and deconvolution propagate to the final statistical measurements and imaging. Reducing the errors of calibration and deconvolution can improve the image quality, and help in building an accurate sky model. Grating lobes can be reduced by using random array configurations.

4. Discussion and Conclusion

We analyze more than 60 hr of data taken from the 21CMA observations centered on the NCP. The longest baseline used is 2780 m, which allows the angular resolution to reach about 2′ at 200 MHz, and the number of sources detected are increased compared to the source catalog given in PAPER 1. We calibrate the primary beam and gain using a model with 40 bright sources in the field. We detect 832 radio sources within a radius of 5° around the NCP region. A source catalog from 75 to 200 MHz is given in this work. We estimate a completeness of 50% at the flux of about 0.08 Jy and compare the 21CMA source count with the observations made with GMRT and MWA. The result shows a good agreement.

In PAPER 1, there are 624 radio sources detected in the region of 5° radius around the NCP. In this work, the number of detected radio sources in the same region is increased to 832. There are 631 more sources below 1 Jy at 150 MHz detected in this work. In PAPER 1, in order to reduce the ionospheric effects and

improve the image quality, baselines longer than 1500 m and shorter than 100 m are excluded in the analysis. We included one step to correct the position offsets caused by the ionosphere in this work, and all baselines in the E-W arm are used in the analysis in this work, so that the image resolution is improved. The 21CMA layouts adopt a duplicated configuration, so grating lobes appear in the images especially for short baselines. In this work, although grating lobes of bright sources cannot be completely avoided, more fainter sources can be detected than those in PAPER 1 since the integration time is increased to three days, and the resolution is improved from 5′.6 to 3′.1 at 150 MHz. In conclusion, the reasons for the increase in the number of detected radio sources are listed as follows: a. longer baselines are included and the angular resolution is improved. b. three days of data are used in this work, so fainter sources can be detected because the noise of the images is reduced. c. position offsets caused by the ionosphere are corrected, so CLEAN can perform better to reduce the sidelobes. In this case, we learn that avoiding longer baselines and shorter baselines can improve the image quality in some situations, but for future studies, such as subtracting the large scale structures of 21 cm signals, building a completed sky source model, etc., including all long baselines and short baselines is necessary, but the data reduction techniques are required to be improved.

Although the source catalog presented in this work is improved compared to PAPER 1, there are still some problems in our current data reduction pipelines, including: evident grating lobes of bright sources, errors generated in calibration, deconvolution procedures and so on. As we already discussed in PAPER 1, the spacing between the 21CMA pods is integral multiples of 20 m, therefore the layout generates grating lobes in the images that are quite hard to be removed perfectly (Bracewell & Thompson 1973; Amy & Large 1990). The deconvolution process still needs to be improved, and a better model of the center bright source 3C061.1 is required to improve the data reduction.

Our future work is improving the data reduction technique to reduce the sidelobe effects of bright sources. Then foreground subtraction and statistical measurement of the background signals will be done.

References

Amy, S. W., & Large, M. I. 1990, PASA, 8, 308

Bernardi, G., de Bruyn, A. G., Harker, G., et al. 2010, A&A, 522, A67

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. 2016, MNRAS, 459, 3314

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., Orrú, E., et al. 2015, A&A, 582, A123

Högbom, J. A. 1974, A&AS, 15, 417

Huang, Y., Wu, X.-P., Zheng, Q., et al. 2016, RAA, 16, 16

Hurley-Walker, N., Morgan, J., Wayth, R. B., et al. 2014, PASA, 31, 45

Intema, H. T., Jagannathan, P., Mooley, K. P., & Frail, D. A. 2016, A&A, 598, A78

Intema, H. T., van Weeren, R. J., Röttgering, H. J. A., & Lal, D. V. 2011, A&A, 535, A38

Jacobs, D. C., Aguirre, J. E., Parsons, A. R., et al. 2011, ApJL, 734, L34

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

Paciga, G., Albert, J. G., Bandura, K., et al. 2013, MNRAS, 433, 639

Perley, R. A. 1999, in Imaging in Radio Astronomy II, A Collection of Lectures from the Sixth NRAO/NMIMT Synthesis Imaging Summer School, Astron. Soc. Pac., San Francisco, 180 ed. G. B. Taylor, C. L. Carilli, & R. A. Perley (San Francisco, CA: ASP) , 383

Rees, N. 1990, MNRAS, 244, 233

Roger, R. S., Costain, C. H., & Bridle, A. H. 1973, AJ, 78, 1030

Scheuer, P. A. G. 1957, PCPS, 53, 764

Taylor, G. B., Ellingson, S. W., Kassim, N. E., et al. 2012, JAI, 1, 50004

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

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., et al. 2013, A&A, 549, A55

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

Zheng, Q., Wu, X.-P., Melanie, J.-H., et al. 2016, ApJ, 832, 190