Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **536,** 2187–2211 (2025) https://doi.org/10.1093/mnras/stae2528
Advance Access publication 2024 November 9

MIGHTEE: the continuum survey Data Release 1

C. L. Hale [1,2★] I. Heywood [1,3,4★] M. J. Jarvis [1,5★] I. H. Whittam [1] P. N. Best,[2] Fangxia An [6,7] R. A. A. Bowler [8] I. Harrison,[9] A. Matthews,[10] D. J. B. Smith,[11] A. R. Taylor [12,13,5] and M. Vaccari [12,7,14]

[1]*Astrophysics, Department of Physics, University of Oxford, Keble Road, Oxford, OX1 3RH, UK*
[2]*Institute for Astronomy, Royal Observatory Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, UK*
[3]*Centre for Radio Astronomy Techniques and Technologies, Department of Physics and Electronics, Rhodes University, PO Box 94, Makhanda, 6140, South Africa*
[4]*South African Radio Astronomy Observatory, 2 Fir Street, Black River Park, Observatory, Cape Town, 7925, South Africa*
[5]*Department of Physics and Astronomy, University of the Western Cape, Robert Sobukwe Road, 7535 Bellville, Cape Town, South Africa*
[6]*Purple Mountain Observatory, Chinese Academy of Sciences, 10 Yuanhua Road, Qixia District, Nanjing 210023, People's Republic of China*
[7]*Inter-University Institute for Data Intensive Astronomy, and Department of Physics and Astronomy, University of the Western Cape, Robert Sobukwe Road, 7535 Bellville, Cape Town, South Africa*
[8]*Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, School of Natural Sciences, The University of Manchester, Manchester, M13 9PL, UK*
[9]*School of Physics and Astronomy, Cardiff University, The Parade, Cardiff, Wales CF24 3AA, UK*
[10]*Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA*
[11]*Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK*
[12]*Inter-University Institute for Data Intensive Astronomy, Department of Astronomy, University of Cape Town, 7701 Rondebosch, Cape Town, South Africa*
[13]*Department of Astronomy, University of Cape Town, Rondebosch, Cape Town 7701, South Africa*
[14]*INAF – Istituto di Radioastronomia, via Gobetti 101, 40129 Bologna, Italy*

Accepted 2024 October 7. Received 2024 September 27; in original form 2024 May 24

ABSTRACT

The MeerKAT International GHz Tiered Extragalactic Exploration Survey (MIGHTEE) is one of the large survey projects using the MeerKAT telescope, covering four fields that have a wealth of ancillary data available. We present Data Release 1 of the MIGHTEE continuum survey, releasing total intensity images and catalogues over ∼20 deg^2, across three fields at ∼1.2–1.3 GHz. This includes 4.2 deg^2 over the Cosmic Evolution Survey (COSMOS) field, 14.4 deg^2 over the XMM Large-Scale Structure (XMM-LSS) field and deeper imaging over 1.5 deg^2 of the Extended Chandra Deep Field South (CDFS). We release images at both a lower resolution (7–9 arcsec) and higher resolution (∼ 5 arcsec). These images have central rms sensitivities of ∼ 1.3−2.7 μJy beam^{-1} (∼ 1.2−3.6 μJy beam^{-1}) in the lower (higher) resolution images, respectively. We also release catalogues comprised of ∼144 000 (∼114 000) sources using the lower (higher) resolution images. We compare the astrometry and flux-density calibration with the Early Science data in the COSMOS and XMM-LSS fields and previous radio observations in the CDFS field, finding broad agreement. Furthermore, we extend the source counts at the ∼10 μJy level to these larger areas (∼ 20 deg^2) and, using the areal coverage of MIGHTEE we measure the sample variance for differing areas of sky. We find a typical sample variance of 10 –20 per cent for 0.3 and 0.5 deg^2 subregions at $S_{1.4} \leq 200\mu$Jy, which increases at brighter flux densities, given the lower source density and expected higher galaxy bias for these sources.

Key words: catalogues – surveys – radio continuum: galaxies – radio continuum: general.

1 INTRODUCTION

Radio surveys provide a dust-free view of the active Universe. At relatively high radio flux density (e.g. at 1.4 GHz, $S_{1.4} \gtrsim 1$ mJy), radio surveys can cover huge swathes of sky quickly (e.g. Becker, White & Helfand 1995; Condon et al. 1998; Mauch et al. 2003; Lacy et al. 2020; McConnell et al. 2020), providing a census of the rare, bright radio source population over the vast majority of cosmic time. Surveys at these relatively high flux densities are dominated by sources where the radio emission arises from synchrotron radiation from jets and lobes as a result of accretion on to the central supermassive black holes in galaxies (known as active galactic nuclei, or AGN) and observations of these sources provides a plethora of information on the physical conditions of these high-energy phenomena. The fact that we are able to observe these AGN across the vast majority of cosmic time (e.g. De Breuck et al. 2000; Jarvis et al. 2009; Saxena et al. 2018), means we are also able to trace the evolution of radio-loud AGN (e.g. Jarvis et al. 2001; Willott et al. 2001; Clewley & Jarvis 2004; Rigby et al. 2011). This allows insights into their relation to their host galaxy properties (e.g.

★ E-mail: Catherine.Hale@physics.ox.ac.uk (CLH); Ian.Heywood@physics.ox.ac.uk (IH); Matt.Jarvis@physics.ox.ac.uk (MJJ)

McLure et al. 1999; Dunlop et al. 2003; Gürkan, Hardcastle & Jarvis 2014; Saxena et al. 2019) and their larger scale environment (e.g. Rawlings & Jarvis 2004; McNamara & Nulsen 2007; Fabian 2012; Magliocchetti 2022) to be studied. Such surveys have revealed how radio AGN populations are related to their accretion rate (e.g. Best & Heckman 2012; Heckman & Best 2014; Mingo et al. 2014; Whittam et al. 2018) through both positive and negative feedback processes (e.g. Bower et al. 2006; Croton et al. 2006; Hardcastle, Evans & Croston 2007; Fernandes et al. 2015; Kalfountzou et al. 2017).

In addition to the synchrotron radiation that is emitted due to the accretion activity on to supermassive black holes, radio surveys are also able to detect the integrated synchrotron emission from the cosmic rays accelerated due to the young massive stars becoming supernovae (e.g. Condon 1992). Given that the radio emission has a long wavelength in relation to the size of dust grains, the radio waves are able to penetrate any dusty regions in the host galaxies, or along the line of sight, and thus potentially provide a dust-free measurement of the star formation rate in galaxies. As such, radio surveys provide an important tracer of both star formation and AGN activity across cosmic time. The wide and shallow surveys, highlighted above, are only able to detect these star-forming galaxies (SFGs) at relatively low redshift (e.g. Yun, Reddy & Condon 2001; Bell 2003; Mauch & Sadler 2007; Jarvis et al. 2010). In order to probe SFGs (and low-luminosity AGN) across a wide redshift range, radio surveys much deeper than the mJy regime (at 1.4 GHz) are required. This is demonstrated by recent cross-matched and classified radio data from both the LOw Frequency Array (LOFAR) Deep Fields (Kondapally et al. 2022; Best et al. 2023) and the MeerKAT International GHz Tiered Extragalactic Exploration Survey (MIGHTEE) Early Science (ES; Whittam et al. 2022, 2024), which show that SFGs have a typical median redshift of $z \sim 0.1$ for $S_{1.4\mathrm{GHz}} \geq 1$ mJy, increasing to $z \sim 0.8-0.9$ across the full, deeper, samples.

Indeed, we have seen significant progress in our understanding of how radio emission may trace star formation and low-luminosity AGN in distant galaxies. This is from both wider surveys which probe the sub-mJy regime, e.g. the Faint Images of the Radio Sky Survey (FIRST; White et al. 1997) and the LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al. 2022) and deep surveys over smaller areas (e.g. Bondi et al. 2003; Smolčić et al. 2017a; Heywood et al. 2020; Sabater et al. 2021). SFGs become a more dominant population for deeper surveys, and their star formation rate is often investigated through the study of the far-infrared–radio correlation (FIRC e.g. Garrett 2002; Appleton et al. 2004; Ivison et al. 2010; Delhaize et al. 2017). Several recent studies have investigated the mass dependence of the FIRC or the relationship of radio luminosity-to-star formation rate. These studies demonstrate that the observed evolution in the such relationships with redshift could be largely explained by introducing a stellar-mass dependence (e.g. Gürkan et al. 2018; Delvecchio et al. 2021; Smith et al. 2021), with other work considering its luminosity dependence (see Matthews et al. 2021b). Due to the fact that the ancillary data that are required to measure the stellar masses of galaxies detected at both far-infrared and radio wavelengths is flux-limited, those galaxies that are observed at higher redshifts necessarily also have higher average stellar masses than the galaxies observed in the low-redshift Universe. Alternatively, the dependence could also be related to the morphology of the galaxy (Molnár et al. 2018), which is in turn also correlated with the mass, with ellipticals generally being more massive than spirals in flux-limited samples. Thus, the combination of deep radio data and deep multiwavelength data is essential for such studies, allowing the possibility of redshifts, host galaxies, and galaxy properties to be identified for the radio sources. This is the main reason why the new generation of deep radio surveys, such as the focus of this paper namely the MIGHTEE (Jarvis et al. 2018) survey, are designed to overlap with those extragalactic deep fields with the best multiwavelength data.

These recent deeper radio surveys have also been used to measure the radio emission due to AGN activity to much lower luminosities, demonstrating that the space–density evolution is luminosity dependent (e.g. Rigby et al. 2015; Yuan et al. 2017; Ceraj et al. 2018; Šlaus et al. 2020; Kondapally et al. 2022) and that the correspondence between radio-luminosity and accretion rate (e.g. Mingo et al. 2014; Whittam et al. 2018), host galaxy properties (e.g. Radcliffe et al. 2021; Delvecchio et al. 2022; Ji et al. 2022; Best et al. 2023), and the underlying dark matter haloes (e.g. Lindsay, Jarvis & McAlpine 2014; Hale et al. 2018; Alonso et al. 2021) is complicated. Furthermore, these surveys are deep enough to begin to probe the radio emission from large samples of those AGN classified as radio quiet (e.g. White et al. 2015, 2017; Panessa et al. 2019; Ceraj et al. 2020; Macfarlane et al. 2021).

Thus, we are entering a realm of detailed understanding of evolution of star formation and AGN activity in galaxies from their radio emission. However, the interdependencies between star formation and AGN activity also highlights the need to combine deep radio observations with deep multiwavelength data covering a wide enough area to fully sample the radio luminosity function for SFGs and all types of AGN, above and below the knee in respective luminosity functions. These interdependencies also need to be studied in the context of whilst enabling the role of the environment to be investigated, whilst mitigating the effects of cosmic variance.

The MIGHTEE Survey (Jarvis et al. 2018) is a radio survey using the MeerKAT telescope (Jonas 2009; Jonas & MeerKAT Team 2018) aiming to provide this combination of depth and area, over some of the most well-studied extragalactic deep fields accessible from the Southern hemisphere, in order to accelerate our understanding of galaxy and AGN evolution. The ES Data Release from this survey covered an area of 1.6 deg^2 in the Cosmic Evolution Survey (COSMOS) field and 3.5 deg^2 in the XMM Large-Scale Structure (XMM-LSS) field (Heywood et al. 2022), and provided the basis for investigations across the full range of science outlined above, including the discovery of two giant radio galaxies (Delhaize et al. 2021), investigations of the FIRC to high redshift (An et al. 2021; Delvecchio et al. 2021), probing the accretion rates to much lower radio luminosities than previously possible (Whittam et al. 2022) and measuring the contribution of SFGs and AGN to the radio sky background temperature (Hale et al. 2023).

In this paper, we describe the processing of the MIGHTEE data over an extended area in each of the COSMOS and XMM-LSS fields (which are the final MIGHTEE *L*-band images of these fields), along with a very deep single pointing exposure centred on the Chandra Deep Field South (CDFS, denoted in this work as CDFS-DEEP). This is known as MIGHTEE Data Release 1 (DR1). The deep observations of the CDFS field are taken as part of the Looking At the Distant Universe with the MeerKAT Array (LADUMA; Blyth et al. 2018) survey, whose main science goals are to determine the role of neutral atomic Hydrogen in galaxy evolution. We describe the observations and data processing, including a description of the effective frequency maps over all fields, in Section 2. We outline how the radio source and component catalogues are constructed in Section 3. In Section 4, we provide an overview of the quality of the data in terms of astrometric uncertainties and the flux-density scale, and in Section 5, we present the source counts across the three fields, which comprise the deepest radio continuum data ever released over

several degree-scale areas of sky. We also present results on the level of sample variance between the three fields as a function of flux-density and survey area. We summarize the results of these analyses and draw conclusions in Section 6.

2 OBSERVATIONS AND DATA PROCESSING

This work makes use of a total of 86 individual MeerKAT tracks at *L*-band (856–1711 MHz) totalling 709.2 h, with 22, 45, and 19 pointings respectively in COSMOS, XMM-LSS, and CDFS-DEEP. The total on-target time in each field is 139.6 h (COSMOS), 297.9 h (XMM-LSS), and 126.7 h (CDFS-DEEP). For the COSMOS and XMM-LSS fields, the images released in this DR1 are the final *L*-band observations for these fields. A larger area of the CDFS field will be released in a future *L*-band data release, for which the expected coverage is shown in appendix C1 of Heywood et al. (2022) and will consist of 291 h of on source time. Alongside this there will additionally be *L*-band observations of the European Large Area ISO Survey (ELAIS) S1 field, totalling 72 h on source. Further details of the observations and pointings for each field are provided in Appendix A in Tables A1–A3. The data processing took place on a per-track basis up to the mosaicking point, and followed a similar procedure to that described in detail by Heywood et al. (2022), with scripts available online in the OXKAT repository (Heywood 2020). As for the ES data of Heywood et al. (2022), primary calibrators were observed at least twice within a block for 5–10 min, with secondary calibrators visited regularly throughout the target observations but for a shorter duration (∼2–3 min). The underlying software packages were containerized using SINGULARITY (Kurtzer, Sochat & Bauer 2017), and the majority of the data processing took place on the *ilifu* cluster,[1] in Cape Town. A brief summary of the procedure is given as follows.

2.1 Flagging, calibration, and imaging

The data were retrieved in Measurement Set format from the South African Radio Astronomy Observatory archive,[2] and were immediately averaged down to 1024 frequency channels (836 kHz in width) in the process.[3] Following initial flagging of the data, bandpass, delay and time-dependent gain solutions were derived using the primary and secondary calibrators using CASA (CASA Team et al. 2022). The solutions were applied to the target data, which were flagged using the TRICOLOUR software (Hugo et al. 2022), which implements the SUMTHRESHOLD algorithm (Offringa et al. 2010) optimized for MeerKAT and accelerated using DASK-MS (Perkins 2022).

The flagged and reference-calibrated target data were then imaged using WSCLEAN (Offringa et al. 2014), with deconvolution proceeding blindly, terminating with a relatively shallow threshold where the peak residual drops below 20 μJy beam^{-1}. The same algorithm that is implemented in the BREIZORRO (Ramaila, Smirnov & Heywood 2023) tool was then used to construct a mask from this initial image for a subsequent round of constrained deconvolution. The second round is deeper, forming a highly complete sky model, with the deconvolution within the masked region terminating at a peak residual of 1 μJy beam^{-1}. The model visibilities corresponding to the spectral clean component model derived from this second round of deconvolution were used as the basis for a single round of phase and delay self-calibration using the CUBICAL package (Kenyon et al. 2018), and the self-calibrated data were re-imaged.

Visual inspection occurs at this stage, and if it is deemed necessary to subtract a strong problematic source from the field then the data were re-imaged at higher spectral resolution with a refined deconvolution mask, and a direction-dependent peeling,[4] step was performed using CUBICAL. In total, peeling was required for 14 of the 45 pointings in XMM-LSS, only one of the 22 pointings in COSMOS, and all of the 19 pointings in CDFS-DEEP.

Whether a source has been peeled or not, the final stage of the processing for each MeerKAT track involved imaging of the data using DDFACET (Tasse et al. 2018, 2023). Residual direction-dependent errors (DDEs) were solved for by manually identifying 10–20 sources across the field of view. These sources are generally ones that are bright enough to provide high signal-to-noise ratio (SNR) gain solutions, and exhibit residual point spread function-like structures due to the the DDEs. At *L*-band, the DDEs are dominated by the antenna primary beam pattern coupled with stochastic pointing errors that differ from antenna to antenna, thus this subset of sources is generally uniformly distributed around the flank of the main lobe of the primary beam, and into the sidelobes. A tesselation algorithm partitions the sky according to their positions, and direction-dependent gain terms are solved for on a per-tessel basis, with solution intervals of 5 min in time and 107 MHz in frequency, using the KILLMS package (Smirnov & Tasse 2015; Tasse 2023). Two subsequent rounds of imaging with DDFACET applying these directional solutions resulted in the final images, with each run producing an image at two different resolutions via two different values (0.0 and −1.2) of the Briggs (1995) robust parameter.

2.2 Mosaicking

For each target field (including CDFS-DEEP), the images produced from each individual MeerKAT observation were combined in the image domain. First, the constituent pointings were all homogenized to a common resolution, namely a circular Gaussian beam with a full width at half-maximum (FWHM) dictated by the largest value of the fitted beam major axis from the set of images of a given field. The clean component model image was directly convolved with this Gaussian, whereas the residual image was convolved with a homogenization kernel computed using the PYPHER package (Boucaud et al. 2016) prior to being summed with the convolved model. Following this, each convolved image was primary beam corrected, using a model of the MeerKAT Stokes *I* beam evaluated at the nominal band-centre frequency using the KATBEAM[5] library. The Stokes *I* beam model was also azimuthally averaged in order to average out the asymmetries in the MeerKAT main lobe. The resulting homogenized, primary beam corrected images were then linearly mosaicked using the MONTAGE[6] toolkit, using the square of

[1] https://www.ilifu.ac.za/
[2] https://archive.sarao.ac.za/
[3] The total number of channels in each observations are given in Tables A1–A3.
[4] The 'peeling' process has a firm original definition (Noordam 2004) however it is commonly (mis)used to describe any process that removes a problematic source from the visibility database via a method that is more advanced than a straightforward model subtraction. We are also guilty of misusing the term here, as the subtraction of problematic sources from the field of view in our processing is more accurately described as a differential gains approach (Smirnov 2011) with only a single additional solvable gain term in the Measurement Equation.
[5] https://github.com/ska-sa/katbeam
[6] http://montage.ipac.caltech.edu/

the primary beam (equivalent to variance weighting) as a weighting function.

A comparison of the final images for each of the three fields at both Briggs' weightings are presented in Fig. 1. We also show a comparison of selected regions within this data release compared to the MIGHTEE ES data in Fig. 2. These regions were chosen to include areas near the edge of the Heywood et al. (2022) ES image, as these areas are mosaicked with additional data in this data release and therefore can show an improvement in sensitivity. For COSMOS and XMM-LSS, the DR1 images are the final *L*-band continuum images from the MIGHTEE survey. For further details of these additional pointings, see appendix C of Heywood et al. (2022).

2.3 Effective frequency maps

Following Heywood et al. (2022), effective frequency maps are generated for each field to provide information of the typical frequency at a source location when effects such as the primary beam and flagging are accounted for. Such maps were generated using the same methodology as for the ES data by using linear mosaics of the frequency values in eight sub-bands, with a weighting function that is the primary beam response evaluated at that sub-band frequency, and weighted further by the inverse of the square of the sub-band image noise (please refer to section 3.2 of Heywood et al. 2022, for a more detailed explanation).

We present these effective frequency mosaics for each of the three fields in Fig. 3. The patterns caused by the overlap of the different pointings in the COSMOS and XMM-LSS fields are clearly visible and these maps are found to have typically lower effective frequencies in the outer regions of the image mosaics. The effective frequencies of the high-resolution images have, in general, lower frequency values compared to the lower resolution images. This is due to the fact that the lower resolution images are produced with a higher weighting of the short spacings, which are more prone to radio frequency interference (RFI, see Mauch et al. 2020). Since the majority of the RFI in MeerKAT's *L* band occupies the lower half of the band, this serves to raise the effective frequency of the lower resolution maps.

Whilst the effective frequency is reported within the catalogues of sources (see Section 3), we do not apply any corrections to the flux density of the images or those recorded in the catalogue. Any such correction would need to assume a spectral shape for the source, given by the spectral index, α.[7] Different scientific studies and use cases may require the assumption of different spectral indices or may want to measure the spectral indices directly. We therefore leave it to the users of to appropriately apply any frequency corrections to the measured source flux densities, or images, as necessary for their work. Not accounting for such an effect could lead up to an $\sim$8 per cent offset in the ratio of the flux densities from the minimum and maximum frequency values across the image, when scaled to a common frequency (and assuming $\alpha = 0.7$).

3 RADIO CATALOGUES

To detect radio sources and create a catalogue of their source properties, we make use of the Python Blob Detection and Source Finder (PyBDSF; Mohan & Rafferty 2015).[8] The general practice of PyBDSF makes use of a sliding box to first derive root mean square (rms) and sky background maps for an image, after which islands of emission which exceed a user determined threshold are identified. For these islands, the emission is modelled using either a single Gaussian component or a combination of multiple Gaussian components, which are used to build up a source model. For the ES MIGHTEE data (Heywood et al. 2022), PyBDSF was used with a typical standard configuration where a 3σ threshold on flux within the image to detect islands of emission was used to ensure high completeness for both point and extended sources, but only those sources which satisfied at least a 5σ peak emission threshold criteria were included in the final source catalogue,[9] alongside the associated Gaussian component catalogue.

For this work, we use PyBDSF in a modified method relative to the ES data to produce the source and component catalogues. Such a change is employed in order to extract maximal information from the images, which are limited by confusion at the sensitivity and resolution of these data products. A similar approach is also being used in deep radio observations with LOFAR (Shimwell et al., in preparation). As discussed in Heywood et al. (2022) and in Section 2, the MIGHTEE continuum imaging data are affected by confusion and as such, this will affect our ability to measure the noise in the image reliably. When determining the rms maps, PyBDSF uses a sliding box to generate a varying rms map across the full field. The size of such a box can either be determined by PyBDSF itself or specified by the user.[10] Such an rms map is generated prior to the identification (and hence removal) of any sources already within the image. Therefore, when there is a high source density, the measured rms will increase due to the sources responsible for elevating the confusion noise. An elevated rms due to this source confusion will adversely affect the detection of sources across the image, reducing the number of sources which reach the threshold needed to be considered 'detected' by PyBDSF. Due to the high source density in the MIGHTEE images we therefore adopt a two-step process for the source finding and catalogue generation, first removing sources from the image to calculate a more accurate rms map and using this rms map for source detection.

More formally, we first run PyBDSF on the image using the detection parameters:

```
bdsf.process_image(image, thresh_isl = 3.,
thresh_pix = 5., mean_map =`zero', rms_box =(120,30))
```

where `image` is the file name of the image, `thresh_isl` is the threshold for detecting an island of emission to then be fit with Gaussian components, and `thresh_pix` is the criterion which determines which sources are included within the final source catalogue. The `mean_map` is the same shape as the input image, reflecting the background emission within the image. We set this to zero at all locations within the image. Finally, `rms_box` is a parameter used to quantify the size of the box used to determine the rms in terms of pixels (where our pixel scale within the images are 1.1 arcsec), as well as the step size used to move this sliding box across the image. Such values were chosen to ensure consistent box sizes were used across the three fields and to avoid too large a box size which could smooth over noise variations.

[7] Where a spectral index is assumed in this work we use the relation $S_\nu \propto \nu^{-\alpha}$ to relate the frequency, ν, to the spectral flux-density measurements, S_ν and will assume a value of $\alpha = 0.7$, unless otherwise stated.

[8] https://pybdsf.readthedocs.io/

[9] Note this 5σ threshold is based on image values, and so some catalogued sources may appear to have peak SNR slightly less than 5σ.

[10] See the guidance at https://pybdsf.readthedocs.io/en/latest/process_image.html#term-rms_box for appropriate rms box scales to use.

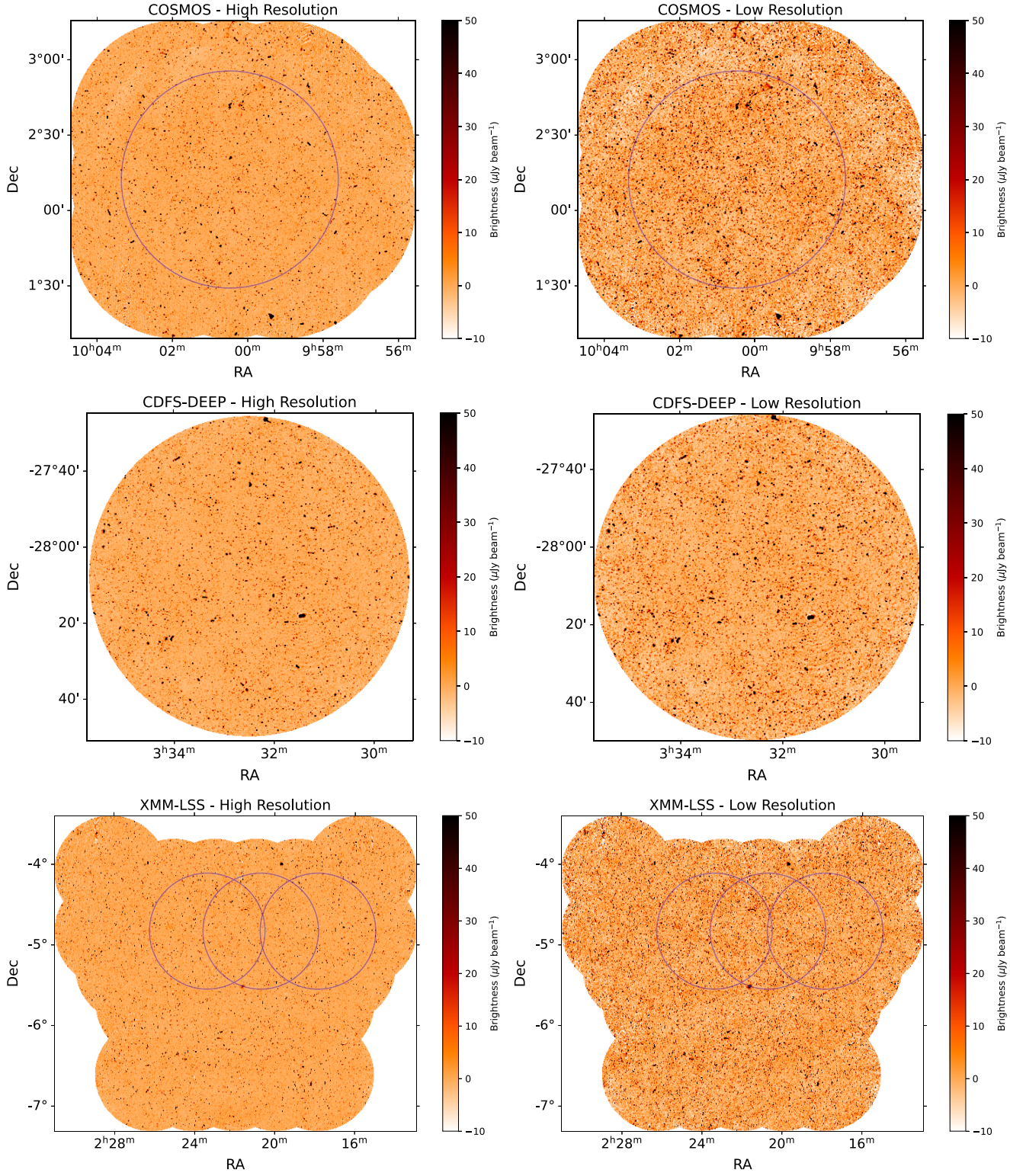


Figure 1. The brightness distributions for each of the MIGHTEE DR1 images which are released with this work. This is shown for COSMOS (top row), the deep pointing in CDFS (middle row), and XMM-LSS (bottom row) for the high-resolution images (left: 5.2, 5.5, and 5.0 arcsec beam FWHM, each field respectively) and the low-resolution images (right: 8.9, 7.3, and 8.9 arcsec beam FWHM, respectively). For the COSMOS and XMM-LSS fields, the extent of the pointings of the ES data of Heywood et al. (2022) are indicated by purple circles.

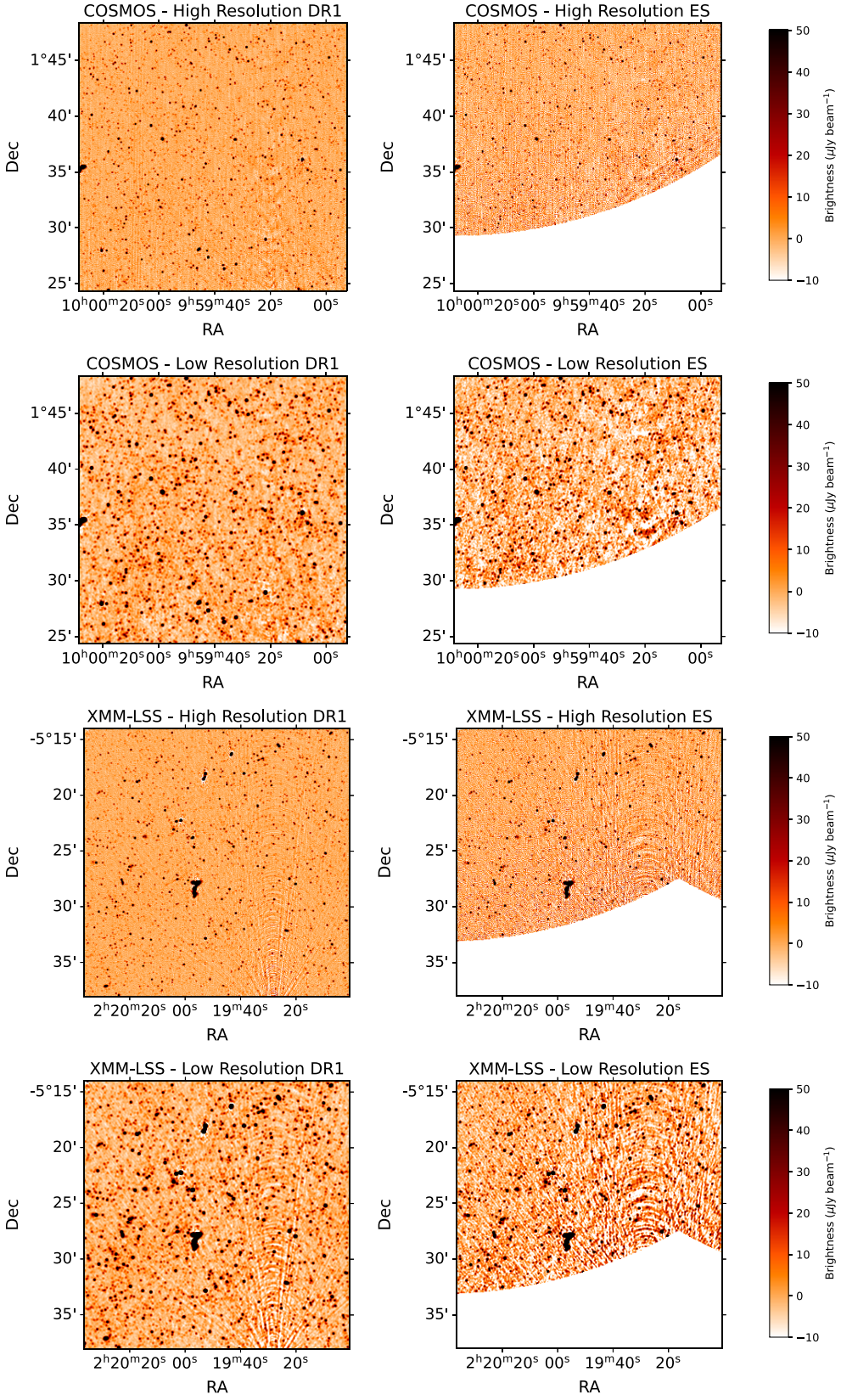


Figure 2. Comparisons of the brightness images for a zoomed-in region of the DR1 images (left) released with this work compared to the same region within the ES data (right) of Heywood et al. (2022) for the COSMOS high-resolution images (first row) and low-resolution images (second row), and the high-resolution images of the XMM-LSS field (third row) and the XMM-LSS low-resolution images (fourth row). The regions are chosen to cover $\sim 0.4^\circ \times 0.4^\circ$ over a region that is towards the outer regions of a pointing in the ES data to the south of the field which has not been mosaicked with neighbouring pointings, as it has been in DR1.

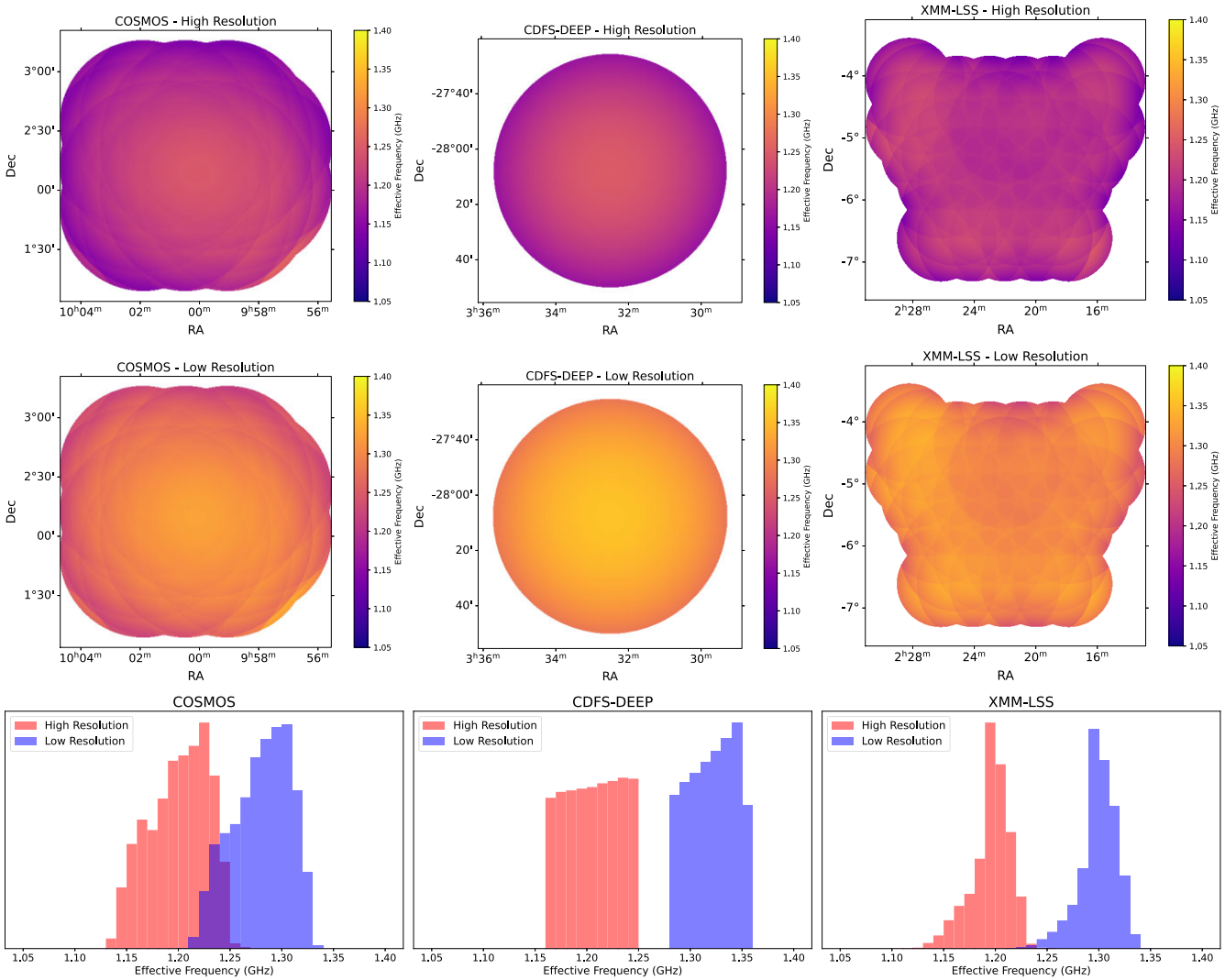


Figure 3. Effective frequency maps for COSMOS (left), CDFS-DEEP (centre), and XMM-LSS (right) fields for the high-resolution images (top row) and low-resolution images (middle row). Also shown (bottom row) are the histograms of the effective frequency pixel values within the images for the high (red) and low (blue) resolution images.

After running PyBDSF, we save the residual map which is generated from PyBDSF, i.e. the image with the source model of the Gaussian components of 'detected sources' subtracted. This should therefore give a better representation of the background, with bright sources removed. We again use PyBDSF, with the parameters described above, on the residual image to generate the rms map. This rms image generated in the second run of PyBDSF should provide an improved estimate of the true rms of the image, with a reduced number of sources. Of course, due to the confused nature of the MIGHTEE data, there will be remaining sources of emission within the residual image and, as such, the rms will still be elevated, though this will be reduced compared to the rms measured using the initial PyBDSF run. This rms image is then used to improve the source detection for the final catalogues.

To obtain the final source and component catalogues, we perform a final run with PyBDSF in which we supply the original MIGHTEE image and use the parameters described above, but supply the rms map generated above as an input (`rms_map`), instead of allowing PyBDSF to internally generate such a map. To do this, we use:

```
rmsmean_map_filename = [mean_map, rms_map]
```

where the `mean_map` is again a map consisting solely of zeros over the image. From this final PyBDSF run, we obtain source and Gaussian component catalogues for each of the three fields for both the low- and high-resolution images. The total number of sources and components detected by PyBDSF are shown in Table 1.

As shown in Table 1, we detect a total of 143 837 sources (154 314 Gaussian components) across the three fields (totalling 20.1 deg^2) in the lower resolution images and 114 225 sources (121 647 Gaussian components) in the corresponding higher resolution images. We release both the source and Gaussian-component catalogues separately for the three fields and provide an example of the first five rows of the source catalogue from the COSMOS field in Table B1. The Gaussian catalogue contains the same columns as the source catalogue, with an additional Gaussian component ID. Measured properties in the Gaussian catalogue such as flux densities, positions, and shapes relate to the individual Gaussian components which are used to model a source, unless otherwise stated, as described in https://pybdsf.readthedocs.io.

Table 1. Tables describing the images and catalogues of the MIGHTEE DR1 catalogues and images. Given are the area and resolution of the images, alongside the number of sources and Gaussian components detected with PyBDSF, as described in Section 3. Additionally, the thermal rms noise within the image, the measured median rms sensitivity within the PyBDSF determined rms maps across the full image, and within a smaller $0.5^\circ \times 0.5^\circ$ central region of the image, are also provided. The latter estimate demonstrates the typical sensitivity in the most sensitive regions that are not affected by the sensitivity roll off in the outer, non-mosaicked, regions. In the case of the mosaics, the thermal noise is estimated by mirroring the negative values in a histogram of pixel brightnesses and determining the standard deviation of a Gaussian fit to the resulting distribution. In the case of the CDFS-DEEP images, the thermal noise is measured in a region far from the phase centre where the primary beam attenuation is sufficient to provide a source-free region. For the COSMOS and XMM-LSS fields where there is ES data from Heywood et al. (2022), we also give a comparison of the area observed and number of Gaussian components detected.

	CDFS-DEEP		COSMOS		XMM-LSS	
Resolution (arcsec)	5.5	7.3	5.2	8.9	5.0	8.9
Area (sq. deg)	1.5	1.5	4.2	4.2	14.4	14.4
Thermal rms sensitivity (μJy beam^{-1})	0.9	0.6	2.2	1.6	3.4	1.5
Median measured rms sensitivity (μJy beam^{-1})	1.9	2.0	5.6	3.5	5.1	3.2
Median measured central rms sensitivity (μJy beam^{-1})	1.2	1.3	2.4	2.1	3.6	2.7
Number of sources	21 152	17 866	20 886	28 267	72 187	97 684
Number of Gaussian components	22 660	19 291	22 420	30 466	76 567	104 557
Ratio of area, cf. Heywood et al. (2022)	–	–	2.8	2.8	4.1	4.1
Ratio of Gaussian components, cf. Heywood et al. (2022)	–	–	4.9	3.0	7.1	5.1

4 CATALOGUE VALIDATION

In this section, we present comparison of the astrometry and flux density of our data to previous radio observations. As noted in Hale et al. (2023), differences in the MeerKAT baseline distribution compared to other telescopes or survey configurations could lead to emission being resolved into multiple sources in one survey (where only one source is detected in another survey), or extended emission being resolved out entirely. This may affect the comparison of flux densities of sources which are resolved. We therefore compare the COSMOS and XMM-LSS fields to the MIGHTEE ES data over these fields from Heywood et al. (2022). These cover overlapping areas, but the mosaicking of many more pointings could affect the measured properties of sources even within the regions which were covered by the ES release, e.g. primary beam corrections between different observation tracks and smearing with distance from phase centres is exacerbated by the mosaicking. Moreover, the improved sensitivity within regions of DR1 that were previously the outer edges of the ES data will now be less affected by the larger noise peaks/troughs in the ES data, which would affect the accuracy of the flux-density measurements. This means that the measurement of flux densities for sources within the ES region of MIGHTEE (especially towards the outer edges) may differ from the measured properties of the same sources in the DR1 region. As we use different methods to detect and characterize sources (albeit both with PyBDSF), this may also introduce differences in the measured source properties. In their work, Heywood et al. (2022) make comparisons of the ES data to Very Large Array (VLA) observations of the COSMOS (Schinnerer et al. 2010) and XMM-LSS fields (Heywood et al. 2020) using catalogues extracted with the same method as in Heywood et al. (2022). We will discuss their findings in the appropriate sections of astrometric (Section 4.2) and flux-density (Section 4.3) accuracy. Details of the location of each of the individual pointings that make the final mosaicked images are presented in Appendix A. The outline of the ES regions compared to DR1 can be seen in Fig. 1.

For the deep data within the CDFS-DEEP field, there are no previous MIGHTEE data to compare with. We therefore compare to observations using the VLA (Miller et al. 2013) and the Australian Telescope Compact Array (ATCA; through the Australia Telescope Large Area Survey, or ATLAS Data Release 3; Franzen et al. 2015), both at 1.4 GHz.[11] The overlap between the coverage of the MIGHTEE CDFS-DEEP data and the data of Miller et al. (2013) and Franzen et al. (2015) is shown in Fig. 4. These data are significantly shallower than the MIGHTEE data, with the catalogues of Franzen et al. (2015) having significantly lower source density, but covering a larger area than that of Miller et al. (2013).

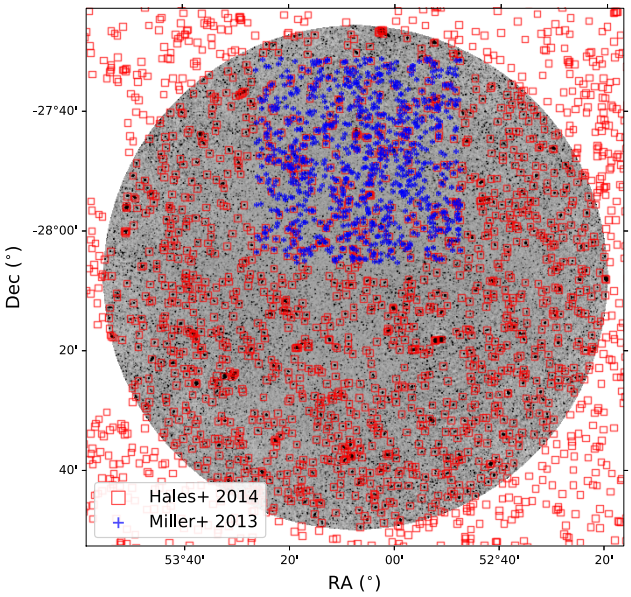


Figure 4. Comparison of the overlap region between sources in CDFS from MIGHTEE DR1 (grey background image) to previous observations with the VLA (blue plus symbols; Miller et al. 2013) and from ATCA with the ATLAS survey (red squares; Franzen et al. 2015).

[11]We note that the ATLAS catalogue also has frequency variations across the field of view between $\sim$ 1.4−1.5 GHz and therefore we use the 1.4 GHz scaled flux densities from their catalogue.

4.1 Matching of catalogues

To ensure a like-for-like comparison, we identify isolated sources within the DR1 catalogue (and the matching catalogue) which have no neighbours within the 2×FWHM of the restoring beam of the MIGHTEE data (which is typically the limiting resolution). This is to avoid sources which may be blended in one catalogue but not the other. The only exception to this is for comparison to the work of Franzen et al. (2015), where the angular resolution is larger than that of MIGHTEE, 16 arcsec × 7 arcsec. For these comparisons, the FWHM assumed is 16 arcsec and the appropriate cut from that is instead adopted. We then match the DR1 catalogues to the survey used for comparison, using a matching radius given by the angular resolution of the MIGHTEE data and only compare those sources which are defined as 'single' sources in PyBDSF (`S_Code = S`). We make a second catalogue with more stringent constraints, where only unresolved, high SNR sources are compared. To do this, we restrict to sources which have high SNR in both the DR1 and comparison data (peak SNR $\geq$10) and only use sources which are defined as unresolved in MIGHTEE, using the method implemented in Hale et al. (2021) and similar studies (e.g. Bondi et al. 2008; Smolčić et al. 2017a; Heywood et al. 2022), where the SNR is used to define an envelope between the integrated-to-peak flux-density ratio (S_I/S_P) as a function of SNR, within which sources in MIGHTEE are considered unresolved. Sources that were identified as single sources in PyBDSF were used to define such an envelope, which would contain 95 per cent of unresolved sources, and the envelopes used to define unresolved MIGHTEE DR1 sources are given in Appendix C, in Table C1.

4.2 Astrometry

In Fig. 5, we present a comparison of the astrometry between the DR1 and ES observations of Heywood et al. (2022) for the COSMOS and XMM-LSS fields using the two comparison catalogues described above for both the low- and high-resolution images. In the ES data, Heywood et al. (2022) find subpixel offsets in RA and Dec., with mean offsets (RA, Dec.) of (−0.27 arcsec, −0.19 arcsec) in the COSMOS field, and (−0.20 arcsec, −0.43 arcsec) in the XMM-LSS field. The astrometric comparison to multiwavelength data is also presented in Whittam et al. (2024), with a mean offset of (−0.24 arcsec, −0.40 arcsec) in a subregion of the COSMOS field. As shown in Fig. 5, the median astrometric offsets for our more permissive comparison catalogue are typically limited to $\lesssim$0.06 arcsec, with the uncertainties derived from the 16th and 84th percentiles $\sim 0.3 - 0.4$ arcsec. The density of sources which can be seen for the astrometric offsets demonstrate that most offsets are well within 1.1 arcsec (the image pixel size in both the high- and low-resolution images), suggesting that any effects of smearing due to the mosaicking of pointings are having a negligible impact on the measured locations of sources. The comparison of the more restrictive matched catalogues (with the additional SNR and unresolved criterion) show a similar offset magnitude compared to the more permissive catalogue, with a reduction in the percentile derived uncertainties to $\sim 0.15 - 0.2$ arcsec. These comparisons demonstrate that the astrometry in DR1 is consistent with that of the ES data.

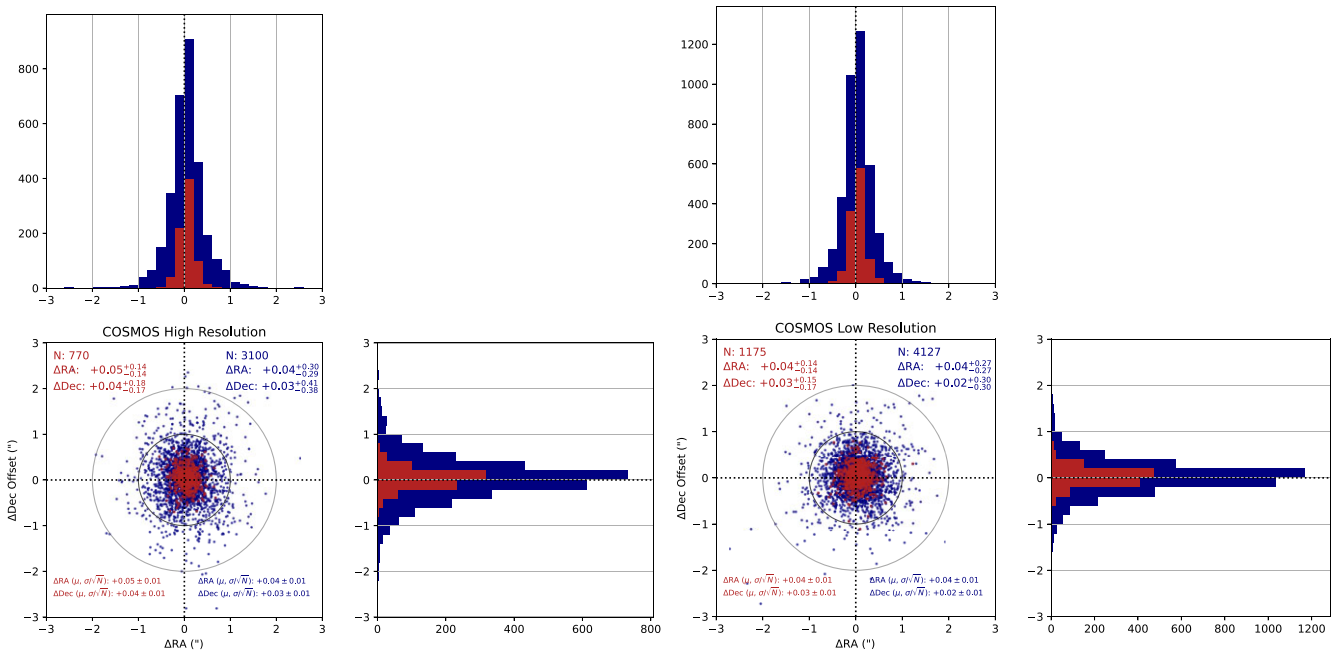
For CDFS-DEEP, we compare the MIGHTEE data to VLA (Miller et al. 2013) and ATCA (Franzen et al. 2015) imaging in Fig. 6. We note that there are far fewer sources to compare with in this field, due to the lower sensitivity and smaller area of the ATCA and VLA data, respectively. Such surveys from the VLA and ATCA have been compared to each other in Franzen et al. (2015) and find mean offsets in RA and Dec. of ∼0.1 arcsec for sources with SNR $\geq$ 20 in the ATCA data. The median offset to the ATCA data is ∼ 0.25 arcsec in RA and ∼ 0.35 arcsec in Dec., much smaller than the pixel resolution of both the high- and low-resolution images. Comparing with the VLA data, we find a similar offset in RA, but a smaller offset in Dec. of ∼ 0.20 arcsec. Both are well within the 1.1 arcsec pixel size of the MIGHTEE data. Whilst the majority of sources have consistent astrometry, within a pixel, when compared to the work of Miller et al. (2013), there are a significant number of sources which appear to have larger astrometric uncertainties when compared to the work of Franzen et al. (2015), which appear to be typically more dispersed in declination. As this is significantly less notable in the comparison with Miller et al. (2013), this may relate to the ATLAS observations, which have an elongated beam in the north–south direction (due to the array configuration) and so any offsets may be more likely in this direction. Franzen et al. (2015) do not see such large declination offsets compared to the data of Miller et al. (2013), which suggests this relates to differences in the selection criteria used to match sources to MIGHTEE.

As in the work of Whittam et al. (2024), the astrometric accuracy of these catalogues will be understood better through the cross-matching of the MIGHTEE detected sources to multiwavelength host galaxies. Such analysis will be presented in future work.

4.3 Flux-density comparisons

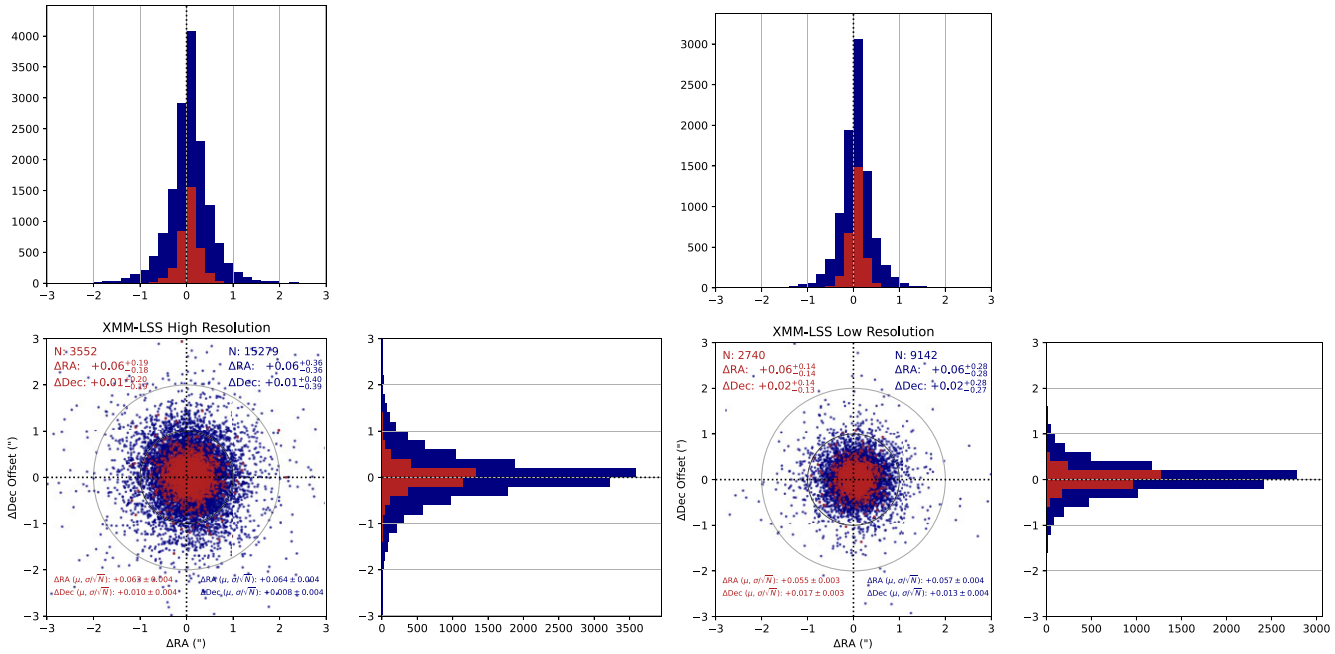
In this section, we compare the flux-density scale of these DR1 images to that of the ES data in the COSMOS and XMM-LSS fields in Fig. 7 for the more permissive cross-matched catalogue, and in Fig. 8 when the additional SNR and unresolved criteria are applied. As discussed, the ES data have been compared to previous surveys from the VLA in Heywood et al. (2022), where they run the same source finder, PyBDSF, over the VLA image data to make source comparisons using a consistent source finding process. Heywood et al. (2022) identify median flux-density ratios between the MIGHTEE and VLA data of 0.95 (1.0) for the COSMOS (XMM-LSS) field. These more restrictive cuts are likely to have a much larger effect on the flux-density comparisons. This is because low SNR sources are much more likely to be affected by confusion, Eddington bias (Eddington 1913) and measurement biases (see discussions for ES data in Hale et al. 2023) which will all affect the measured flux density of sources. This will be especially important in the outer regions of the ES data where the rms is larger and so the measurement of faint sources is more challenging. In DR1, the effect of mosaicing together multiple pointings reduces the noise in the outer regions of the ES area, which are now fully within the DR1 mosaic, as can be seen in Fig. 2. As these DR1 images have different effective frequency maps compared to the ES data, we scale the data to a common frequency of 1.4 GHz. We similarly use a frequency of 1.4 GHz when making comparison to the data of Miller et al. (2013) and Franzen et al. (2015) in the CDFS field.

As can be seen in Fig. 7, at the highest flux densities the comparison between ES and DR1 follows an approximate 1-to-1 relationship, with larger deviations from this found at lower flux densities. This scatter reduces when the additional SNR and unresolved criterion are applied (Fig. 8), with median ratios of the flux densities $\lesssim$ 3 per cent. In the COSMOS high-resolution images, though, this median offset is more similar to 7 per cent, with the reason for the larger offset unclear. In Figs 7 and 8, sources are coloured based on their peak SNR in the ES data, which highlights that the sources which significantly deviate from the 1-to-1 relation are at the lowest



(a) COSMOS High Resolution

(b) COSMOS Low Resolution

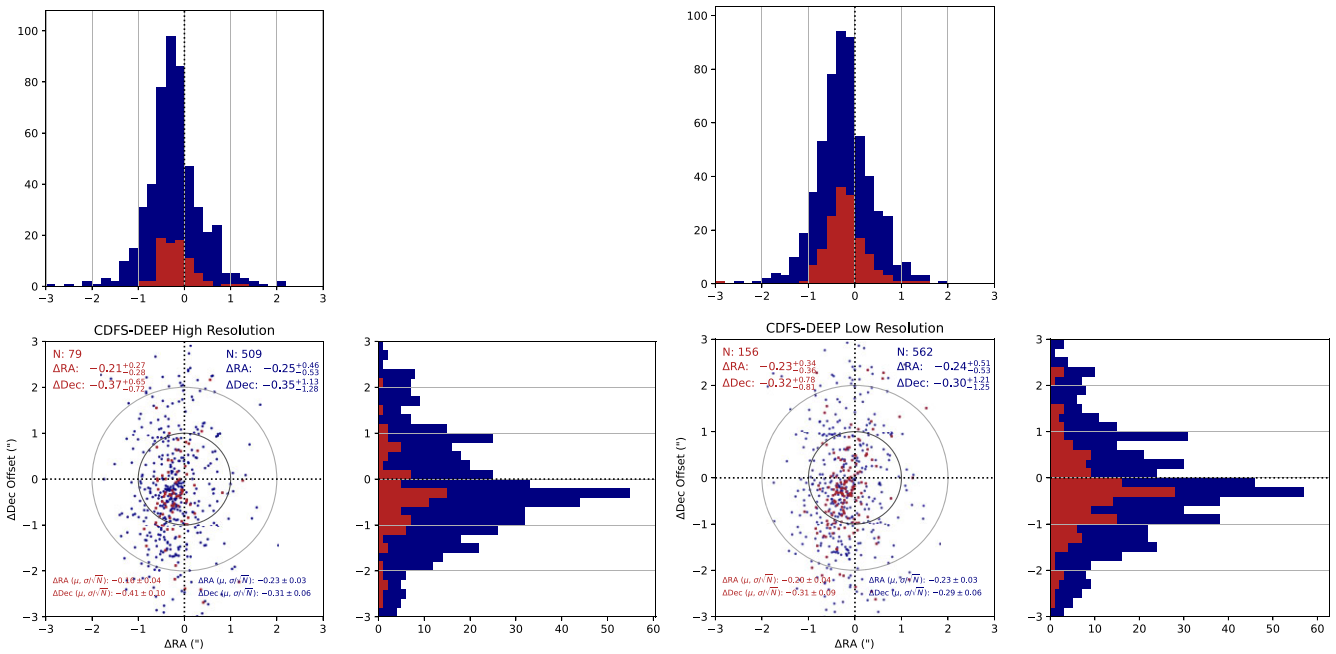


(c) XMM-LSS High Resolution

(d) XMM-LSS Low Resolution

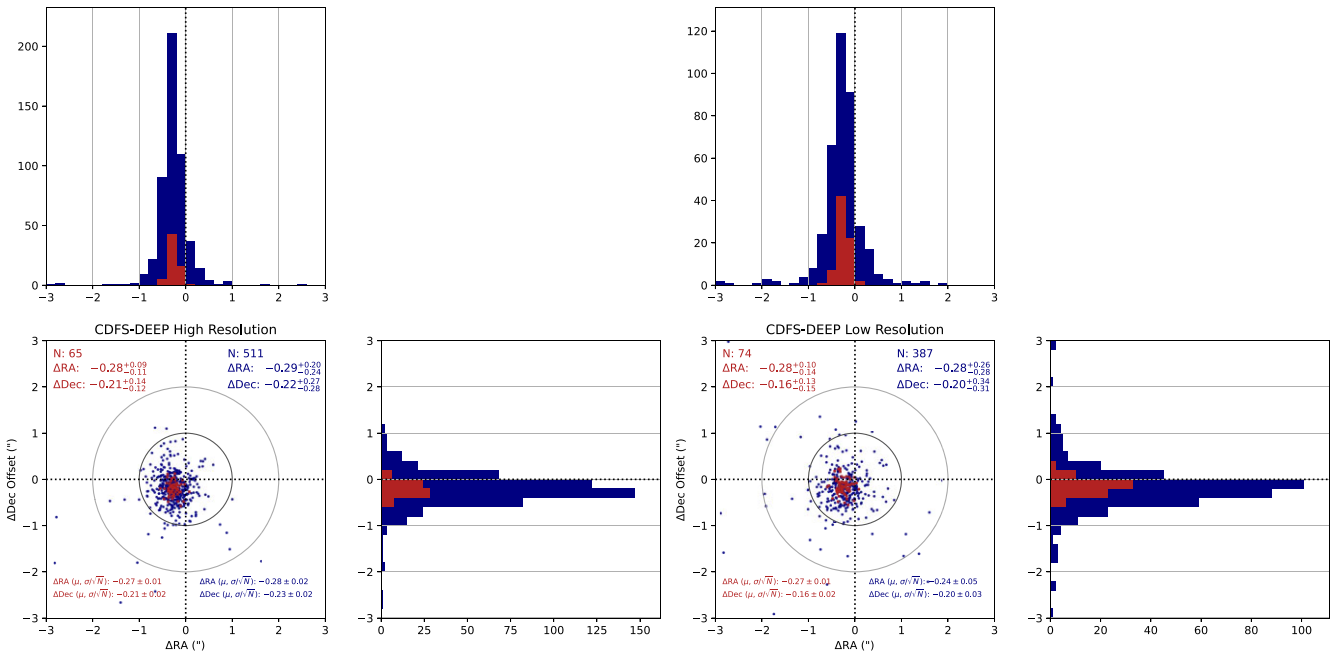
Figure 5. Comparisons of the astrometric offsets between DR1 and the ES data from Heywood et al. (2022) for the COSMOS (upper) and XMM-LSS (lower) fields, for the high- (left) and low-resolution (right) catalogues. Shown are the offsets for both the single-component sources, as discussed in Section 4 (navy) as well as with the additional SNR and unresolved criterion discussed in Section 4 (red). The numbers quoted in the upper part of the scatter panel presents the number of sources which contribute to the comparisons as well as the median offset in RA and Dec. (with $\Delta X = X_{\rm DR1} - X_{\rm ES}$, for both RA and Dec.) and the associated errors from the 16th and 84th percentiles reported in arcseconds. The colour of the text is chosen to match the data plotted. The circles of radii of 1 and 2 arcsec are included to guide the eye. The numbers provided in the lower part of the scatter plot quote the mean and standard error on the mean for the RA and Dec. offsets.

SNR ($\lesssim 10$). These sources appear to have an excess flux density in the ES data compared to DR1. This can be understood due to the higher noise of the ES data, which may cause an apparent boost in the flux densities due to Eddington bias compared with DR1. This is because fainter sources in the ES data may be preferentially boosted to be detected above the detection threshold, given the larger noise in the ES data. This causes the median ratio of the flux densities (shown in the histogram inset) to appear to deviate from the expected ratio of 1 by up to ∼20 per cent. In comparison, with the more stringent cuts applied to the data to ensure sources



(a) CDFS-DEEP High Resolution to Franzen et al. (2015)

(b) CDFS-DEEP Low Resolution to Franzen et al. (2015)



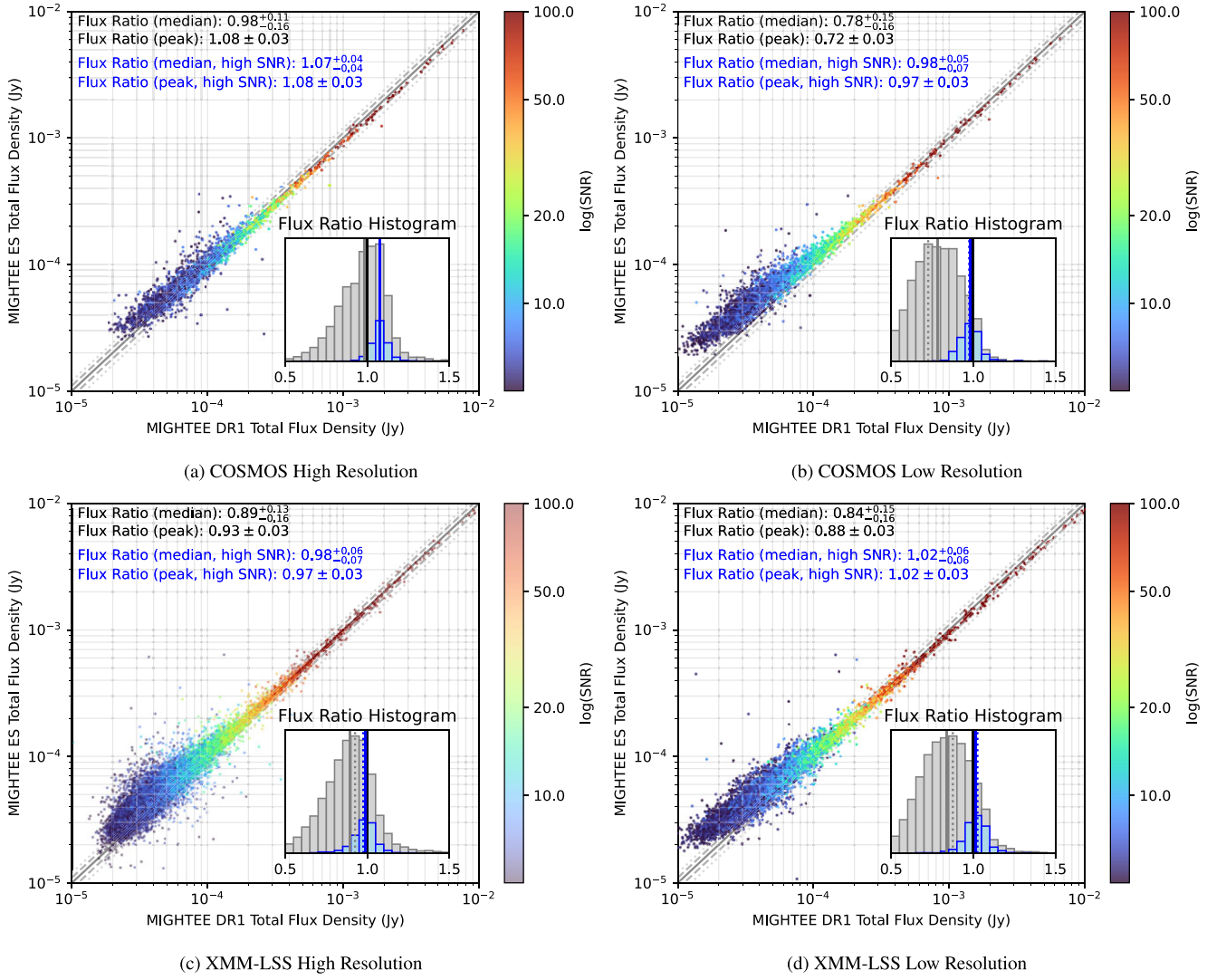
(c) CDFS-DEEP High Resolution to Miller et al. (2013)

(d) CDFS-DEEP Low Resolution to Miller et al. (2013)

Figure 6. As for Fig. 5 for the astrometric offsets, now for the CDFS-DEEP field for the high- (left) and low-resolution (right) catalogues compared to data from ATCA (Franzen et al. 2015, upper) and the VLA (Miller et al. 2013, lower).

are unresolved and at high SNR, the median flux-density ratio is closer to a ratio of 1 with median offsets typically constrained to within ∼10 per cent. The boosting of flux in the ES data compared to DR1 is more apparent in the low-resolution images where, as discussed, the combination of confusion and Eddington bias will be more prevalent.

For CDFS-DEEP, we again compare to the work of Miller et al. (2013) and Franzen et al. (2015). This introduces challenges in comparing flux densities due to the different baseline configurations between MeerKAT compared to the VLA and ATCA, which could lead to extended emission being resolved out, particularly in the VLA data. The observations of Miller et al. (2013) are much higher resolution (2.8 arcsec× 1.6 arcsec), whilst Franzen et al. (2015) have a more elongated beam, but at more comparable resolution (12 arcsec× 6 arcsec). For comparisons with Miller et al. (2013), it may therefore be the case that sources which are unresolved in MIGHTEE still have emission resolved out in the VLA data of Miller et al. (2013). Furthermore, there are also differences in the source



(a) COSMOS High Resolution

(b) COSMOS Low Resolution

(c) XMM-LSS High Resolution

(d) XMM-LSS Low Resolution

Figure 7. Comparisons of the integrated flux densities (scaled to 1.4 GHz) between DR1 and the ES data from Heywood et al. (2022) for the COSMOS (upper) and XMM-LSS (lower) fields, for the high- (left) and low-resolution (right) catalogues for the single-component sources matched with the matching criteria discussed in Section 4. Sources are coloured by their peak SNR in the ES data and the inset shows the histogram of the ratio of the flux-density measurement in the ES data compared to DR1 for all sources plotted (grey) and for sources with SNR in ES$\geq$ 20 (blue). For each histogram, the vertical lines are the ratio of 1 (black), and the coloured lines are the peak (dotted) and median (solid) values of the ratio, using the same colour scheme as the histogram. The black solid line indicates a 1-to-1 ratio, grey dashed lines indicate ratios of the flux densities of 0.9 and 1.1, and the grey dotted lines indicate flux-density ratios of 0.8 and 1.2.

finding algorithms that were used to generate the source catalogues which may also affect detections and characterizations of sources (see comparisons in e.g. Hopkins et al. 2015; Bonaldi et al. 2021; Boyce et al. 2023).

The comparison of the flux density of MIGHTEE CDFS-DEEP sources, when the more stringent cuts are applied, are shown in Fig. 9. This results in a small number of sources available for comparison. We find a good agreement between our observations and the catalogue of Miller et al. (2013), especially when making comparisons to the high-resolution MIGHTEE data, with a median offset of $\lesssim 5$ per cent. However, we find a larger offset in the source flux density compared to Franzen et al. (2015), with the MIGHTEE sources appearing to have an excess in flux density compared to those in the ATCA catalogue, with a median offset similar to $\sim$10 per cent. The reasons for any remaining differences are unclear, although we note the difference in resolution, baseline distributions and source extraction software could all have an impact. We note, though, that in their work, Franzen et al. (2015) find a typical 2 −3 per cent offset to the surveys of Miller et al. (2013), again suggesting that our work has a different selection of sources for comparison with MIGHTEE than used in their work.

5 COMBINED 1.4 GHzSOURCE COUNTS

In this section, we use our source catalogues to measure the Euclidean normalized source counts within the three deep fields.[12] We take a similar approach to a number of works which consider the completeness in radio images (see e.g. Smolčić et al. 2017a; Hale et al. 2021; Shimwell et al. 2022) and in the work of Hale et al. (2023) by injecting sources into an image and considering the recovery of

[12]Unless otherwise stated, the completeness and sources counts presented have been scaled to 1.4 GHz.

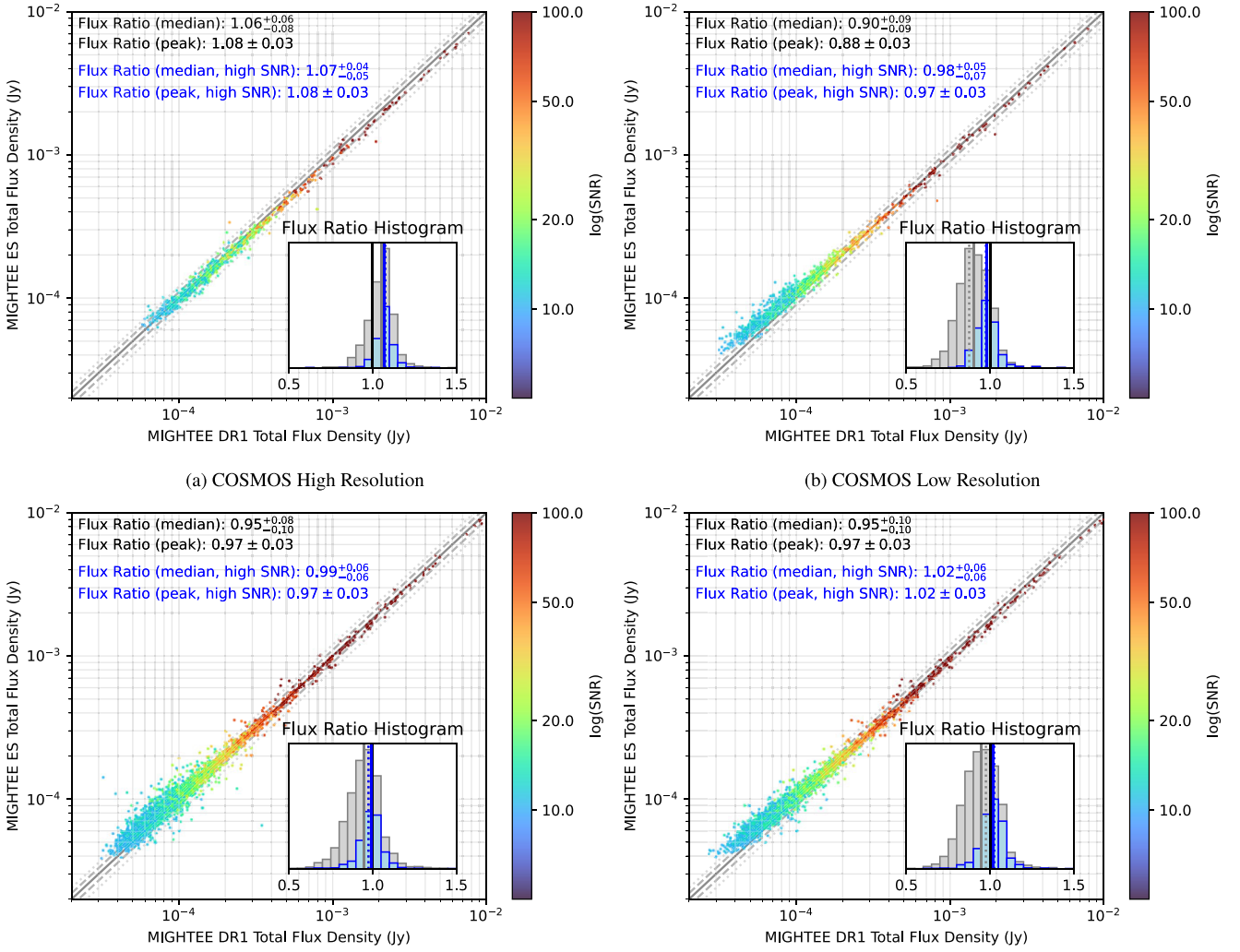
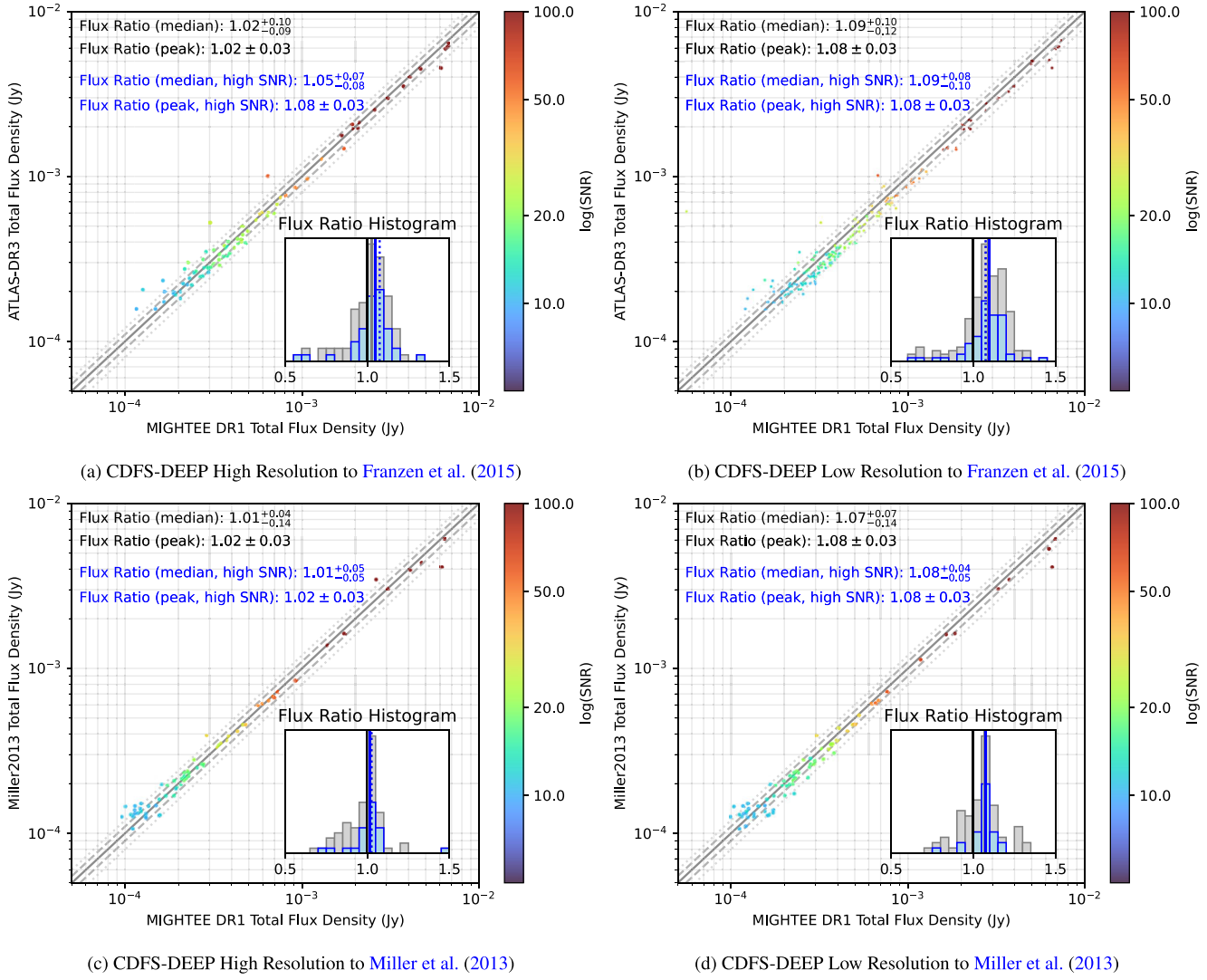


Figure 8. As for Fig. 7, the additional SNR and unresolved criteria applied to the COSMOS and XMM-LSS sources, as discussed in Section 4.

sources by the source finder, PyBDSF. However, such confused images present challenges to accurately measure and account for incompleteness. For example, any sources injected in the image would add to the confusion noise, whereas injecting sources into the residual image will mean that the effects of source confusion and source blending are likely to be underestimated. Furthermore, injecting a realistic number of sources into the residual image that is consistent with the observed source counts, will add to the background of $< 1\sigma$ sources which have not been removed from the image when determining the rms map, significantly increasing the sky density of these faint sources. Finally, one could also create images which do not make use of the MIGHTEE images (or residual images) and instead theoretically account for the rms variations and primary beam attenuation based on theoretical predictions from the telescopes known properties. However, such images also present challenges in accurately including observational limitations of e.g. artefacts, deconvolution issues as well as accurately including the true morphology of sources in the image.

In this work, we take the approach of directly injecting sources into the image and recovering sources. This allows us to better account for effects of confusion within this deep imaging however, as the image is already confused, it already contains a significant number of sources both above and below the detection threshold. This approach is likely to be the most accurate in estimating and accounting for the image and source detection systematics, provided that the number of sources injected into the image is restricted to minimize additional contribution to the confusion noise. For such simulations, we make use of the Tiered Radio Extragalactic Continuum Simulations (T-RECS, Bonaldi et al. 2019, 2023) and inject a number of sources directly into the original image (2000 in CDFS-DEEP, 2500 in COSMOS and 7500 in XMM-LSS, or $\sim$10 per cent of the detected sources per field) using a limiting 1.4 GHz flux density of 4 μJy in COSMOS and XMM-LSS and 2 μJy in CDFS-DEEP (as CDFS-DEEP is deeper, see Table 1). As the effective frequency changes across the field of view, we scale the flux density of simulated sources using the effective frequency (and a spectral index of $\alpha = 0.7$) at the random locations. Each source is then injected in the image using to different source models depending on the source type, discussed in Bonaldi et al. (2019, 2023). This makes use of models from the GALSIM package (Rowe et al. 2015) and injects using



(a) CDFS-DEEP High Resolution to Franzen et al. (2015)

(b) CDFS-DEEP Low Resolution to Franzen et al. (2015)

(c) CDFS-DEEP High Resolution to Miller et al. (2013)

(d) CDFS-DEEP Low Resolution to Miller et al. (2013)

Figure 9. Comparisons of the integrated flux densities (scaled to 1.4 GHz) between MIGHTEE CDFS-DEEP data to Franzen et al. (2015, upper panels) and Miller et al. (2013, lower panels), for the high- (left) and low-resolution (right) catalogues for the catalogues made with the more stringent criteria discussed in Section 4. Sources are coloured by their peak SNR in the non-MIGHTEE data and the inset shows the histogram of the ratio of the flux-density measurement in the non-MIGHTEE data compared to DR1.

the prescriptions used in the SIMUCLASS package[13] (Harrison et al. 2020).

After injecting these sources into the image, the final rms map discussed in Section 3 is used to re-extract the sources, using the final stage of the source finding prescription discussed in Section 3. In this work, our procedure of supplying the rms map used in the DR1 source catalogue generation ensures that the number of simulated sources injected in the image are not artificially inflating the measured noise in the image which would be measured by PyBDSF. We repeat this process to generate a total of 4000 simulations for each field.

As we supply the PyBDSF rms map which was used for source detection to produce the final continuum catalogue, the detected sources will be a combination of the original sources in the image and the simulated image. Therefore, when measuring the completeness we use a different approach to that of Hale et al. (2023). We do not impose any matching radius when comparing our simulated sources that we detect to those injected in the image. Instead we calculate the completeness by comparing the flux distributions of the total detected sources with the total number of injected sources as well as those within the image already. This allows us to better measure the true completeness, accounting for the blending and grouping of sources. More formally, we consider the total completeness (C) within a given flux-density bin (S to $S + \mathrm{d}S$) to be:

$$C(S, S{+}\mathrm{d}S) = \frac{N_{\text{Output sims}}(S, S{+}\mathrm{d}S) - N_{\text{sims}} N_{\text{Image}}(S, S{+}\mathrm{d}S)}{N_{\text{Input sims}}(S, S{+}\mathrm{d}S)}, \quad (1)$$

where $N_{\text{Output sims}}(S, S + \mathrm{d}S)$ is the number of sources detected by PyBDSF across all simulations with measured flux densities in the given flux-density bin ($S \rightarrow S + \mathrm{d}S$). $N_{\text{Image}}(S, S + \mathrm{d}S)$ is the number of sources within the DR1 catalogues (Section 3) in the given flux-density bin and N_{sims} is the number of simulations (4000 here for the default simulations). $N_{\text{Input sims}}(S, S + \mathrm{d}S)$ is the total number of simulated sources injected across all of these simulations

[13] https://github.com/itrharrison/simuclass-public/

Table 2. Details of the number (N) and minimum 1.4 GHz flux-density limit ($S_{\rm min}$) of sources injected into the images of the three fields (CDFS-DEEP, COSMOS, and XMM-LSS) fields respectively, for the completeness simulations used in Section 5.

Label	CDFS-DEEP		COSMOS		XMM-LSS	
	$S_{\rm min}$ (μJy)	N	$S_{\rm min}$ (μJy)	N	$S_{\rm min}$ (μJy)	N
Default	2	2000	4	2500	4	7500
High flux 0	10	1000	10	2500	10	7500
High flux 1	25	1000	25	2500	25	7500
High flux 2	100	1000	100	2500	100	7500

with simulated input flux densities within the given flux-density bin. The numerator accounts for the fact that the output catalogue for each simulation is comprised of both the simulated sources and the real sources. We note that the flux bins considered in this calculation are at 1.4 GHz and so the simulated input flux densities and measured output flux densities are first scaled to 1.4 GHz using the effective frequency maps, prior to the calculation of the completeness.

Calculating the total completeness using such a method accounts for the combined effects of (i) incompleteness due to sensitivity variations, (ii) Eddington bias (Eddington 1913), (iii) incompleteness due to the source finder, and (iv) measurement errors introduced from the source finder. It also accounts for the effects of confusion and the blending of sources which are in close proximity, as well as the effects of source splitting where for e.g. the lobes of AGN are identified as two separate sources, instead of being attributed to a single source. This would not be properly probed if an exclusion radius around sources already in the original image was imposed, which is important to measure the true incompleteness of the catalogues.

We note that, due to the expected source counts distribution, the majority of sources injected into the image will be faint. Therefore to better understand the completeness at brighter flux densities, we generate additional simulations where we adopt a higher minimum flux density for the simulated sources. The additional flux-density limits used and number of sources injected into the image are presented in Table 2. To combine the simulations together, we use the 'Default' simulations (Table 2) up to 10 times the minimum of the 'High Flux 0' simulations. For each of the subsequent 'High Flux' simulations, we use them to determine the completeness above 10× the minimum flux-density limit of the simulation, having run 1000 iterations for each simulation.

We estimate the associated uncertainties on the completeness measurements through combining subsets of simulations ($N_{\rm sims,\ subset}$) into independent samples. To determine how many simulations should be combined (denoted $N_{\rm sims,\ subset}$), we calculate the number of sources we expect to observe over the image area for each field, using the counts distribution of T-RECS (Bonaldi et al. 2019, 2023). We then compare how many simulations need to be combined (in each flux-density bin used for completeness, up to 10 mJy) such that the total number of sources matches the expected number of sources in T-RECS. Using the median of the integer number of simulations which should be combined, we generate a number of different 'combined' samples. We calculate the completeness in each of these combined samples, using equation (1) and use the standard deviation between these subsamples for the uncertainty, as a function of flux density. Typically this results in $\sim 30-40$ independent samples per field to calculate the completeness when using the default simulations and increases when the 'High Flux' simulations are used. For comparison, we also consider the completeness we would measure from the Default simulations up to the highest flux densities (> 0.1 mJy). Above 0.1 mJy all the simulations described in Table 2 are combined together to estimate the completeness and their associated errors.

The completeness measurements are shown in Fig. 10, colour coded by the simulations used. For each of the fields, these can be seen to rise steeply before peaking around a completeness of ∼1.2, and then declining to the expected value around a completeness of 1. This shows some differences between the completeness from combining the simulations together at 10× the minimum flux-density limit of the higher flux-density simulations, compared to only at the highest flux densities (> 0.1 mJy). We discuss the effect on the source counts below. At the highest flux densities, the completeness can appear to drop below a value of 1. This may be in part due to bright, jetted source being split into multiple components, but could also relate to the source models used by the simulation which if, for example, led to too large sources, may affect their completeness (see e.g. the discussion of source sizes in TRECS from Asorey & Parkinson 2021).

For the final corrected source counts for each field, we divide the source counts from sources within the catalogue of each field (with the flux densities scaled to 1.4 GHz) by the completeness measured in the given flux-density bin. We only consider flux densities within the range from 5× the minimum flux density of the 'Default' simulations (20 μJy for COSMOS and XMM-LSS, and 10 μJy for CDFS-DEEP) to 10 mJy. To determine the uncertainty on the source counts, we combine in quadrature the errors from both the Poissonian uncertainties from the number counts using the method of Gehrels (1986) with the uncertainties from the completeness simulations (we discuss the effects of cosmic variance in Section 5.1). The raw and corrected source counts for each field are presented in Fig. 11, alongside the model source counts from T-RECS (Bonaldi et al. 2023) and previous source counts from various surveys, converted to 1.4 GHz, from the works of: de Zotti et al. (2010), Smolčić et al. (2017a), Mauch et al. (2020), van der Vlugt et al. (2021), and Matthews et al. (2021a) and from the MIGHTEE ES data (Hale et al. 2023). An estimate of the subthreshold source counts using a $P(D)$ analysis in the MeerKAT DEEP2 field from Matthews et al. (2021a) is also shown. The results from the two methods used to combine the completeness simulations typically show small differeneces in the sources counts. A table of the raw and corrected source counts is provided in Table D1.

Fig. 11 shows that the corrected source counts for the deep pointing within CDFS are in good agreement with previous measurements of the source counts from der Vlugt et al. (2021), Matthews et al. (2021a), and Hale et al. (2023) (which are some of the deepest source counts currently available) and the simulated catalogues of Bonaldi et al. (2019). The source counts from COSMOS and XMM-LSS are lower than for CDFS-DEEP, across the flux-density range of $\sim 50-200$ μJy, but in agreement with Smolčić et al. (2017a). Notably, in this range the raw source counts are generally larger than the corrected source counts measurements, due to a completeness value>1. As discussed in Hale et al. (2023), a completeness greater than 1 reflects the movement of sources within flux-density bins, leading to an apparent excess of sources compared to the number of simulated sources that were injected. This will be in part due to Eddington bias, which preferentially boosts the more abundant fainter flux-density objects to high fluxes, whereas fewer sources are shifted to fainter fluxes resulting in a net positive gain in certain flux-density bins. This is also exacerbated by the merging/splitting of sources when they are in close proximity to each other (as may be the case in these images) and will also be affected by the fitting of source by PyBDSF. These effects are dependent on the underlying true

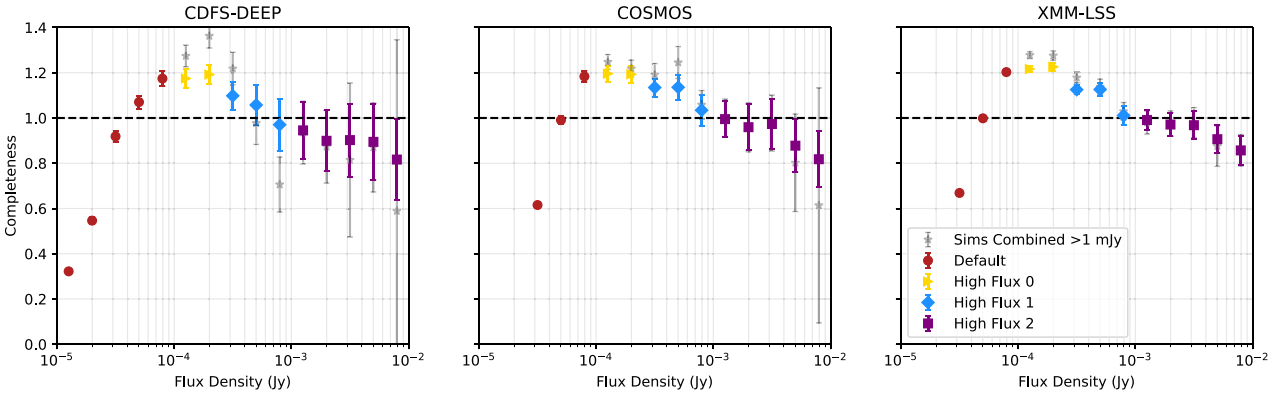


Figure 10. Measured completeness as a function of flux density as described in Section 5 for the CDFS-DEEP (left), COSMOS (centre), and XMM-LSS (right) fields. The completeness values associated with the different simulations in the analysis, as presented in Table 2 are indicated by different markers: Default (red circles), High Flux 0 (gold triangles), High Flux 1 (blue diamonds), and High Flux 2 (purple squares). Finally, the completeness when the default simulations is combined with the completeness from all the simulations in Table 2 together are also shown (grey stars).

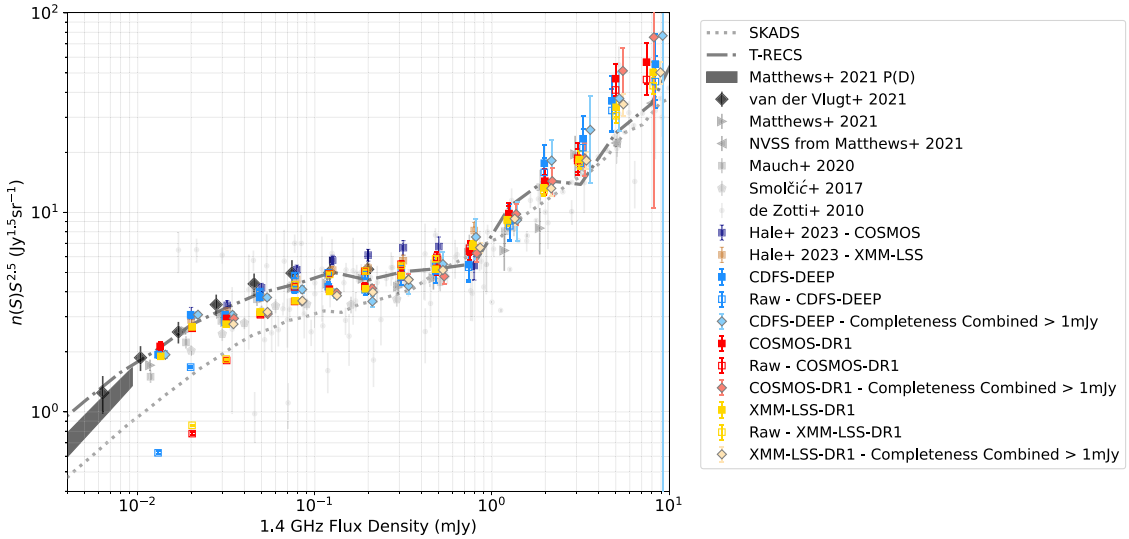


Figure 11. Euclidean normalized 1.4 GHz source counts for the COSMOS (red), CDFS-DEEP (blue), and XMM-LSS (yellow) fields for the MIGHTEE DR1 for both the corrected (filled markers) and raw (open markers) counts. The corrected source counts where the completeness is instead calculated from combining simulations above 1 mJy are also given as coloured diamonds with grey outlines and are artificially offset in flux density for visibility. Also shown are source counts from the T-RECS simulations, Bonaldi et al. (2019, 2023, grey dot–dashed line); the SKADS simulated catalogues, Wilman et al. (2008, grey dotted line) as well as counts from previous works of: Zotti et al. (2010, light grey circles), Smolčić et al. (2017a, light grey pentagons), Mauch et al. (2020, grey squares), Matthews et al. (2021a, grey triangles), and der Vlugt et al. (2021, black diamonds) as well as the subthreshold counts from the $P(D)$ analysis of Matthews et al. (2021a). Source counts are presented at the median flux density of observed sources within the bin.

source counts, as discussed in Hale et al. (2023), thus a reasonable model of the underlying sources counts is important to determine the correct completeness.

The results from DR1 are also generally lower in the range $S_{1.4} \sim 0.1 - 1.0$ mJy compared to the ES results in the COSMOS field (Hale et al. 2023). We note that the source counts in that regime for the COSMOS ES field were, in general, larger than both the models from T-RECS and some deep source count observations from previous surveys. This may be as a result of differences in accounting for the effects of confusion and the blending of sources in this work and the work of Hale et al. (2023) or may relate to differences in the flux scale when comparing the flux densities in the ES and DR1 data, as can be seen in Figs 7 and 8. As demonstrated, these figures appear to suggest an excess in the flux density of the ES data compared to DR1 for those sources which are at low SNR, which will lead to differences in the raw source counts. However, whilst in Hale et al. (2023) evidence was presented that a small number of sources with extended host galaxies were detected in MIGHTEE and not in the VLA 3 GHz COSMOS survey and so Hale et al. (2023) attributed the larger source counts observed in their work compared to Smolčić et al. (2017a) to be in part due to resolution bias. However, it may be the case that a combination of effects are in play. The VLA 3 GHz COSMOS survey may indeed be less sensitive to extended emission within the field, however it is also significantly less affected by the effects of confusion. In this work we have allowed confusion to be more readily included in our completeness calculations through not forcing any radial constraints on simulated sources, prior to the calculation of completeness. Therefore, confusion may be better accounted for in

this work compared to the ES source counts of Hale et al. (2023) and so this may explain the lower source counts in the COSMOS and XMM-LSS fields. However, the work of der Vlugt et al. (2021) also use much higher resolution observations than in this work but agree better with the results from the CDFS-DEEP field in this work. The source counts in CDFS-DEEP are also in better agreement with that of Matthews et al. (2021a) and Bonaldi et al. (2023). In reality, a combination of effects will contribute to the differences observed in the source counts presented in this work compared to previous studies both with the MIGHTEE ES data and other work. These include, differences in the source finding strategies, cosmic variance, flux-scale offsets, and differences in accounting for incompleteness and frequency variations across the image. Most notably, in this work accounting for the effects of confusion and source blending is challenging and, whilst we have presented a method to account for this, there may be remaining systematics.

5.1 Sample variance

As can be seen from Fig. 11 and discussed in Section 5, there are differences between the three fields which are a result of sample variance in the number of sources in different regions of the sky, although we note that there may be residual observational effects, e.g. flux offsets between fields, as well as the scatter due to the Poisson statistics alone. One of the key advantages of MIGHTEE is its ability to probe large regions of sky that will probe independent sightlines and therefore different parts of the large-scale structure. Larger sky surveys such as FIRST (Becker et al. 1995; Helfand, White & Becker 2015) and LoTSS (Shimwell et al. 2019, 2022) allow larger areas to be probed and sample variance to be overcome on much larger scales. However, these surveys are typically shallower than MIGHTEE, so can only constrain sample variance for brighter populations. In this section, we compare the variance in the source counts as a function of areal extent. We make comparisons to the work of Heywood, Jarvis & Condon (2013), who used the SKA Simulated Skies (SKADS, Wilman et al. 2008) catalogue to probe the expected sample variance.

To do this, we consider the angular areas used in Heywood et al. (2013) which can be sampled within the MIGHTEE area, across each of the three fields (i.e. area $\lesssim 1.5\,\mathrm{deg}^2$) and where we can also define $\geq$10 subregions. This restricts us to areas of 0.1, 0.3, and $0.5\,\mathrm{deg}^2$. For each area considered, we then estimate uncertainty due to sample variance using our three fields. In doing so, we average the effects of clustering and large-scale structure across a range of redshifts. Fig. 12 presents the subregions we use for this analysis. Following the same method as described in Section 5, we calculate the raw source counts along with the completeness over each subregion.[14] Using this, we estimate the fractional sample variance from the mean and standard deviation of the completeness corrected source counts over these subregions.

A comparison of the fractional sample variance (σ/μ) measured across the subregions is shown in Fig. 13 (top panel) as a function of flux density for the three areal regions considered (0.1, 0.3, and $0.5\,\mathrm{deg}^2$). We find a typical sample variance of 10–20 per cent for the areal subregions at $S_{1.4} \leq 100\mu$Jy. The sample variance increases in all cases to brighter flux densities as expected, given the lower source density and the expected higher galaxy bias for these sources (e.g. Lindsay et al. 2014; Magliocchetti et al. 2017; Hale et al. 2018; Mazumder, Chakraborty & Datta 2022), which will be dominated by AGN populations as opposed to the SFGs which dominate at fainter flux densities (see e.g. Smolčić et al. 2017b; Whittam et al. 2022; Best et al. 2023).

We also show this as a ratio to the work of Heywood et al. (2013) in Fig. 13 (bottom panel) over the same flux density and areas. We find that the fractional sample variance is comparable when measured to that by Heywood et al. (2013) from SKADS, across a wide range of flux densities ($\gtrsim 0.1$ mJy), increasing to a factor of 2−3 at fainter flux densities. Heywood et al. (2013) defined the sample variance ($\sigma_{S,\,\mathrm{per\,cent}}$) as the combination of the relative percentage errors due to clustering (or cosmic variance, $\sigma_{\mathrm{CV,\,per\,cent}}$) and the error due to the Poisson counts ($\sigma_{\mathrm{P,\,per\,cent}}$) of sources within the field:

$$\sigma^2_{S,\,\mathrm{per\,cent}} = \sigma^2_{\mathrm{CV,\,per\,cent}} + \sigma^2_{\mathrm{P,\,per\,cent}}. \quad (2)$$

However, we note that this is an idealized situation in which there are no incompleteness effects across the full area investigated. Our observations also include uncertainty from the completeness simulations on our measurements of the source counts. Therefore, to compare with the clustering uncertainty presented in Heywood et al. (2013), we should combine the variance from the Poissonian counts, cosmic variance, *and* completeness uncertainties. Thus, a direct comparison to the relative uncertainties due to cosmic variance (or clustering of sources) is challenging to measure. This is especially true here, as our completeness corrections vary for different subregions, principally due to the position within the field and their proximity to bright sources and regions of higher noise. Therefore, we are unable to conclude that the excess in sample variance at faint flux densities can be wholly attributed to an underestimate in the work of Heywood et al. (2013).

6 DISCUSSION AND CONCLUSIONS

In this work, we present catalogues and images of the COSMOS, XMM-LSS, and CDFS fields from the MIGHTEE survey DR1, using the MeerKAT telescope. These observations total $\sim$20 deg^2 across the three fields with 1.5 deg^2 in CDFS-DEEP, 4.2 deg^2 in COSMOS, and 14.4 deg^2 in XMM-LSS. Images are released at two angular resolutions, one prioritizing resolution ($\sim$5 arcsec) and the second prioritizing sensitivity at lower resolution[15] ($\sim 7-9$ arcsec). This results in images with central rms sensitivities of $\sim 1.3-2.7$ μJy beam^{-1} in the lower resolution images and $\sim 1.2-3.6$ μJy beam^{-1} in the higher resolution images. This increases to $\sim 2.0-3.5$ μJy beam^{-1} (lower resolution) and $\sim 1.9-5.6$ μJy beam^{-1} (higher resolution) across the full area. Source finding and characterization using PyBDSF identified a total of 143 817 sources in the lower resolution images and 114 225 sources in the higher resolution images, using a multistage source finding process that more accurately characterizes the rms across the areas observed. These observations present the largest areal observations at GHz frequencies, which reach an rms sensitivity $\sim$ a few μJy beam^{-1} depth. This allows a high density of sources to be obtained, with $\gtrsim$ 6000 ($\gtrsim$ 12000) sources per deg^2 in the maximal sensitivity for the COSMOS and XMM-LSS (CDFS-DEEP) fields, respectively, whilst also allowing the effects of sample variance to be overcome.

We compare the astrometry and flux-density measurements of these catalogues to the ES MIGHTEE data of Heywood et al. (2022) for the COSMOS and XMM-LSS fields, and previous observations of Miller et al. (2013) and Franzen et al. (2015) for the CDFS-DEEP

[14]Where we combine the different simulations using the 10× minimum flux-density limit of the high flux-density simulations of Table 2.

[15]Though for the CDFS-DEEP imaging, both are of comparable sensitivity.

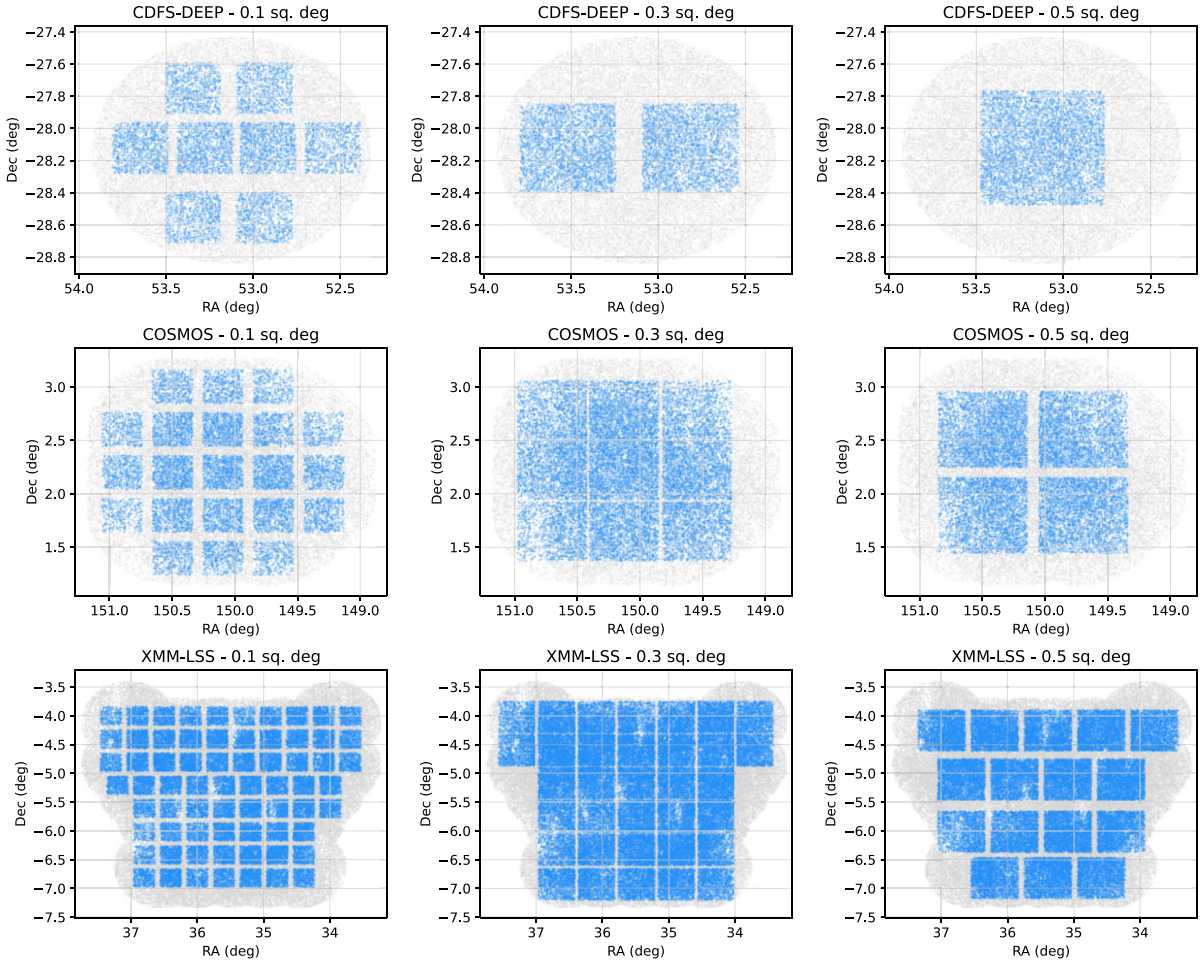


Figure 12. Regions used to define the cosmic variance within the fields. Grey points indicate the MIGHTEE radio sources across the full fields and the blue points indicate those in the smaller regions used to probe the sample variance. Shown are such regions for CDFS-DEEP (top row), COSMOS (middle row), and XMM-LSS (bottom row) for the 0.1, 0.3, and 0.5 deg^2 regions from left to right.

fields. These comparisons show that astrometric offsets are typically constrained to within a pixel of the MIGHTEE data. Comparison of source flux densities show agreement with ES observations at high SNR, with an excess seen in the flux-density measurements of the ES data compared to the observations presented in this work, for sources that were low SNR in the ES data. We attribute this to effects of the higher noise in the ES data, which exacerbate Eddington bias in the ES catalogue at faint flux densities in these confused images, together with an improvement in the PyBDSF methodology used in this work. For the CDFS field, differences in the flux densities are seen between previous CDFS field observations of Franzen et al. (2015) using the ATCA telescope compared to the sources presented in this work. Such offsets are reduced when compared to Miller et al. (2013) at high SNR. The reasons for this are unclear, though we note that there are significant differences in both the telescopes used, their configurations and the source finding algorithms used, which will all contribute to observed variations.

Next, we consider the source counts distribution within the fields, compared to the results of the ES data of Hale et al. (2023), simulations from T-RECS (Bonaldi et al. 2019, 2023) and previous deep radio observations. We calculate the raw and completeness corrected source counts using a method that better accounts for the merging of sources within the highly confused fields discussed in this work. Our corrected counts for the CDFS-DEEP field show good agreement with previous observations and simulations (e.g. Bonaldi et al. 2019; der Vlugt et al. 2021; Matthews et al. 2021a) in the faintest flux-density bins ($\lesssim$100 μJy), with the COSMOS and XMM-LSS fields exhibiting lower source counts to these works and more consistent with Smolčić et al. (2017a). These investigations demonstrate the challenges in accurately determining the source counts within such deep fields where the source density is high as well as variations which can be observed between fields. High-resolution imaging such as from the Square Kilometre Array Observatory (SKAO) and the International LOFAR Telescope (Morabito et al. 2022) will be crucial in the future to overcome the effects of confusion in increasingly deeper observations of deep fields.

Finally, we also use the large area covered by MIGHTEE DR1 to measure the sample variance between subregions in the fields and compare these to the expectations from Heywood et al. (2013). We find a typical sample variance of 10–20 per cent for the areal subregions considered for $S_{1.4} \leq 100\mu$Jy. The sample variance increases to brighter flux densities as expected, given the lower source density and the expected higher galaxy bias for these sources. Our results suggest a larger sample variance (up to a factor of $\sim 2-3$) at the faintest flux densities $\lesssim 0.1$ mJy compared to the predictions of Heywood et al. (2013), with comparable variance

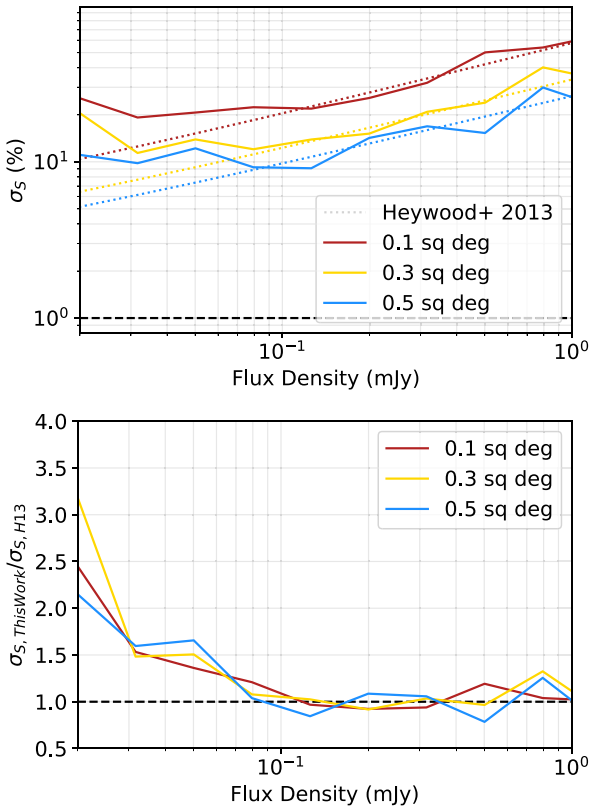


Figure 13. Comparison of the observed sample variance relative error (σ/μ) from field to field variations compared to the predictions from Heywood et al. (2013) for patches of areas: 0.1 deg^2 (red), 0.3 deg^2 (gold), and 0.5 deg^2 (blue). The upper panel shows the sample variance relative error of this work (solid lines) given as a percentage compared to the models of Heywood et al. (2013, dotted) and the lower panel presents the ratio of our work compared to Heywood et al. (2022) for each of the areas studied.

found at brighter flux densities. However, our measurements combine both uncertainties considered in their work (from Poissonian counts and source clustering) as well as the effects arising from the spread of calculated completeness correction factors found.

Alongside this paper we release images of the three fields (at both the resolutions discussed) and rms maps for the images alongside the catalogues presented in this paper. These are available from https://doi.org/10.48479/7msw-r692. Such catalogues will be elevated with future work, currently in progress, which will use the Gaussian and source catalogues to produce a host galaxy associated catalogue using a combination of visual identification (as used in other radio surveys e.g. Banfield et al. 2015; Prescott et al. 2018; Williams et al. 2019; Whittam et al. 2024) as well as statistical matching through the likelihood ratio method (see e.g. McAlpine et al. 2012; Kondapally et al. 2022; Whittam et al. 2024). Furthermore, a future data release from the MIGHTEE field will extend the observations across the CDFS field and extend to the ELAIS-S1 field and will be the final *L*-band continuum observations from the MIGHTEE survey.

ACKNOWLEDGEMENTS

The MeerKAT telescope is operated by the South African Radio Astronomy Observatory, which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation. We acknowledge the use of the *ilifu* cloud computing facility – http://www.ilifu.ac.za, a partnership between the University of Cape Town, the University of the Western Cape, Stellenbosch University, Sol Plaatje University, and the Cape Peninsula University of Technology. The *ilifu* facility is supported by contributions from the Inter-University Institute for Data Intensive Astronomy (IDIA – a partnership between the University of Cape Town, the University of Pretoria and the University of the Western Cape), the Computational Biology division at UCT and the Data Intensive Research Initiative of South Africa (DIRISA).

This research made use of ASTROPY,[16] a community-developed core PYTHON package for Astronomy (Astropy Collaboration 2013, 2018); TOPCAT (Taylor 2005, 2011); DS9 (Smithsonian Astrophysical Observatory 2000; Joye & Mandel 2003); MATPLOTLIB (Hunter 2007); NUMPY (van der Walt, Colbert & Varoquaux 2011; Harris et al. 2020); SCIPY (Virtanen et al. 2020); and TQDM (da Costa-Luis et al. 2021). This work has made use of the Cube Analysis and Rendering Tool for Astronomy (CARTA; Comrie et al. 2021). This research has made use of NASA's Astrophysics Data System. This research made use of Montage, which is funded by the National Research Foundationunder grant no. ACI-1440620, and was previously funded by the National Aeronautics and Space Administration's Earth Science Technology Office, Computation Technologies Project, under Cooperative agreement no. NCC5-626 between NASA and the California Institute of Technology. This research has made use of NASA's Astrophysics Data System.

We thank SVW, FT, and the anonymous referee for the comments which have been helpful in improving the manuscript. CLH acknowledges support from the Leverhulme Trust through an Early Career Research Fellowship. CLH, MJJ, and IHW also acknowledge support from the Oxford Hintze Centre for Astrophysical Surveys which is funded through generous support from the Hintze Family Charitable Foundation. IH and MJJ acknowledge the support of the Science and Technology Facilities Council (STFC) consolidated grant [ST/S000488/1] and [ST/W000903/1] and from a UK Research and Innovation (UKRI) Frontiers Research Grant [EP/X026639/1]. IH acknowledges support from the South African Radio Astronomy Observatory which is a facility of the National Research Foundation (NRF), an agency of the Department of Science and Innovation. IH thanks the Rhodes Centre for Radio Astronomy Techniques and Technologies (RATT) for the provision of computing facilities that were used for processing some of the MIGHTEE data. PNB is grateful for support from the UK STFCvia grant ST/V000594/1. FXA acknowledges the support from the National Natural Science Foundation of China (12303016) and the Natural Science Foundation of Jiangsu Province (BK20242115) RB acknowledges support from an STFC Ernest Rutherford Fellowship [grant no. ST/T003596/1]. IHa acknowledges support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 849169). DJBS acknowledges support from the UK STFCunder grants ST/V000624/1 and ST/Y001028/1. MV acknowledges financial support from the Inter-University Institute for Data Intensive Astronomy (IDIA), a partnership of the University of Cape Town, the University of Pretoria and the University of the Western Cape, and from the South African Department of Science and Innovation's National Research Foundationunder the ISARP RADIOMAP Joint Research Scheme (DSI-NRF grant no. 150551) and the CPRR HIPPO Project (DSI-NRF grant no. SRUG22031677). For the purpose of Open Access, the author has applied a CC BY public copyright licence to

[16] https://www.astropy.org/

any Author Accepted Manuscript (AAM) version arising from this submission.

DATA AVAILABILITY

The radio images and catalogues are accessible from: https://doi.org/10.48479/7msw-r692

REFERENCES

Alonso D., Bellini E., Hale C., Jarvis M. J., Schwarz D. J., 2021, MNRAS, 502, 876
An F. et al., 2021, MNRAS, 507, 2643
Appleton P. N. et al., 2004, ApJS, 154, 147
Asorey J., Parkinson D., 2021, MNRAS, 506, 4121
Astropy Collaboration, 2013, A&A, 558, A33
Astropy Collaboration, 2018, AJ, 156, 123
Banfield J. K. et al., 2015, MNRAS, 453, 2326
Becker R. H., White R. L., Helfand D. J., 1995, ApJ, 450, 559
Bell E. F., 2003, ApJ, 586, 794
Best P. N. et al., 2023, MNRAS, 523, 1729
Best P. N., Heckman T. M., 2012, MNRAS, 421, 1569
Blyth S. et al., 2018, in MeerKAT Science: On the Pathway to the SKA, p. 4
Bonaldi A. et al., 2021, MNRAS, 500, 3821
Bonaldi A., Bonato M., Galluzzi V., Harrison I., Massardi M., Kay S., De Zotti G., Brown M. L., 2019, MNRAS, 482, 2
Bonaldi A., Hartley P., Ronconi T., De Zotti G., Bonato M., 2023, MNRAS, 524, 993
Bondi M. et al., 2003, A&A, 403, 857
Bondi M., Ciliegi P., Schinnerer E., Smolčić V., Jahnke K., Carilli C., Zamorani G., 2008, ApJ, 681, 1129
Boucaud A., Bocchio M., Abergel A., Orieux F., Dole H., Hadj-Youcef M. A., 2016, A&A, 596, A63
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Boyce M. M. et al., 2023, PASA, 40, e027
Briggs D. S., 1995, PhD thesis, New Mexico Institute of Mining and Technology
CASA Team et al., 2022, PASP, 134, 114501
Ceraj L. et al., 2018, A&A, 620, A192
Ceraj L. et al., 2020, A&A, 642, A125
Clewley L., Jarvis M. J., 2004, MNRAS, 352, 909
Comrie A. et al., 2021, CARTA: The Cube Analysis and Rendering Tool for Astronomy. Zenodo, doi:10.5281/zenodo.3377984
Condon J. J., 1992, ARA&A, 30, 575
Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693
Croton D. J. et al., 2006, MNRAS, 365, 11
da Costa-Luis C. et al., 2021, tqdm: A fast, Extensible Progress Bar for Python and CLI, Zenodo, https://doi.org/10.5281/zenodo.5109730
De Breuck C., van Breugel W., Röttgering H. J. A., Miley G., 2000, A&AS, 143, 303
de Zotti G., Massardi M., Negrello M., Wall J., 2010, A&AR, 18, 1
Delhaize J. et al., 2017, A&A, 602, A4
Delhaize J. et al., 2021, MNRAS, 501, 3833
Delvecchio I. et al., 2021, A&A, 647, A123
Delvecchio I. et al., 2022, A&A, 668, A81
Dunlop J. S., McLure R. J., Kukula M. J., Baum S. A., O'Dea C. P., Hughes D. H., 2003, MNRAS, 340, 1095
Eddington A. S., 1913, MNRAS, 73, 359
Fabian A. C., 2012, ARA&A, 50, 455
Fernandes C. A. C. et al., 2015, MNRAS, 447, 1184
Franzen T. M. O. et al., 2015, MNRAS, 453, 4020
Garrett M. A., 2002, A&A, 384, L19
Gehrels N., 1986, ApJ, 303, 336
Gürkan G. et al., 2018, MNRAS, 475, 3010
Gürkan G., Hardcastle M. J., Jarvis M. J., 2014, MNRAS, 438, 1149
Hale C. L. et al., 2021, PASA, 38, e058
Hale C. L. et al., 2023, MNRAS, 520, 2668
Hale C. L., Jarvis M. J., Delvecchio I., Hatfield P. W., Novak M., Smolčić V., Zamorani G., 2018, MNRAS, 474, 4133
Hardcastle M. J., Evans D. A., Croston J. H., 2007, MNRAS, 376, 1849
Harris C. R. et al., 2020, Nature, 585, 357
Harrison I. et al., 2020, MNRAS, 495, 1737
Heckman T. M., Best P. N., 2014, ARA&A, 52, 589
Helfand D. J., White R. L., Becker R. H., 2015, ApJ, 801, 26
Heywood I. et al., 2022, MNRAS, 509, 2150
Heywood I., 2020, Astrophysics Source Code Library, record ascl:2009.003
Heywood I., Hale C. L., Jarvis M. J., Makhathini S., Peters J. A., Sebokolodi M. L. L., Smirnov O. M., 2020, MNRAS, 496, 3469
Heywood I., Jarvis M. J., Condon J. J., 2013, MNRAS, 432, 2625
Hopkins A. M. et al., 2015, PASA, 32, e037
Hugo B. V., Perkins S., Merry B., Mauch T., Smirnov O. M., 2022, in Ruiz J. E., Pierfedereci F., Teuben P.eds, ASP Conf. Ser., Vol. 532, Astronomical Data Analysis Software and Systems XXX. Astron. Soc. Pac., San Francisco, p. 541
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Ivison R. J. et al., 2010, A&A, 518, L31
Jarvis M. et al., 2018, in MeerKAT Science: On the Pathway to the SKA, Vol. 277, p. 6
Jarvis M. J. et al., 2010, MNRAS, 409, 92
Jarvis M. J., Rawlings S., Willott C. J., Blundell K. M., Eales S., Lacy M., 2001, MNRAS, 327, 907
Jarvis M. J., Teimourian H., Simpson C., Smith D. J. B., Rawlings S., Bonfield D., 2009, MNRAS, 398, L83
Ji Z., Giavalisco M., Kirkpatrick A., Kocevski D., Daddi E., Delvecchio I., Hatcher C., 2022, ApJ, 925, 74
Jonas J. L., 2009, IEEE Proc., 97, 1522
Jonas J., MeerKAT Team, 2018, in MeerKAT Science: On the Pathway to the SKA, Vol. 277, p. 1
Joye W. A., Mandel E., 2003, in Payne H. E., Jedrzejewski R. I., Hook R. N.eds, ASP Conf. Ser. Vol. 295, Astronomical Data Analysis Software and Systems XII, p. 489
Kalfountzou E. et al., 2017, MNRAS, 471, 28
Kenyon J. S., Smirnov O. M., Grobler T. L., Perkins S. J., 2018, MNRAS, 478, 2399
Kondapally R. et al., 2022, MNRAS, 513, 3742
Kurtzer G. M., Sochat V., Bauer M. W., 2017, PLoS ONE, 12, e0177459
Lacy M. et al., 2020, PASP, 132, 035001
Lindsay S. N., Jarvis M. J., McAlpine K., 2014, MNRAS, 440, 2322
Macfarlane C. et al., 2021, MNRAS, 506, 5888
Magliocchetti M., 2022, A&AR, 30, 6
Magliocchetti M., Popesso P., Brusa M., Salvato M., Laigle C., McCracken H. J., Ilbert O., 2017, MNRAS, 464, 3271
Matthews A. M., Condon J. J., Cotton W. D., Mauch T., 2021a, ApJ, 909, 193
Matthews A. M., Condon J. J., Cotton W. D., Mauch T., 2021b, ApJ, 914, 126
Mauch T. et al., 2020, ApJ, 888, 61
Mauch T., Murphy T., Buttery H. J., Curran J., Hunstead R. W., Piestrzynski B., Robertson J. G., Sadler E. M., 2003, MNRAS, 342, 1117
Mauch T., Sadler E. M., 2007, MNRAS, 375, 931
Mazumder A., Chakraborty A., Datta A., 2022, MNRAS, 517, 3407
McAlpine K., Smith D. J. B., Jarvis M. J., Bonfield D. G., Fleuren S., 2012, MNRAS, 423, 132
McConnell D. et al., 2020, PASA, 37, e048
McLure R. J., Kukula M. J., Dunlop J. S., Baum S. A., O'Dea C. P., Hughes D. H., 1999, MNRAS, 308, 377
McNamara B. R., Nulsen P. E. J., 2007, ARA&A, 45, 117
Miller N. A. et al., 2013, ApJS, 205, 13
Mingo B., Hardcastle M. J., Croston J. H., Dicken D., Evans D. A., Morganti R., Tadhunter C., 2014, MNRAS, 440, 269
Mohan N., Rafferty D., 2015, Astrophysics Source Code Library, record ascl:1502.007
Molnár D. C. et al., 2018, MNRAS, 475, 827
Morabito L. K. et al., 2022, A&A, 658, A1

Noordam J. E., 2004, in Oschmann J. M. J.ed., Proc. SPIE Conf. Ser., Vol. 5489, Ground-based Telescopes. p. 817
Offringa A. R. et al., 2014, MNRAS, 444, 606
Offringa A. R., de Bruyn A. G., Biehl M., Zaroubi S., Bernardi G., Pandey V. N., 2010, MNRAS, 405, 155
Panessa F., Baldi R. D., Laor A., Padovani P., Behar E., McHardy I., 2019, Nat. Astron., 3, 387
Perkins S., 2022, in Ruiz J. E., Pierfedereci F., Teuben P.eds, ASP Conf. Ser. Vol. 532, p. 337
Prescott M. et al., 2018, MNRAS, 480, 707
Radcliffe J. F., Barthel P. D., Thomson A. P., Garrett M. A., Beswick R. J., Muxlow T. W. B., 2021, A&A, 649, A27
Ramaila A. J., Smirnov O., Heywood I., 2023, Astrophysics Source Code Library, record ascl:2305.009
Rawlings S., Jarvis M. J., 2004, MNRAS, 355, L9
Rigby E. E., Argyle J., Best P. N., Rosario D., Röttgering H. J. A., 2015, A&A, 581, A96
Rigby E. E., Best P. N., Brookes M. H., Peacock J. A., Dunlop J. S., Röttgering H. J. A., Wall J. V., Ker L., 2011, MNRAS, 416, 1900
Rowe B. T. P. et al., 2015, Astron. Comput., 10, 121
Sabater J. et al., 2021, A&A, 648, A2
Saxena A. et al., 2018, MNRAS, 475, 5041
Saxena A. et al., 2019, MNRAS, 489, 5053
Schinnerer E. et al., 2010, ApJS, 188, 384
Shimwell T. W. et al., 2019, A&A, 622, A1
Shimwell T. W. et al., 2022, A&A, 659, A1
Šlaus B. et al., 2020, A&A, 638, A46
Smirnov O. M., 2011, A&A, 527, A108
Smirnov O. M., Tasse C., 2015, MNRAS, 449, 2668
Smith D. J. B. et al., 2021, A&A, 648, A6
Smithsonian Astrophysical Observatory, 2000, Astrophysics Source Code Library, record ascl:0003.002
Smolčić V. et al., 2017a, A&A, 602, A1
Smolčić V. et al., 2017b, A&A, 602, A2
Tasse C. et al., 2018, A&A, 611, A87
Tasse C. et al., 2023, Astrophysics Source Code Library, record ascl:2305.008
Tasse C., 2023, Astrophysics Source Code Library, record ascl:2305.005
Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R.eds, ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV, p. 29
Taylor M., 2011, Astrophysics Source Code Library, record ascl:1101.010
van der Vlugt D. et al., 2021, ApJ, 907, 5
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Virtanen P. et al., 2020, Nat. Methods, 17, 261
White R. L., Becker R. H., Helfand D. J., Gregg M. D., 1997, ApJ, 475, 479
White S. V., Jarvis M. J., Häußler B., Maddox N., 2015, MNRAS, 448, 2665
White S. V., Jarvis M. J., Kalfountzou E., Hardcastle M. J., Verma A., Cao Orjales J. M., Stevens J., 2017, MNRAS, 468, 217
Whittam I. H. et al., 2022, MNRAS, 516, 245
Whittam I. H. et al., 2024, MNRAS, 527, 3231
Whittam I. H., Prescott M., McAlpine K., Jarvis M. J., Heywood I., 2018, MNRAS, 480, 358
Williams W. L. et al., 2019, A&A, 622, A2
Willott C. J., Rawlings S., Blundell K. M., Lacy M., Eales S. A., 2001, MNRAS, 322, 536
Wilman R. J. et al., 2008, MNRAS, 388, 1335
Yuan Z., Wang J., Zhou M., Qin L., Mao J., 2017, ApJ, 846, 78
Yun M. S., Reddy N. A., Condon J. J., 2001, ApJ, 554, 803

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: LIST OF MEERKAT OBSERVATIONS

Tables A1-A3 provide details of the 709.2 h of MeerKAT observations that were used to produce the data products presented here, for each of the three target fields.

Table A1. MeerKAT observations that were used to produce the COSMOS mosaics. Columns included are the date of observation, observation ID, field, central RA and Dec. of the pointing, time in hours of the full observation track and on-source time, number of spectral channels, number of antennas used in the observation, and finally, the primary and secondary calibrator sources.

Date	ID	Field	RA (J2000)	Dec. (J2000)	Track (h)	On-source (h)	$N_{\rm chan}$	$N_{\rm ant}$	Primary	Secondary
2018-04-19	1524147354	COSMOS	$10^{\rm h}00^{\rm m}29^{\rm s}$	$02^\circ12'21\farcs0$	8.65	6.10	4096	64	J0408-6545	3C237
2018-05-06	1525613583	COSMOS	$10^{\rm h}00^{\rm m}29^{\rm s}$	$02^\circ12'21\farcs0$	8.39	5.10	4096	62	J0408-6545	3C237
2019-07-16	1563267356	COSMOS_1	$09^{\rm h}59^{\rm m}46^{\rm s}$	$02^\circ01'44\farcs6$	7.00	6.33	4096	59	J0408-6545	J1008+0740
2019-07-27	1564215117	COSMOS_2	$09^{\rm h}59^{\rm m}46^{\rm s}$	$02^\circ22'57\farcs4$	7.95	6.98	4096	61	J0408-6545	J1008+0740
2019-07-28	1564301832	COSMOS_3	$10^{\rm h}01^{\rm m}11^{\rm s}$	$02^\circ01'44\farcs6$	7.96	6.97	4096	60	J0408-6545	J1008+0740
2019-08-16	1565939836	COSMOS_4	$10^{\rm h}01^{\rm m}11^{\rm s}$	$02^\circ22'57\farcs4$	7.99	6.97	4096	58	J0408-6545	J1008+0740
2019-08-23	1566542621	COSMOS_1	$09^{\rm h}59^{\rm m}46^{\rm s}$	$02^\circ01'44\farcs6$	7.97	6.98	4096	61	J0408-6545	J1008+0740
2020-03-28	1585413022	COSMOS_5	$09^{\rm h}59^{\rm m}04^{\rm s}$	$02^\circ12'21\farcs0$	8.00	6.25	32768	59	J0408-6545	J1008+0740
2020-03-29	1585498873	COSMOS_6	$10^{\rm h}01^{\rm m}54^{\rm s}$	$02^\circ12'21\farcs0$	8.00	6.25	32768	59	J0408-6545	J1008+0740
2020-03-31	1585671638	COSMOS_7	$10^{\rm h}00^{\rm m}29^{\rm s}$	$01^\circ51'08\farcs2$	8.00	6.25	32768	60	J0408-6545	J1008+0740
2020-04-02	1585844155	COSMOS_8	$10^{\rm h}00^{\rm m}29^{\rm s}$	$02^\circ33'33\farcs8$	8.00	6.25	32768	60	J0408-6545	J1008+0740
2020-04-30	1585928757	COSMOS_9	$10^{\rm h}01^{\rm m}54^{\rm s}$	$02^\circ33'33\farcs8$	8.00	6.25	32768	60	J0408-6545	J1008+0740
2020-04-04	1586016787	COSMOS_10	$09^{\rm h}59^{\rm m}04^{\rm s}$	$02^\circ33'33\farcs8$	8.03	6.25	32768	60	J0408-6545	J1008+0740
2020-04-06	1586188138	COSMOS_11	$09^{\rm h}58^{\rm m}21^{\rm s}$	$02^\circ22'57\farcs4$	8.00	6.25	32768	59	J0408-6545	J1008+0740
2020-04-07	1586274966	COSMOS_12	$09^{\rm h}58^{\rm m}21^{\rm s}$	$02^\circ01'44\farcs6$	8.00	6.25	32768	60	J0408-6545	J1008+0740
2020-04-12	1586705155	COSMOS_13	$09^{\rm h}59^{\rm m}04^{\rm s}$	$01^\circ51'08\farcs2$	8.00	6.25	32768	59	J0408-6545	J1008+0740
2020-04-13	1586791316	COSMOS_14	$10^{\rm h}01^{\rm m}53^{\rm s}$	$01^\circ51'08\farcs2$	8.00	6.25	32768	60	J0408-6545	J1008+0740
2020-04-26	1587911796	COSMOS	$10^{\rm h}00^{\rm m}29^{\rm s}$	$02^\circ12'21\farcs0$	7.98	6.25	32768	59	J0408-6545	J1008+0740
2021-04-07	1617809470	COSMOS_1	$09^{\rm h}59^{\rm m}46^{\rm s}$	$02^\circ01'44\farcs6$	8.00	6.25	32768	60	J0408-6545	J1008+0740
2021-05-02	1619963656	COSMOS_2	$09^{\rm h}59^{\rm m}46^{\rm s}$	$02^\circ22'57\farcs4$	7.96	6.25	32768	62	J0408-6545	J1008+0740
2021-05-15	1621083675	COSMOS_4	$10^{\rm h}01^{\rm m}11^{\rm s}$	$02^\circ22'57\farcs4$	7.97	6.25	32768	61	J0408-6545	J1008+0740
2021-05-30	1622376680	COSMOS_3	$10^{\rm h}01^{\rm m}11^{\rm s}$	$02^\circ01'44\farcs6$	8.00	6.70	32768	61	J0408-6545	J1008+0740

Table A2. MeerKAT observations that were used to produce the CDFS-DEEP image. For all CDFS observations, the primary calibrator was PKS B0408-65, and the secondary was J0240-2309. The pointing centre for all of these observations was J2000 $03^{\rm h}32^{\rm m}30.4^{\rm s}-28^\circ07^{\rm m}57\farcs0$. These observations all have 32 768 channels and the column descriptions are given in Table A1.

Date	ID	Track (h)	On-source (h)	$N_{\rm ant}$
2019-12-12	1576162858	9.91	7.26	59
2020-01-03	1578058860	9.92	7.26	61
2020-01-24	1579878660	7.47	5.60	61
2020-01-26	1580039158	9.90	7.25	61
2020-02-08	1581162358	9.85	7.25	62
2020-02-09	1581248760	10.18	7.59	62
2020-02-22	1582371217	9.81	7.26	58
2020-04-25	1587795059	9.94	7.24	59
2020-09-28	1601325069	9.08	6.61	60
2020-09-29	1601409818	9.09	6.29	58
2020-10-03	1601756163	9.08	6.28	61
2020-10-05	1601928962	9.09	6.28	59
2020-10-07	1602098167	9.06	6.27	58
2020-10-08	1602183844	9.06	6.27	59
2020-10-09	1602270065	9.06	6.27	60
2020-10-14	1602704665	9.06	6.60	59
2020-10-15	1602792067	9.08	6.27	60
2020-10-16	1602873187	9.06	6.60	60
2020-10-17	1602964865	9.10	6.26	59

Table A3. MeerKAT observations that were used to produce the XMM-LSS mosaics. Column descriptions are given in Table A1.

Date	ID	Field	RA (J2000)	Dec. (J2000)	Track (h)	On-source (h)	$N_{\rm chan}$	$N_{\rm ant}$	Primary	Secondary
2018-10-06	1538856059	XMMLSS_12	$02^{\rm h}17^{\rm m}51^{\rm s}$	$-04°49'59\farcs0$	8.02	6.20	4096	59	J1939-6342	J0201-1132
2018-10-07	1538942495	XMMLSS_13	$02^{\rm h}20^{\rm m}42^{\rm s}$	$-04°49'59\farcs0$	8.07	6.22	4096	59	J1939-6342	J0201-1132
2018-10-08	1539028868	XMMLSS_14	$02^{\rm h}23^{\rm m}22^{\rm s}$	$-04°49'59\farcs0$	8.03	6.19	4096	60	J1939-6342	J0201-1132
2018-10-11	1539286252	XMMLSS_12	$02^{\rm h}17^{\rm m}51^{\rm s}$	$-04°49'59\farcs0$	8.05	6.23	4096	63	J1939-6342	J0201-1132
2018-10-12	1539372679	XMMLSS_13	$02^{\rm h}20^{\rm m}42^{\rm s}$	$-04°49'59\farcs0$	8.03	5.92	4096	62	J1939-6342	J0201-1132
2018-10-13	1539460932	XMMLSS_14	$02^{\rm h}23^{\rm m}22^{\rm s}$	$-04°49'59\farcs0$	8.00	6.24	4096	62	J1939-6342	J0201-1132
2019-07-27	1564271932	XMMLSS_15	$02^{\rm h}26^{\rm m}22^{\rm s}$	$-04°37'58\farcs8$	8.00	7.00	4096	59	J1939-6342	J0201-1132
2019-08-02	1564788958	XMMLSS_12p5	$02^{\rm h}19^{\rm m}16^{\rm s}$	$-04°49'58\farcs8$	7.99	6.97	4096	62	J0408-6545	J0201-1132
2019-08-03	1564874467	XMMLSS_13p5	$02^{\rm h}22^{\rm m}06^{\rm s}$	$-04°49'58\farcs8$	8.07	7.00	4096	62	J1939-6342	J0201-1132
2020-01-06	1578317762	XMMLSS_12	$02^{\rm h}17^{\rm m}51^{\rm s}$	$-04°49'59\farcs0$	8.00	6.50	32768	62	J1939-6342	J0201-1132
2020-08-10	1597099565	XMMLSS_6_8	$02^{\rm h}15^{\rm m}47^{\rm s}$	$-04°49'58\farcs8$	8.00	6.85	32768	60	J1939-6342	J0201-1132
2020-08-13	1597359662	XMMLSS_8_8	$02^{\rm h}19^{\rm m}54^{\rm s}$	$-04°49'58\farcs8$	8.00	6.83	32768	60	J1939-6342	J0201-1132
2020-08-14	1597445461	XMMLSS_9_8	$02^{\rm h}21^{\rm m}58^{\rm s}$	$-04°49'58\farcs8$	8.00	6.83	32768	61	J1939-6342	J0201-1132
2020-08-15	1597534262	XMMLSS_10_8	$02^{\rm h}24^{\rm m}02^{\rm s}$	$-04°49'58\farcs8$	8.00	6.83	32768	62	J1939-6342	J0201-1132
2020-08-16	1597617063	XMMLSS_11_8	$02^{\rm h}26^{\rm m}06^{\rm s}$	$-04°49'58\farcs8$	8.00	6.83	32768	62	J1939-6342	J0201-1132
2020-08-17	1597703462	XMMLSS_12_8	$02^{\rm h}28^{\rm m}09^{\rm s}$	$-04°49'58\farcs8$	8.00	6.83	32768	61	J1939-6342	J0201-1132
2020-08-24	1598306465	XMMLSS_6_9	$02^{\rm h}16^{\rm m}48^{\rm s}$	$-04°23'16\farcs8$	8.00	6.85	32768	58	J1939-6342	J0201-1132
2020-08-27	1598564761	XMMLSS_7_9	$02^{\rm h}18^{\rm m}52^{\rm s}$	$-04°23'16\farcs8$	8.00	6.83	32768	59	J1939-6342	J0201-1132
2020-08-30	1598823546	XMMLSS_8_9	$02^{\rm h}20^{\rm m}56^{\rm s}$	$-04°23'16\farcs8$	7.99	6.83	32768	59	J1939-6342	J0201-1132
2020-09-03	1599168664	XMMLSS_9_9	$02^{\rm h}22^{\rm m}59^{\rm s}$	$-04°23'16\farcs8$	7.99	6.83	32768	58	J1939-6342	J0201-1132
2020-09-10	1599770763	XMMLSS_10_9	$02^{\rm h}25^{\rm m}03^{\rm s}$	$-04°23'16\farcs8$	7.99	6.83	32768	58	J1939-6342	J0201-1132
2020-09-11	1599858231	XMMLSS_11_9	$02^{\rm h}27^{\rm m}07^{\rm s}$	$-04°23'16\farcs8$	7.99	6.83	32768	58	J1939-6342	J0201-1132
2020-09-21	1600722156	XMMLSS_6_7	$02^{\rm h}16^{\rm m}49^{\rm s}$	$-05°16'40\farcs8$	8.05	6.85	32768	60	J1939-6342	J0201-1132
2020-09-23	1600893068	XMMLSS_7_7	$02^{\rm h}18^{\rm m}53^{\rm s}$	$-05°16'40\farcs8$	7.97	6.78	32768	59	J1939-6342	J0201-1132
2020-09-25	1601066490	XMMLSS_8_7	$02^{\rm h}20^{\rm m}57^{\rm s}$	$-05°16'40\farcs8$	7.97	6.78	32768	60	J1939-6342	J0201-1132
2020-10-01	1601583156	XMMLSS_9_7	$02^{\rm h}23^{\rm m}01^{\rm s}$	$-05°16'40\farcs8$	7.98	6.78	32768	59	J1939-6342	J0201-1132
2020-10-02	1601667467	XMMLSS_10_7	$02^{\rm h}25^{\rm m}04^{\rm s}$	$-05°16'40\farcs8$	7.97	6.78	32768	61	J1939-6342	J0201-1132
2020-10-18	1603049464	XMMLSS_11_7	$02^{\rm h}27^{\rm m}08^{\rm s}$	$-05°16'40\farcs8$	8.00	6.78	32768	59	J1939-6342	J0201-1132
2021-03-20	1616233334	XMMLSS_7_6	$02^{\rm h}17^{\rm m}52^{\rm s}$	$-05°43'22\farcs8$	3.99	3.12	32768	63	J0408-6545	J0201-1132
2021-04-17	1618640777	XMMLSS_7_6	$02^{\rm h}17^{\rm m}52^{\rm s}$	$-05°43'22\farcs8$	8.00	6.74	32768	63	J0408-6545	J0201-1132
2021-04-18	1618726873	XMMLSS_8_6	$02^{\rm h}19^{\rm m}56^{\rm s}$	$-05°43'22\farcs8$	8.00	6.74	32768	64	J0408-6545	J0201-1132
2021-04-23	1619161272	XMMLSS_9_6	$02^{\rm h}21^{\rm m}59^{\rm s}$	$-05°43'22\farcs8$	8.00	6.74	32768	60	J0408-6545	J0201-1132
2021-04-26	1619245873	XMMLSS_10_6	$02^{\rm h}24^{\rm m}03^{\rm s}$	$-05°43'22\farcs8$	8.00	6.74	32768	63	J0408-6545	J0201-1132
2021-04-25	1619330180	XMMLSS_11_6	$02^{\rm h}26^{\rm m}07^{\rm s}$	$-05°43'22\farcs8$	8.00	6.74	32768	63	J0408-6545	J0201-1132
2021-04-26	1619416874	XMMLSS_7_5	$02^{\rm h}18^{\rm m}54^{\rm s}$	$-06°10'04\farcs8$	8.35	6.74	32768	61	J0408-6545	J0201-1132
2021-05-02	1619933473	XMMLSS_10_5	$02^{\rm h}25^{\rm m}06^{\rm s}$	$-06°10'04\farcs8$	8.03	6.74	32768	62	J0408-6545	J0201-1132
2021-05-04	1620109872	XMMLSS_9_5	$02^{\rm h}23^{\rm m}02^{\rm s}$	$-06°10'04\farcs8$	8.03	6.74	32768	59	J0408-6545	J0201-1132
2021-05-09	1620536533	XMMLSS_8_5	$02^{\rm h}20^{\rm m}58^{\rm s}$	$-06°10'04\farcs8$	8.00	6.70	32768	62	J0408-6545	J0201-1132
2021-05-14	1620967872	XMMLSS_8_4	$02^{\rm h}19^{\rm m}57^{\rm s}$	$-06°36'46\farcs8$	8.00	6.70	32768	59	J0408-6545	J0201-1132
2021-05-22	1621656680	XMMLSS_7_4	$02^{\rm h}17^{\rm m}53^{\rm s}$	$-06°36'46\farcs8$	8.00	6.70	32768	63	J0408-6545	J0201-1132
2021-05-23	1621742776	XMMLSS_10_4	$02^{\rm h}24^{\rm m}05^{\rm s}$	$-06°36'46\farcs8$	8.00	6.70	32768	62	J0408-6545	J0201-1132
2021-06-05	1622863620	XMMLSS_11_4	$02^{\rm h}26^{\rm m}09^{\rm s}$	$-06°36'46\farcs8$	8.00	6.70	32768	61	J0408-6545	J0201-1132
2021-06-06	1622949019	XMMLSS_6_9p5	$02^{\rm h}15^{\rm m}47^{\rm s}$	$-04°05'60\farcs0$	8.00	6.70	32768	61	J0408-6545	J0201-1132
2021-06-12	1623469482	XMMLSS_9_4	$02^{\rm h}22^{\rm m}01^{\rm s}$	$-06°36'46\farcs8$	8.00	6.70	32768	63	J0408-6545	J0201-1132
2021-06-27	1624760792	XMMLSS_12_9p5	$02^{\rm h}28^{\rm m}09^{\rm s}$	$-04°05'60\farcs0$	8.00	6.70	32768	58	J0408-6545	J0201-1132

APPENDIX B: EXTRACT FROM CATALOGUE

Table B1 shows the first five entries from the source catalogue of the low-resolution imaging of the COSMOS field.

Table B1. An example source catalogue, showing the first five lines for the low-resolution (8.9 arcsec) COSMOS catalogue. The columns for each source included from PyBDSF are: the ID of the source (Source_id); the ID of the island of emission associated with the source (Isl_id); the RA and Dec. of the source (in J2000 co-ordinates), the integrated flux density (Total_flux), and peak flux density (Peak_flux); source size information of the major and minor axes and position angle (Maj, Min, and PA) and deconvolved source sizes (indicated by DC_); rms across the island (Isl_rms) and a code to describe the type of source (S_Code where S = Single, M = Multiple, and C = Complex). All errors are indicated by columns which begin with E_. Further details of the PyBDSF columns can be found at https://pybdsf.readthedocs.io. We additionally include a column with the effective frequency of the source, in MHz, at the source position (Eff_freq) and a column which lists the number of Gaussian components that the source is comprised of (NGaus).

Source_id	Isl_id	RA (°)	E_RA (°)	DEC (°)	E_DEC (°)	Total_flux (Jy)	E_Total_flux (Jy)	Peak_flux (Jy beam^{-1})
0	0	151.173585	0.000063	2.216888	0.000053	0.0001580	0.0000135	0.0001373
1	1	151.172523	0.000092	1.854424	0.000112	0.0000810	0.0000121	0.0000706
2	3	151.173070	0.000222	2.541834	0.000289	0.0000756	0.0000183	0.0000452
3	4	151.171141	0.000083	1.845267	0.000160	0.0001294	0.000016	0.0000747
4	5	151.169030	0.000294	1.774813	0.000245	0.0000428	0.0000138	0.0000330

E_Peak_flux (Jy beam^{-1})	Maj (°)	E_Maj (°)	Min (°)	E_Min (°)	PA (°)	E_PA (°)	DC_Maj (°)	E_DC_Maj (°)
0.0000071	0.002774	0.000149	0.002534	0.000125	87.855644	24.131809	0.001260	0.000149
0.0000064	0.002795	0.000265	0.002507	0.000215	173.143515	35.577114	0.001306	0.000265
0.0000073	0.004090	0.000790	0.002500	0.000333	34.249398	17.301097	0.003258	0.000790
0.0000062	0.003945	0.000379	0.002682	0.000194	174.329007	10.942307	0.003075	0.000379
0.0000065	0.003461	0.000820	0.002285	0.000377	126.972961	23.713182	0.002423	0.000820

DC_Min (°)	E_DC_Min (°)	DC_PA (°)	E_DC_PA (°)	Isl_rms (Jy beam^{-1})	S_Code	Eff_freq (MHz)	NGaus
0.000558	0.000125	87.855644	24.131809	0.0000069	S	1225.10	1
0.000420	0.000215	173.143515	35.577114	0.0000062	S	1222.32	1
0.000380	0.000333	34.249398	17.301097	0.0000070	S	1223.50	1
0.001043	0.000194	174.329007	10.942307	0.0000059	S	1222.16	1
0.000000	0.000377	126.972961	23.713182	0.0000064	S	1236.36	1

APPENDIX C: SIGNAL-TO-NOISE ENVELOPES APPLIED

To identify unresolved sources we fit an upper envelope of the form:

$$\frac{S_I}{S_P} = A + B \times \mathrm{SNR}^{-C}, \tag{C1}$$

where S_I is the integrated flux density, S_P is the peak flux density, SNR is the peak signal-to-noise ratio, and A, B, and C are constants which are fit for as in the method used in Hale et al. (2021). The parameters used for the catalogue validation is given in Table C1.

Table C1. Parameters A, B, and C used to define the upper SNR envelopes as in equation (C1), to identify unresolved sources, as used in Section 4.

Field	Image resolution (arcsec)	A	B	C
CDFS-DEEP	5.5	1.02	0.80	0.75
CDFS-DEEP	7.3	1.04	0.60	0.55
COSMOS	5.2	1.06	1.10	0.75
COSMOS	8.9	1.01	0.70	0.60
XMM-LSS	5.0	1.06	1.10	0.70
XMM-LSS	8.9	1.00	0.60	0.55

APPENDIX D: TABLE OF SOURCE COUNTS

Table D1 presents the 1.4 GHz Euclidean normalized source counts in each of the three fields observed.

Table D1. Details of the 1.4 GHz Euclidean normalized source counts presented in Section 5 and Fig. 11. Given are the flux-density ranges for the source counts, the median flux density of sources within the bin considered, the raw number counts and source counts as well as the corrected source counts for each of the three fields studied: CDFS-DEEP, COSMOS, and XMM-LSS both through (i) combining all completeness simulations of the high flux-density simulations at $\geq 10\times$ the minimum input flux density and (ii) combining completeness simulations together for >1 mJy. This is released as Supplementary data alongside the paper, in which the terms 'nerr' and 'perr' are used to indicate the negative and positive errors, respectively.

	CDFS-DEEP			COSMOS			XMM-LSS			Completeness (high flux $\geq 10 \times S_{\rm min}$)			Completeness (combined >1 mJy)		
Flux density range (μJy)	Median flux density (μJy)	Raw number counts	Raw source counts (Jy$^{1.5}$sr^{-1})	Median flux density (μJy)	Raw number counts	Raw source counts (Jy$^{1.5}$sr^{-1})	Median flux density (μJy)	Raw number counts	Raw source counts (Jy$^{1.5}$sr^{-1})	Corrected source counts CDFS-DEEP (Jy$^{1.5}$sr^{-1})	Corrected source counts COSMOS (Jy$^{1.5}$sr^{-1})	Corrected source counts XMM-LSS (Jy$^{1.5}$sr^{-1})	Corrected source counts CDFS-DEEP (Jy$^{1.5}$sr^{-1})	Corrected source counts COSMOS (Jy$^{1.5}$sr^{-1})	Corrected source counts XMM-LSS (Jy$^{1.5}$sr^{-1})
10–15	13	3058	$0.62^{-0.01}_{+0.01}$	–	–	–	–	–	–	$1.94^{-0.08}_{+0.08}$	–	–	$1.94^{-0.08}_{+0.08}$	–	–
15–25	19	4122	$1.68^{-0.03}_{+0.03}$	–	–	–	–	–	–	$3.07^{-0.11}_{+0.11}$	–	–	$3.07^{-0.11}_{+0.11}$	–	–
25–39	31	3468	$2.82^{-0.05}_{+0.05}$	31	6068	$1.81^{-0.02}_{+0.02}$	31	21074	$1.84^{-0.01}_{+0.01}$	$3.07^{-0.10}_{+0.10}$	$2.94^{-0.07}_{+0.07}$	$2.76^{-0.03}_{+0.03}$	$3.07^{-0.10}_{+0.10}$	$2.94^{-0.07}_{+0.07}$	$2.76^{-0.03}_{+0.03}$
39–63	48	2476	$4.01^{-0.08}_{+0.08}$	49	5152	$3.07^{-0.04}_{+0.04}$	49	18141	$3.16^{-0.02}_{+0.02}$	$3.75^{-0.13}_{+0.13}$	$3.10^{-0.07}_{+0.07}$	$3.17^{-0.04}_{+0.04}$	$3.75^{-0.13}_{+0.13}$	$3.10^{-0.07}_{+0.07}$	$3.17^{-0.04}_{+0.04}$
63–100	77	1493	$4.83^{-0.12}_{+0.13}$	76	3564	$4.24^{-0.07}_{+0.07}$	77	12458	$4.34^{-0.04}_{+0.04}$	$4.11^{-0.16}_{+0.16}$	$3.58^{-0.10}_{+0.10}$	$3.60^{-0.05}_{+0.05}$	$4.11^{-0.16}_{+0.16}$	$3.58^{-0.10}_{+0.10}$	$3.60^{-0.05}_{+0.05}$
100–158	118	772	$4.98^{-0.18}_{+0.19}$	120	2080	$4.93^{-0.11}_{+0.11}$	121	7041	$4.89^{-0.06}_{+0.06}$	$4.24^{-0.21}_{+0.22}$	$4.12^{-0.15}_{+0.15}$	$4.02^{-0.06}_{+0.06}$	$3.91^{-0.20}_{+0.21}$	$3.95^{-0.14}_{+0.14}$	$3.83^{-0.06}_{+0.06}$
158–251	192	379	$4.88^{-0.25}_{+0.26}$	191	1072	$5.07^{-0.15}_{+0.16}$	193	3667	$5.08^{-0.08}_{+0.09}$	$4.09^{-0.25}_{+0.26}$	$4.25^{-0.19}_{+0.19}$	$4.14^{-0.09}_{+0.09}$	$3.58^{-0.23}_{+0.24}$	$4.16^{-0.18}_{+0.18}$	$3.98^{-0.09}_{+0.09}$
251–398	307	203	$5.21^{-0.37}_{+0.39}$	306	583	$5.50^{-0.23}_{+0.24}$	307	1961	$5.42^{-0.12}_{+0.13}$	$4.75^{-0.43}_{+0.44}$	$4.85^{-0.26}_{+0.27}$	$4.82^{-0.14}_{+0.14}$	$4.28^{-0.39}_{+0.41}$	$4.62^{-0.27}_{+0.28}$	$4.60^{-0.14}_{+0.14}$
398–630	480	105	$5.38^{-0.52}_{+0.58}$	488	316	$5.95^{-0.33}_{+0.35}$	480	1061	$5.85^{-0.18}_{+0.19}$	$5.09^{-0.66}_{+0.70}$	$5.24^{-0.39}_{+0.40}$	$5.20^{-0.21}_{+0.21}$	$5.51^{-0.75}_{+0.79}$	$4.78^{-0.38}_{+0.39}$	$5.15^{-0.22}_{+0.22}$
630–1000	738	52	$5.32^{-0.73}_{+0.84}$	746	176	$6.61^{-0.50}_{+0.54}$	779	623	$6.86^{-0.27}_{+0.29}$	$5.48^{-0.99}_{+1.08}$	$6.39^{-0.64}_{+0.67}$	$6.78^{-0.39}_{+0.40}$	$7.53^{-1.66}_{+1.76}$	$6.25^{-0.60}_{+0.63}$	$6.66^{-0.37}_{+0.38}$
1000–1584	1260	42	$8.57^{-1.32}_{+1.54}$	1245	131	$9.82^{-0.86}_{+0.94}$	1221	415	$9.11^{-0.45}_{+0.47}$	$9.06^{-1.83}_{+2.02}$	$9.87^{-1.17}_{+1.23}$	$9.20^{-0.61}_{+0.63}$	$9.16^{-1.96}_{+2.13}$	$9.81^{-1.18}_{+1.24}$	$9.30^{-0.67}_{+0.68}$
1584–2511	1973	39	$15.88^{-2.53}_{+2.97}$	1989	92	$13.76^{-1.43}_{+1.59}$	1965	294	$12.88^{-0.75}_{+0.80}$	$17.65^{-3.86}_{+4.23}$	$14.35^{-2.14}_{+2.26}$	$13.26^{-1.06}_{+1.09}$	$18.19^{-4.43}_{+4.77}$	$14.35^{-2.20}_{+2.32}$	$13.20^{-1.08}_{+1.11}$
2511–3981	3264	26	$21.12^{-4.11}_{+5.00}$	3044	61	$18.21^{-2.32}_{+2.64}$	3100	204	$17.83^{-1.25}_{+1.34}$	$23.40^{-6.19}_{+6.95}$	$18.71^{-3.21}_{+3.46}$	$18.42^{-1.74}_{+1.81}$	$25.93^{-11.96}_{+12.46}$	$18.63^{-3.35}_{+3.59}$	$18.20^{-1.77}_{+1.84}$
3981–6309	4756	20	$32.41^{-7.18}_{+8.98}$	4983	69	$41.10^{-4.93}_{+5.57}$	5018	175	$30.52^{-2.30}_{+2.49}$	$36.23^{-10.54}_{+12.15}$	$46.83^{-8.38}_{+8.88}$	$33.68^{-3.45}_{+3.60}$	$37.26^{-11.79}_{+13.33}$	$51.25^{-15.06}_{+15.40}$	$34.87^{-4.42}_{+4.55}$
6309–10000	8358	14	$45.27^{-11.94}_{+15.60}$	7455	39	$46.35^{-7.39}_{+8.67}$	8126	124	$43.15^{-3.87}_{+4.23}$	$55.45^{-19.04}_{+22.67}$	$56.68^{-12.51}_{+13.68}$	$50.37^{-5.83}_{+6.17}$	$76.86^{-100.85}_{+102.28}$	$75.51^{-64.98}_{+65.40}$	$50.34^{-6.10}_{+6.42}$

This paper has been typeset from a TEX/LATEX file prepared by the author.