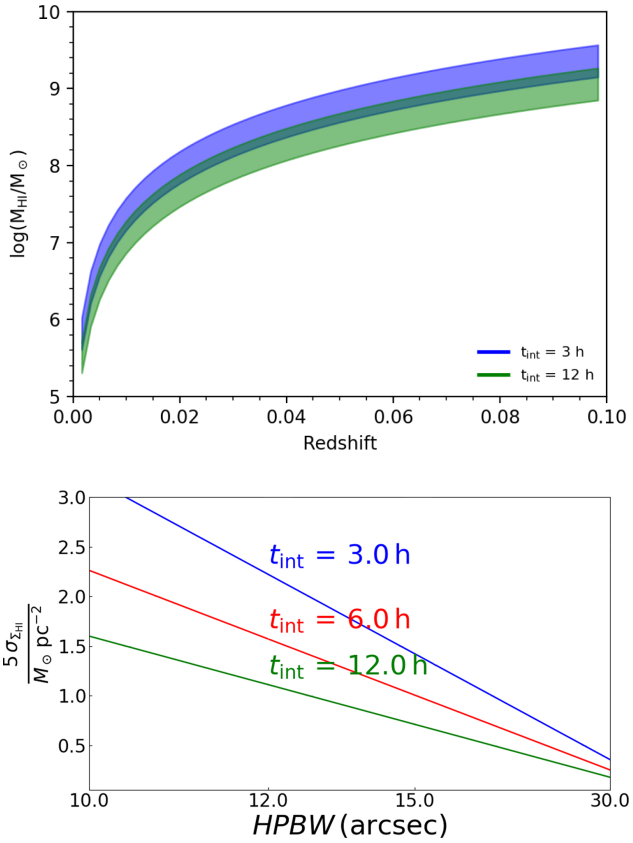

Fig. 2. H I mass sensitivity for the MGCLS. *Top*: logarithm of the total H I mass sensitivity as a function of redshift for integration times between 3 h and 12 h (not all MGCLS datasets have the same usable on-source time for H I analysis). The shaded areas indicate the H I mass limits of the MGCLS data assuming a 5σ detection for a galaxy with line width ranging from 44 km s^{-1} to 300 km s^{-1}. *Bottom*: column-density sensitivity versus angular resolution at $z = 0.03$ for different integration times present in the MGCLS. The horizontal axis scale is proportional to HPBW^{-2}.

velocity resolution of 44.1 km s^{-1} (at $z = 0$). The survey is therefore suitable to approximately resolve the velocity structure of galaxies with a rotational amplitude of $\gtrsim$100 km s^{-1} (depending on their inclination), which can be seen as a rough threshold dividing dwarf galaxies from more massive objects (Lelli et al. 2014).

The top panel of Fig. 2 shows the H I mass sensitivity as a function of redshift for an integration time of 3 and 12 h (the range of usable on-source time for H I analysis in the MGCLS, with most clusters having 6−10 h on-source), demonstrating that the MGCLS observations are able to detect galaxies below the 'knee' of the H I mass function ($\log(M^*_{\rm H\,I}/M_\odot) = 9.94$, Jones et al. 2018) out to $z \lesssim 0.1$. Moreover, the angular resolution provided by the survey, shown in the bottom panel of Fig. 2, enables one to resolve the structure of the H I in galaxies with a resolution between 10″ and 30″, corresponding to a spatial resolution of ∼4−20 kpc at redshifts of $0.02 < z < 0.1$. This means that for clusters in this range, larger galaxies will be spatially resolved and we can conduct studies of resolved H I galaxies while simultaneously probing extended extragalactic structures in the wide field of view – a key discriminating feature with respect to single-dish H I surveys of rich environments. Panel D in Fig. 1 shows an example of a resolved H I detection, covering two consecutive channels, in the Abell 194 field. Examples of H I science from the MGCLS datasets are given in Sect. 9.

4.4. Data quality issues

There are a number of issues with data quality in the DR1 release that affect the visibilities and images. The images have been corrected for a number of these effects, as described below. However, they still affect the visibilities, and impose limitations on the accuracy of the images. Users of the data products should take these into consideration when scientific analyses are performed.

4.4.1. Dynamic range

Fields with very strong sources ($I >$ few 100 s mJy beam^{-1}) are typically limited by residual artefacts from the brighter sources. This is especially true if there are several widely separated bright sources, as the self-calibration cannot correct direction-dependent effects (DDEs). The DDEs that are thought to dominate are asymmetries in the antenna pattern, pointing errors, and ionospheric refraction.

4.4.2. Flux density and spectral index

Uncertainties in the primary beam pattern affect both the derived brightnesses and spectral indices. Near the centre of each field, the total array pattern is very close to that of the individual antennas (Mauch et al. 2020). However, application of this individual antenna pattern to spectra of sources in the outer parts of cluster fields produces non-physical results. This is expected because the array power pattern is broadened by pointing errors in the individual antennas. Derived brightnesses, and especially spectral indices, are therefore not reliable beyond a radius of 36′.

We compared the flux densities of the MGCLS compact sources with those from other radio surveys, as discussed in Sect. 5.2.1, and found them to agree to within about 6%, with the MeerKAT flux densities being, on average, 6% ± 4% lower than those of the other two catalogues. No corrections to the flux densities have been made in either the catalogues or the maps.

While MeerKAT's very wide fractional bandwidth improves the sensitivity and gives us the ability to derive in-band spectral indices, it also creates uncertainties regarding the effective central frequency at each pixel. Simple averaging in frequency space would result in effective reference frequencies that depend on the spectra of the sources in question and their position in the field. When we can adequately fit the spectrum, typically requiring a $S/N \geq 10$, the brightness is calculated correctly at the well-defined reference frequency of 1.28 GHz. The spectral fit was accepted if the χ^2 per degree of freedom was less than 1.5 times the χ^2 per degree of freedom based on assuming the default value of −0.6; otherwise the default value is reported and the central frequency brightness is calculated with that spectral assumption. This results in an underestimate of the brightness by ∼6% (10%) for a true spectral index of −1 (−2). Typically, these are less than the random errors in these low brightness cases.

Since the flux density threshold for switching to the default spectral index depends on the local noise properties of each cluster image, no global value can be specified. However, inspection of the spectral indices as a function of flux density for a selected region reveals the appropriate local threshold below which the default spectral index has been assigned. Caution is still advised when examining regions of maps where spectral indices very close to the default value are encountered.

4.4.3. Largest angular size

MeerKAT has very good short baseline coverage, allowing recovery of extended emission. However, the minimum baseline length of 29 m does restrict the maximum size of a structure that is properly imaged. This maximum size scales inversely with frequency. Angular scales of less than 10′ should be fully recovered. Larger scales are less well imaged, especially at the higher frequencies. This effect leads to negative holes around bright extended structures and an artificial apparent steepening of the spectrum. Very steep fitted spectra of extended emission should be treated with caution.

4.4.4. Astrometry

A number of instrumental issues can affect the accuracy of the astrometry. As detailed in this section, we correct the image astrometry in each field by matching the positions with those of the respective optical hosts. This results in a final accuracy better than 0.3″ in the high S/N limit. However, the astrometric errors are still present in the visibility data, and users that reprocess raw data need to take these into consideration.

Data in the earliest observations suffered from a 2-second time offset in the labelling of the data, and a half-channel frequency error, which propagate into errors in the u, v, w coordinates. The column 'Fix' in Table 1 indicates cluster fields where this issue was corrected in the visibilities. Fields that were not fixed will contain rotation and scaling errors in the positions of sources that depend on position in the field. These can be as large as 2″ at the edge of the fields.

Calibrator position errors were present in the initial MeerKAT calibrator list, affecting several calibrators up to a level of several arcseconds. This results in an approximately constant position offset of sources in the affected cluster fields. Corrections were made in the images when incorrect calibrator positions were discovered, but the final corrections were made at the end, in any case, through optical cross-matching.

A low-accuracy delay model in the correlator in use during the observations can cause a similar, albeit subtler, problem when the calibrator and target are widely separated. An additional bias is possible when the calibrator does not sufficiently dominate the visibility intensities in the field. The model had insufficient accuracy to reliably transfer phases measured on the calibrator to those of the target, especially since many of the astrometric calibrators were 10° or more from the target. This resulted in approximately constant position offsets of up to several arcseconds in individual fields, which had to be corrected.

Astrometric corrections were made for each field centre, removing or reducing the effects noted above. We matched compact radio components to a large number of background quasars, radio galaxies, and star-forming galaxies in each field, using the optical and infrared catalogues from the Dark Energy Camera Legacy Survey (DECaLS; Dey et al. 2019). A flux density-weighted average correction was determined for every cluster field, as indicated in column 'Posn' in Table 1, and appropriate corrections made to each corresponding image. Residual systematic errors should be under 0.1″ for clusters as a whole, although errors may be larger for individual sources.

Additional astrometric checks were carried out after the above correction by cross-matching with sources from the International Celestial Reference Frame (ICRF) catalogue (Charlot et al. 2020). Since the ICRF constitutes the most accurately known set of astronomical positions, they are an ideal check of the astrometry in the MGCLS catalogue described in Sect. 5.2. There are eight ICRF sources in our cluster fields. They are bright (0.3 to 1.8 Jy), and the statistical uncertainty in their MeerKAT position determinations are therefore small. They were chosen for the ICRF for being compact on milliarcsecond scales, and they generally do not have structure on the MeerKAT's ∼8″ scale, which might affect the position determination. We therefore compared our catalogue positions with the ICRF3[8] ones for the eight sources that were included in our fields: ICRF J010645.1−403419, J025612.8−213729, J031757.6−441417, J033413.6−400825, J060031.4−393702, J062552.2−543850, J124557.6−412845, and J133019.0−312259.

We found that most of the MGCLS catalogue positions differed from the ICRF3 ones by less than 1″ in either coordinate. The single exception was ICRF J025612.8−213729 (QSO B0253−218) for which the MGCLS catalogue position was ∼2.3″ north of the ICRF3 position. This source is an exception to the above generalisation about lack of structure, in that Reid et al. (1999) show it to be somewhat extended in a N–S direction at 5 GHz, with several components spread out over ∼10″. This field, MACS J0257.6−2209, is also one of the few for which there were uncorrected timing and frequency errors (Col. 5 in Table 1, discussed earlier in this section) that will cause position errors far from the field centre. In addition, since the source is resolved, the MGCLS position at 1.28 GHz and ∼8″ resolution could differ from the ICRF3 position, which is that of the milliarcsecond core only at 5 GHz.

Ignoring this source, we find that for the remaining seven sources the difference between the MGCLS catalogue positions and the ICRF3 ones were -0.04 ± 0.34″ in RA and -0.02 ± 0.15″ in Dec. This check gives us confidence that the positions in the MGCLS catalogue are accurate, and adopt the uncertainty of 0.36″ based on the ICRF3 comparisons.

4.4.5. Polarisation

The first plane of the Stokes Q and U *5pln* cubes (see Sect. 4.2.1) provide a good indication of where significant polarisation is present, but should not be used quantitatively on their own. Each image was created by a noise-weighted sum of the frequency planes in the full cube, which is strictly correct only when both the RM and the spectral index are zero, and when no depolarisation is present. At any RM ≠ 0, the amplitude of Q and U will be reduced, reaching a factor of 2 reduction at $|\mathrm{RM}| \sim 25\,\mathrm{rad\,m^{-2}}$, depending on the source spectral index and the noise in the different frequency channels. The first plane of the Stokes Q and U *5pln* cubes should therefore be used quantitatively with caution.

We note that polarisation leakage affects the upper half of the band, with the residual polarisation leakage increasing with distance from the field centre within the half-power region of the beam (de Villiers & Cotton, priv. comm.; de Villiers et al. 2021). The instrumental polarisation may reach up to 10% in the upper part of the band, while it is typically less than 2% at the lower frequencies. Users should evaluate how the leakage affects their particular science case.

5. MGCLS source catalogues

We produced source catalogues for all fields in the MGCLS, based on the intensity plane of the full-resolution enhanced products described in Sect. 4.2. Here we detail our source finding

[8] http://hpiers.obspm.fr/icrs-pc/newwww/icrf/index.php

method, as well as the various catalogues being provided with the legacy data products. Limitations to the accuracy of the results are discussed in Sect. 4.4. It is very important to note that these catalogues are not complete, and any statistical analyses must consider their limitations. In particular, the sensitivity depends on the distance from the respective field centres, and regions around bright sources are excluded, as discussed below.

5.1. Source detection

We used the Python Blob Detection and Source Finder (PYBDSF; Mohan & Rafferty 2015) software to create individual source catalogues for all MGCLS fields, using the full-resolution, primary-beam-corrected data products. PYBDSF searches for islands of emission and attempts to fit models consisting of one or more elliptical Gaussians to them. Gaussians are then grouped into sources, and there may be more than one source per island. Each source is given a code: 'S' for single-Gaussian sources that are the only source on its island, 'C' for single-Gaussian sources that share an island with other sources, and 'M' for sources composed of two or more Gaussian fits. We used the default $3\sigma_{\rm img}$ island boundary threshold and $5\sigma_{\rm img}$ source detection threshold, where $\sigma_{\rm img}$ is the local image rms. As many images have variable image noise levels across the field, we allow PYBDSF to calculate the 2D rms map during the source finding.

For sources in regions of high image noise, for example those near bright sources with strong sidelobes, the typical statistical uncertainty in peak source brightness is a factor of ~2 larger than for sources elsewhere in the same field. Spurious source detections are common around very bright sources, with PYBDSF sometimes cataloguing sidelobes as sources. To mitigate spurious detections in our DR1 catalogues, we excised all catalogue entries around bright sources. A source was considered bright if its peak brightness was higher than the bright source limit for that field, $I^{\rm bs}_{\rm lim}$. This limit is connected to the image dynamic range, such that

$$I^{\rm bs}_{\rm lim} = 10^{-4} \times \frac{I_{\rm max}}{\sqrt{\sigma_{\rm global}}}\ {\rm Jy\,beam^{-1}}, \quad (2)$$

where $I_{\rm max}$ is the maximum source brightness in the image, and $\sigma_{\rm global}$ is the median image rms, both in units of Jy beam^{-1}. The region around a bright source within which catalogue entries were excised, $r_{\rm cut}$, scales with the source brightness:

$$r_{\rm cut} = 0.005 \times \left(1 + \log_2 \frac{I^{\rm bs}_{\rm peak}}{I^{\rm bs}_{\rm lim}}\right) {\rm deg}, \quad (3)$$

where both $I^{\rm bs}_{\rm lim}$ and $I^{\rm bs}_{\rm peak}$, the peak brightness of the bright source, are in Jy beam^{-1}. A median of 2.6% of nominally detected sources were removed per field through this process.

5.2. Compact source catalogue

From our PYBDSF results, we compiled a single MGCLS compact source catalogue from all fields, only including sources that could be fit with a single Gaussian component (source codes 'S' or 'C'), after the spurious source excision. The full DR1 catalogue[9] contains ~626 000 sources from the 115 cluster fields, with an excerpt shown in Table 2. The catalogue columns, described in the Table caption, contain standard radio source information including the integrated flux density, peak brightness, and source size, with catalogue source positions provided in decimal degrees. The source identifier (first column of the catalogue) uses an International Astronomical Union (IAU) classification, with the designation MKTCS JHHMMSS.ss±DDMMSS.s, where the decimal positional information is truncated, rather than rounded.

5.2.1. Comparison with previous radio catalogues

To verify the MGCLS compact source flux densities, we compare them to those from other radio surveys. To cover all MGCLS pointings, we use catalogues from both the 1.4 GHz NRAO VLA Sky Survey (NVSS; north of −40° Dec; Condon et al. 1998) and the 843 MHz SUMSS survey (south of −30° Dec; Bock et al. 1999). To scale the NVSS and SUMSS flux densities to the MGCLS reference frequency of 1.28 GHz, we assume a power law $S_\nu \propto \nu^\alpha$ with a fiducial spectral index of $\alpha = -0.7$ (Smolčić et al. 2017).

To avoid incompleteness effects due to differences in sensitivity between the three surveys, we consider only sources with $S/N \geq 50$ in all three surveys. This high limit also minimises effects from additional faint MGCLS sources within the larger NVSS and SUMSS beams, as noted below. To cross-match MGCLS compact sources with their NVSS and SUMSS counterparts, we use a 5″ radius. This radius is a compromise between maximising the number of real counterparts and minimising the number of spurious matches. By shifting the MGCLS sources by 1′ and repeating the cross-matching, we determine that the expected percentage of spurious matches is 4.1% and 3.8% for NVSS and SUMSS, respectively. These have a negligible effect on the flux density comparisons.

Our cross-matching yields a total of 398 and 565 compact MGCLS sources with NVSS and SUMSS counterparts, respectively. Figure 3 shows the flux density scale comparison between the MGCLS compact sources and the scaled NVSS (top panel) and scaled SUMSS (bottom panel) sources. The MGCLS flux densities are in good agreement with those of both NVSS and SUMSS. We fit a power law of the form

$$S_{\rm scaled\_cat} = \kappa \times S^{\gamma}_{\rm MGCLS} \quad (4)$$

to the MGCLS and scaled NVSS and SUMSS flux densities. Values of 1 for both κ and γ would indicate an exact one-to-one correspondence. We obtain fit values of $\kappa_{\rm NVSS} = 1.06 \pm 0.02$, $\gamma_{\rm NVSS} = 1.00 \pm 0.01$, and $\kappa_{\rm SUMSS} = 1.06 \pm 0.04$, $\gamma_{\rm SUMSS} = 1.00 \pm 0.01$ for the NVSS and SUMSS comparisons, respectively. This is consistent with a linear relation between the MGCLS and scaled fluxes, with the scaled NVSS and SUMSS fluxes being marginally higher than those from MGCLS. The MGCLS sources were not chosen to be isolated, so the poorer resolution of the sky surveys (~45″) may lead to a single NVSS or SUMSS source having contributions from additional faint MGCLS sources; at the lowest fluxes, this could result in a small bias in opposite spectral index directions for the two surveys. However, the equality of the fitted flux densities show that any such effects are negligible.

We also examine the spectral index distribution of the radio cross-matched MGCLS compact sources. For this purpose, we include 148 MHz data from the TIFR GMRT Sky Survey (TGSS; north of −53° Dec; Intema et al. 2017) in addition to the NVSS and SUMSS surveys. Given the different sky coverages, some MGCLS sources have flux densities at only one additional frequency, whereas others have three. We fit a single $S_\nu \propto \nu^\alpha$ power-law and inspected each fit to discard both spurious matches and sources showing spectral curvature. Figure 4 shows the resulting spectral index distributions. The distributions of

[9] Available at https://doi.org/10.48479/7epd-w356

Table 2. Excerpt of the MGCLS compact source catalogue at 1.28 GHz.

(1) Src. name MKTCS	(2) RA_{J2000} (deg)	(3) Dec_{J2000} (deg)	(4) ΔRA (deg)	(5) ΔDec (deg)	(6) $S_{tot}^{1.28\,GHz}$ (mJy)	(7) ΔS_{tot} (mJy)	(8) $I_{peak}^{1.28\,GHz}$ (mJy b^{-1})	(9) ΔI_{peak} (mJy b^{-1})	(10) s_{max} ($''$)	(11) s_{min} ($''$)	(12) $s_{p.a.}$ ($^\circ$)	(13) Field
J001059.77−190940.3	2.7491	−19.1612	0.0000	0.0000	0.695	0.020	0.669	0.011	7.6	7.5	6	Abell_13
J001059.94−190654.9	2.7498	−19.1153	0.0001	0.0002	0.113	0.028	0.081	0.013	9.8	7.8	21	Abell_13
J001059.14−195204.2	2.7464	−19.8679	0.0002	0.0003	0.106	0.027	0.056	0.010	14.2	7.3	143	Abell_13
J001059.23−194540.7	2.7468	−19.7613	0.0002	0.0003	0.077	0.024	0.045	0.010	11.1	8.4	153	Abell_13
J001059.50−192405.3	2.7479	−19.4015	0.0001	0.0001	0.050	0.012	0.060	0.007	8.0	5.8	179	Abell_13
J002318.34−254121.6	5.8264	−25.6894	0.0001	0.0002	0.053	0.016	0.061	0.010	7.4	6.5	10	Abell_22
J002317.08−253627.2	5.8212	−25.6076	0.0000	0.0000	0.939	0.021	0.827	0.011	8.3	7.6	36	Abell_22
J002318.08−253621.5	5.8253	−25.6060	0.0002	0.0001	0.073	0.021	0.067	0.011	8.5	7.1	62	Abell_22
J002317.47−252111.6	5.8228	−25.3532	0.0001	0.0001	0.242	0.030	0.198	0.015	8.6	7.8	143	Abell_22
J002317.15−261532.9	5.8215	−26.2592	0.0003	0.0001	0.573	0.095	0.156	0.021	19.6	10.3	80	Abell_22

Notes. The full catalogue, which includes all cluster fields, is available online at https://doi.org/10.48479/7epd-w356. Columns: (1) MGCLS source ID using the IAU designation of the form MKTCS JHHMMSS.ss±DDMMSS.s, where the decimals are truncated; (2)–(5) J2000 RA and Dec, and associated 1σ uncertainty, respectively; (6) and (7) Total integrated Stokes-I flux density and associated 1σ uncertainty at the reference frequency, respectively; (8) and (9) Peak Stokes-I brightness and associated 1σ uncertainty, respectively; (10) and (12) Source size: FWHM of the major and minor axes of the source, and source p.a.; (13) Cluster field of the source. All uncertainties are statistical only and are determined from the images as per Condon (1997).

spectral indices for sources with two, three, and four frequency points are similar, although the $N_\nu = 4$ distribution has a lower median spectral index. Subtle differences may reflect differences in the spectral populations of the different surveys, but are beyond the scope of the current work.

5.2.2. Optical cross-matching

We created optical cross-match catalogues for the compact sources in the Abell 209 and Abell S295 fields using data from DECaLS (Dey et al. 2019). These fields were selected due to their existing DECaLS coverage and their decent MGCLS dynamic range. Cross-match catalogues for other MGCLS fields will be compiled in follow-up works.

To identify DECaLS counterparts for the MGCLS compact sources, we used the likelihood ratio (LR) method (Sutherland & Saunders 1992; Laird et al. 2009; Smith et al. 2011). The LR here is defined as the ratio of the probability that an optical source (at a given distance from the radio position and with a given optical magnitude) is the true counterpart, to the probability that the same source is a spurious alignment, that is, $\mathrm{LR} = (q(m) \times f(r))/n(m)$, where $q(m)$ is the expected number of true optical counterparts with magnitude m, $f(r)$ is the probability distribution function of the positional uncertainties in both the radio and the optical source catalogues, and $n(m)$ is the background density of optical galaxies of magnitude m in the DECaLS r band (or g or z band). The magnitudes are AB magnitudes from DECaLS DR8.

The a priori probability $q(m)$ is determined as follows. First, the radio and optical source catalogues are matched by finding the closest counterpart within a fixed search radius of $4''$. We chose this radius based on our distributions of cross-matches, as shown in Fig. 5, where the distributions ($\mathrm{LR} > 0.5$) of the RA and Dec angular separations between the positions of radio sources can be approximated by Gaussians. The search radius of $4''$ is the optimal radius where we detect the most true sources. Radii greater than $4''$ will sharply increase the number of spurious matches. The number of spurious matches (of magnitude m) is estimated by scaling $n(m)$ to the area of $4''$ radius within which we search for counterparts. This is then subtracted from the number of counterparts (as a function of magnitude) to determine the number of true associations, $q(m)$.

In our case, the probability distribution $f(r)$ is a two-dimensional Gaussian distribution:

$$f(r) = \frac{1}{2\pi\delta^2} \exp\left(-\frac{r^2}{2\delta^2}\right). \tag{5}$$

Here, r is the angular distance (in arcsec) from the radio source position, and δ is the combined positional error given by $\sqrt{\delta_{\mathrm{decals}}^2 + \delta_{\mathrm{mgcls}}^2}$, where δ_{decals} is the positional uncertainty from the DECaLS catalogue, and δ_{mgcls} is the positional uncertainty from the MGCLS compact source catalogue. For each source in the MGCLS DR1 compact source catalogue, we adopted an elliptical Gaussian distribution for the positional errors, with the uncertainties in RA and Dec on the radio position reported in the radio catalogue. We assume a systematic optical position uncertainty of $0.2''$ in both RA and Dec for the DECaLS catalogue (Dey et al. 2019).

The optical cross-match catalogue includes all matches within $4''$ of a compact MGCLS source detection, along with their LR probabilities and flags (no LR cutoff imposed). Table 3 shows an excerpt from the optical cross-match catalogue for the Abell 209 field, with the full catalogues for it and the Abell S295 fields available online[10]. The presence of more than one counterpart for a particular radio source provides additional information to that contained in the LR itself, which can then be used to estimate the reliability of the counterpart source, or the probability that a particular source is the correct counterpart. The reliability for radio source i, as defined by Sutherland & Saunders (1992), is calculated as

$$\mathrm{REL}_i = \frac{\mathrm{LR}_i}{\sum \mathrm{LR}_{\mathrm{search\,radius}} + (1 - Q)}, \tag{6}$$

where $\sum \mathrm{LR}_{\mathrm{search\,radius}}$ is the sum of LR for all possible DECaLS counterparts to the radio source within our search radius of $4''$, and Q is the fraction of MGCLS compact radio sources with optical counterparts above the DECaLS magnitude limit. Comparison of $\sum \mathrm{REL}_i$ with the total number of counterparts with

[10] https://doi.org/10.48479/7epd-w356

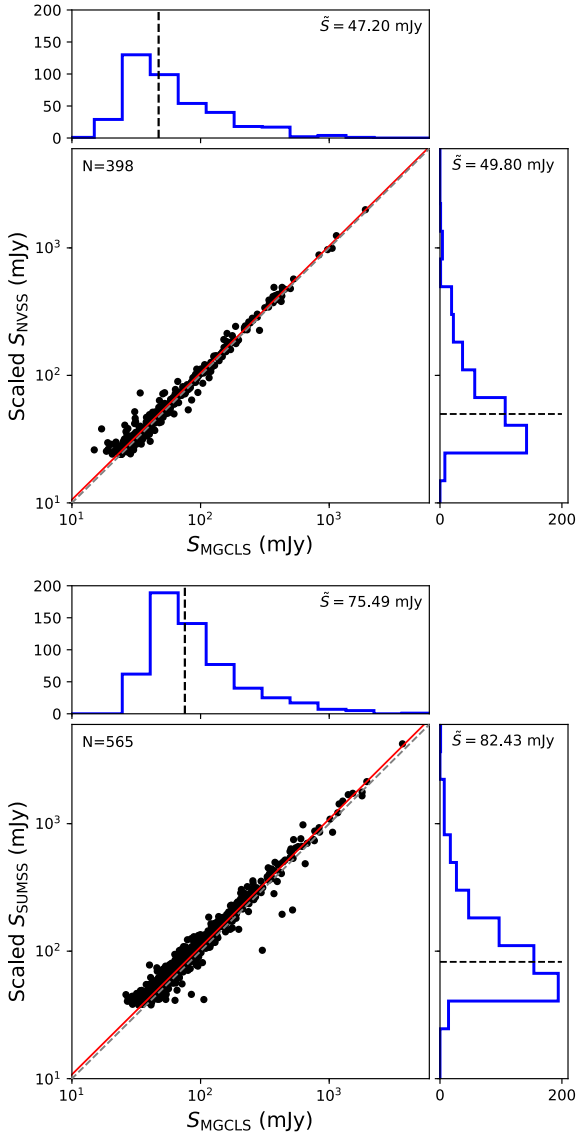

Fig. 3. Comparison between the integrated flux density of MGCLS compact sources with their counterparts in the NVSS (*top*) and SUMSS (*bottom*) catalogues, with the latter two being scaled to the MGCLS frequency of 1.28 GHz. The best-fit relation in each case is shown by the solid red line and is consistent with a linear relationship, with the number of sources in the fit shown in the upper left corner. The dashed grey line shows the exact one-to-one relationship. The scaled sky survey flux densities are typically 6% higher than their MGCLS counterparts, so this represents a possible small bias in the MGCLS flux scale. Histograms show the relevant flux density distributions, with the dashed black lines indicating the respective median values, $\tilde{S}$.

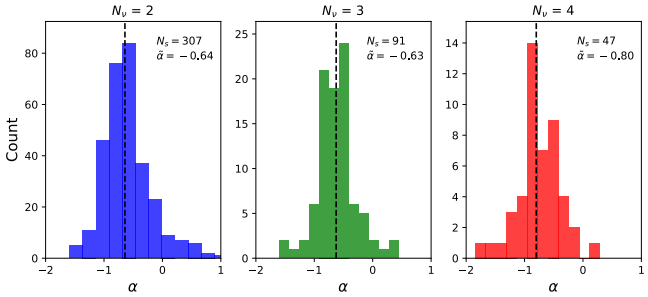

Fig. 4. Spectral index distributions for the cross-matched MGCLS compact sources, using data from the MGCLS (1.28 GHz), NVSS (1.4 GHz), SUMSS (843 MHz), and TGSS (148 MHz) catalogues. *Left to right*: distributions for MGCLS sources with flux density measurements at two, three, and four frequencies, respectively. The number of compact sources, N_s, and the median spectral index, $\tilde{\alpha}$, are indicated in the upper right corner of each panel, and the respective median spectral indices are indicated by vertical dashed lines.

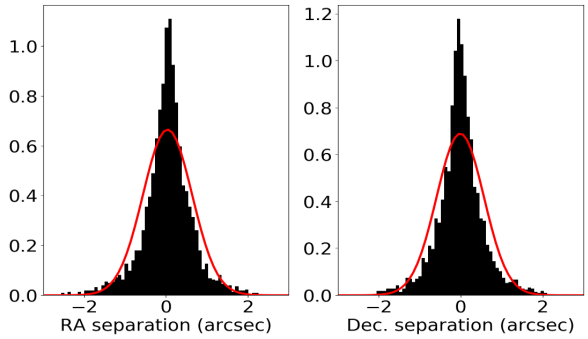

Fig. 5. Histograms of the RA and Dec angular separations between the positions of radio sources in the Abell 209 compact source catalogue and their optical counterparts, for LR > 0.5. In each panel, the red line shows the normalised Gaussian distribution.

LR > LR_{cutoff} provides an estimate of the spurious identification rate, or error rate (ER). The choice of the cutoff in LR is a trade-off between maximum completeness and maximum purity. Completeness is defined as the fraction of radio catalogue sources that have an optical counterpart, and purity (given by 1 – ER) is the fraction of radio-optical source matches that are real.

Figure 6 shows the completeness and purity for both Abell 209 and Abell S295, with no LR cutoff imposed. Our chosen LR_{cutoff} of 0.5 is indicated by the vertical dotted line. A value of LR_{cutoff} = 0.5 corresponds to an estimated spurious identification rate of 4.5% in Abell 209, with 59% (2723 of 4581) of radio sources in the Abell 209 compact source catalogue having optical counterparts.

5.3. Extended sources

We do not provide catalogues for the extended sources, that is, those for which the PYBDSF fit required multiple Gaussians, indicated in the PYBDSF output with code 'M'. These need to be verified visually, a process that is extremely time consuming. However, to provide an indication of the number of extended sources in the MGCLS, we performed this verification for sources in the Abell 209 and Abell S295 fields.

Extended sources can be separated into two categories: blended sources (i.e. those with overlapping Gaussian components) and multi-component sources with multiple, non-overlapping, and often visually separable components. There are 158 and 347 blended sources in the Abell 209 and Abell S295 fields, respectively, roughly 3–5% of the number of compact sources in the respective fields. Their integrated flux densities range from 85 μJy to 105 mJy, with the largest blended sources being just over 30″ across. Extrapolating these numbers to the full MGCLS sample, there are of the order of 29 000 blended sources in the full survey.

We defined the multi-component sources as those with distinct structures such as jets, cores, or lobes. Identifying the different components that comprise a single source can be difficult,