

# SuperCLASS – I. The super cluster assisted shear survey: Project overview and data release 1

Richard A. Battye,<sup>1</sup>★ Michael L. Brown,<sup>1</sup> Caitlin M. Casey,<sup>2</sup> Ian Harrison<sup>1,3</sup>,  
Neal J. Jackson,<sup>1</sup> Ian Smail<sup>1,4</sup>, Robert A. Watson,<sup>1</sup> Christopher A. Hales,<sup>5,6</sup>  
Sinclair M. Manning<sup>1,2</sup>, Chao-Ling Hung<sup>1,2</sup>, Christopher J. Riseley,<sup>7,8,9</sup>  
Filipe B. Abdalla,<sup>10</sup> Mark Birkinshaw,<sup>11</sup> Constantinos Demetroullas,<sup>1,12</sup>  
Scott Chapman,<sup>13</sup> Robert J. Beswick,<sup>1</sup> Tom W. B. Muxlow,<sup>1</sup> Anna Bonaldi<sup>1,14</sup>,  
Stefano Camera<sup>1,15,16</sup>, Tom Hillier,<sup>1</sup> Scott T. Kay<sup>1</sup>, Aaron Peters,<sup>1</sup>  
David B. Sanders,<sup>17</sup> Daniel B. Thomas,<sup>1</sup> A. P. Thomson,<sup>1</sup> Ben Tunbridge,<sup>1</sup>  
and Lee Whittaker<sup>1,10</sup> (SuperCLASS Collaboration)

*Affiliations are listed at the end of the paper*

Accepted 2020 February 25. Received 2020 February 18; in original form 2019 November 15

## ABSTRACT

The SuperCLuster Assisted Shear Survey (SuperCLASS) is a legacy programme using the *e*-MERLIN interferometric array. The aim is to observe the sky at *L*-band (1.4 GHz) to a r.m.s. of  $7 \mu\text{Jy beam}^{-1}$  over an area of  $\sim 1 \text{ deg}^2$  centred on the Abell 981 supercluster. The main scientific objectives of the project are: (i) to detect the effects of weak lensing in the radio in preparation for similar measurements with the Square Kilometre Array (SKA); (ii) an extinction free census of star formation and AGN activity out to  $z \sim 1$ . In this paper we give an overview of the project including the science goals and multiwavelength coverage before presenting the first data release. We have analysed around 400 h of *e*-MERLIN data allowing us to create a Data Release 1 (DR1) mosaic of  $\sim 0.26 \text{ deg}^2$  to the full depth. These observations have been supplemented with complementary radio observations from the Karl G. Jansky Very Large Array (VLA) and optical/near infrared observations taken with the Subaru, Canada-France-Hawaii, and *Spitzer* Telescopes. The main data product is a catalogue of 887 sources detected by the VLA, of which 395 are detected by *e*-MERLIN and 197 of these are resolved. We have investigated the size, flux, and spectral index properties of these sources finding them compatible with previous studies. Preliminary photometric redshifts, and an assessment of galaxy shapes measured in the radio data, combined with a radio-optical cross-correlation technique probing cosmic shear in a supercluster environment, are presented in companion papers.

**Key words:** galaxies: evolution – large-scale structure of Universe – cosmology: observations – radio continuum: galaxies.

## 1 INTRODUCTION

The extragalactic radio source population is a powerful tool to probe a range of astrophysical processes and a number of large scale surveys have been performed covering a wide range of radio frequencies. The two largest (in terms of area) radio source surveys performed at *L*-band (around 1.4 GHz) are the Faint Images of

the Radio Sky at Twenty centimeters (FIRST) (Becker, White & Helfand 1995) and the NRAO-VLA Sky Survey (NVSS) (Condon et al. 1998) which have each detected  $\sim 10^6$  sources and probed the source population to  $\sim 1 \text{ mJy}$  over significant ( $> 10^4 \text{ deg}^2$ ) areas of the sky. A significant fraction of these sources are AGN/jet-driven sources, but as one goes lower in flux density it is expected that the radio source population will become dominated by lower luminosity star-forming galaxies (SFGs). This has been confirmed by a number of deeper but much smaller area surveys that have been used to probe the source counts down to flux densities  $\sim 20 \mu\text{Jy}$ , for example,

★ E-mail: richard.battye@manchester.ac.uk

Owen & Morrison (2008), Smolčić et al. (2017), Schinnerer et al. (2010) Richards et al. (1999), Muxlow et al. (2005), Mainieri et al. (2008), Padovani et al. (2009), Ibar et al. (2009), Bonzini et al. (2013), Padovani et al. (2015). Most notably in the present context is a survey of the Hubble Deep Field - North (HDF-N), using the Very Large Array (VLA) and the Multi-Element Radio Linked Interferometer Network (MERLIN) (Muxlow et al. 2005), which detected a sample of  $\sim 100$  radio sources with flux density greater than  $40 \mu\text{Jy}$  in an area of  $\sim 10 \text{ arcmin} \times 10 \text{ arcmin}$  with a maximum baseline corresponding to a resolution of  $\sim 200 \text{ mas}$ .

The advent of the Square Kilometre Array (SKA) sometime in the next decade will allow one to perform surveys at  $\mu\text{Jy}$  flux densities over significant areas of sky with sub-arcsec resolution. This will enable a wide-range of science applications as elucidated in Braun et al. (2015). In addition to investigating the nature of the radio sources, and hence the astrophysical processes involved, the increased source density will allow the radio sources to be used as cosmological probes, including weak-lensing surveys (Bacon et al. 2015; Brown et al. 2015; Camera et al. 2015; Jarvis et al. 2015). Large  $L$ -band surveys with SKA pathfinders – MIGHTEE with MeerKAT (Jarvis et al. 2016) and EMU with ASKAP (Norris et al. 2011) – will provide important insights into the source populations for these studies, but cannot measure the small angular scales necessary for detailed morphology studies.

Gravitational lensing is the deflection of electromagnetic radiation by mass concentrations predicted by the theory of General Relativity that can be used to probe the so-called ‘dark components’ of the Universe – dark matter and dark energy (e.g. Copeland, Sami & Tsujikawa 2006) – and possible modifications to the theory of gravity (e.g. Clifton et al. 2012). Depending on the level of alignment between the source and lensing matter, it can be either strong, whereby multiple images are formed, or weak which leads to the distortion of the shapes of galaxies. This weak lensing effect is coherent across the sky, known as cosmic shear, and is now considered to be one of the primary cosmological tests (Bartelmann & Schneider 2001). If galaxies are randomly orientated, the cosmic shear two-point correlation function, or equivalently the shear power spectrum, can be inferred by measuring large numbers of galaxy shapes with the main source of statistical uncertainty being due to the dispersion in the intrinsic shapes of the galaxies.

The effects of cosmic shear have been measured in the optical waveband. This started with small areas (Tyson, Valdes & Wenk 1990; Mould et al. 1994; Bacon, Refregier & Ellis 2000; Kaiser, Wilson & Luppino 2000; Van Waerbeke et al. 2000; Wittman et al. 2000) and now the state-of-the-art surveys cover  $\sim 100 \text{ deg}^2$  (Hildebrandt et al. 2017; Troxel et al. 2018; Hikage et al. 2019). Future surveys to be performed with the *Euclid* satellite (Laureijs et al. 2011) and LSST (The LSST Dark Energy Science Collaboration 2018) – the so-called stage IV surveys – should achieve source densities of more than  $10 \text{ arcmin}^{-2}$  over areas of  $\sim 10^4 \text{ deg}^2$ . Interestingly, the large-scale surveys performed by the SKA should reach similar levels with sub-arcsec resolution in the next decade and therefore it is realistic to imagine performing competitive and complementary lensing surveys using the SKA (Bonaldi et al. 2016; Harrison et al. 2016; Camera et al. 2017; Square Kilometre Array Cosmology Science Working Group et al. 2018).

As already pointed out source densities in the radio waveband are typically much lower than the optical. Nonetheless, cosmic shear has been detected at the  $3\sigma$ -level in the radio waveband using the FIRST survey with a source density of  $\sim 100 \text{ deg}^{-2}$  (Chang, Refregier & Helfand 2004). This pioneering work was followed up in Patel et al.

(2010) where ellipticities were measured for radio sources detected in the MERLIN/VLA observations of HDF-N. More recently Demetroullas & Brown (2016, 2018) have found shear correlations between the radio and optical, while Tunbridge, Harrison & Brown (2016) and Hillier et al. (2019) have investigated shear correlations in the COSMOS field, at 1.4 and 3 GHz respectively, with both being limited by the lack of small angular scales observable with the VLA.

The primary science goal of the SuperCLASS project is to detect the effects of weak lensing in the radio band in a survey which is representative of the source population and resolution of that which will ultimately be observed by the SKA. The high-resolution afforded by the combination of two telescopes which have both been recently upgraded – *e*-MERLIN and the Karl G. Jansky Very Large Array (VLA) – is very similar to that expected for the SKA and has sufficient sensitivity to probe the population of star-forming galaxies expected to dominate SKA surveys, albeit over much smaller areas of the sky. It will allow the development of the tools necessary for shape measurement and a quantitative assessment, for the first time, of the physical properties of the radio sources which can be used for cosmic shear measurements such as the size distribution and the expected r.m.s. ellipticity. Shape measurement techniques in the optical waveband are well developed (e.g. Massey et al. 2013, and references therein), but high resolution radio observations necessarily require an interferometric array. This can be a positive in terms of reducing the impact of the atmosphere, but it also means that the ideas used in optical shape measurement need to be reassessed at the very least and, indeed, the fact that the observations are done in the Fourier domain allows for new possibilities (Rivi et al. 2016; Rivi & Miller 2018; Rivi et al. 2019).

There are a number of reasons why it is interesting to pursue radio weak lensing:

(i) The key instrumental systematic in ground-based optical shear measurements is the stability, both in time and space, of the telescope point spread function (PSF) which has to be deconvolved in order to infer ellipticities. This problem is due to the fact that the observations are affected by atmospheric seeing. In contrast, observations at radio frequencies around  $L$ -band will be diffraction limited and, in principle, the PSF can be calculated directly from the optical properties of the telescope(s) involved.

(ii) It has been suggested that there is extra information which can be gleaned from the radio observations that can be used to mitigate some of the astrophysical systematics that can hamper cosmic shear measurements. These include the use of polarization and the H I line profile to infer information about the intrinsic orientation of a galaxy and hence mitigate the impact of intrinsic alignments (Brown & Battye 2011a, b; Whittaker, Brown & Battye 2015), as well as the possibility of statistically measuring the redshift distribution of the sources directly from the radio observations (Harrison, Lochner & Brown 2017).

(iii) Whatever observational systematics remain in the radio inferred shape measurements, they will be of very different origin to those present in those from the optical waveband. Therefore, if one were to have radio and optical observations of the same fields from *Euclid*, LSST, and SKA, as would seem prudent, then it would be possible to cross-correlate the two, removing any uncorrelated systematic effects from the telescopes, such as incorrectly deconvolved PSFs (Patel et al. 2010; Demetroullas & Brown 2016; Camera et al. 2017).

To date, there have been no surveys explicitly designed to measure weak-lensing with radio galaxies. Furthermore, the SKA pathfinders

are not suitable for such explorations due to a lack of long baselines accessing small angular scales. This work aims to fill this gap using a specific combination of radio telescopes.

An additional and equally important science goal is to use the fact that radio emission is an unbiased tracer of star-formation to disentangle the effects of star-formation and active galactic nucleus (AGN) activity out to  $z \sim 1$ . This is of particular interest in the context of Dusty Star-Forming Galaxies (DSFGs) (Smail, Ivison & Blain 1997; Chapman et al. 2005; Casey, Narayanan & Cooray 2014) which have implied star-formation rates (SFRs) of  $\sim 100\text{--}1000\text{ M}_\odot\text{ yr}^{-1}$  and are believed to contribute significantly to the SFR of the Universe at high redshifts. Due to the significant levels of dust obscuration present in these sources, it is not always possible to perform morphological analyses in the optical waveband. Relatively small pilot studies have suggested that such classifications are possible using high resolution radio continuum observations (Biggs & Ivison 2008; Casey et al. 2009a,b; Thomson et al. 2019). Fortunately, the survey characteristics necessary to find larger samples, in terms of resolution, depth, and area covered, are very similar to those required for weak lensing.

Key to both of these science goals are the unique capabilities of *e*-MERLIN in the pre-SKA era. *e*-MERLIN is an upgrade of MERLIN that comprises seven telescopes based in the UK with a maximum baseline length of 217 km corresponding to angular scales of  $\approx 200$  mas at *L*-band (Thomasson 1986). The upgrade has involved replacement of the microwave links which transferred data to the central correlator with fibre optic cables allowing a substantial increase in bandwidth (512 MHz at *L*-band) and improved receivers; the overall sensitivity improvement when compared to MERLIN at *L*-band is around a factor of 5. Importantly for the science goals discussed in the next section, *e*-MERLIN has the unique ability to resolve star-forming galaxies at these frequencies allowing us to measure their ellipticities and perform a separation of this population from that which is AGN/jet-driven using morphological analysis.

SuperCLASS is part of the *e*-MERLIN legacy programme which comprises 12 large projects which have been allocated  $\sim 50$  per cent of the observing time over the first  $\sim 5$  yr of operation.<sup>1</sup> A number of these programmes have complementary goals to those of SuperCLASS in that they are trying to understand the nature of star/galaxy formation across cosmic time. These ‘sibling’ projects include the *e*-MERLIN Galaxy Evolution survey (e-MERGE,<sup>2</sup> Muxlow et al in preparation) which is performing a very deep  $1\text{ }\mu\text{Jy}$  r.m.s. high resolution survey of the HDF-N following up on Muxlow et al. (2005), Legacy *e*-MERLIN Multi-Band Imaging of Nearby Galaxies (LEMMINGS) whose goal is to make high resolution radio images of a number of nearby galaxies and Astrophysics of Galaxy Transformation and Evolution (AGATE) that will probe the radio source population in the direction of known massive clusters. The combined knowledge gleaned from these projects should lead to a significant advance in understanding the nature of galaxy formation and evolution.

The outline of this paper is as follows: in Section 2 we describe the selection of the SuperCLASS field and the rationale behind the multiwavelength observations which have been taken as part of the project. In Section 3 we describe the individual multiwavelength observations in more detail. Section 4 describes the definition of

our Data Release 1 (DR1) survey region, the creation of source catalogues for the radio data, and the cross-matching of these sources with catalogues from the optical data. In Section 5 we describe basic properties of these sources, in particular their sizes, fluxes, and Sérsic profiles as constrained by the data available to us. Section 6 presents our summary and conclusions.

For more detail on the optical redshifts and Spectral Energy Distributions of the radio sources, we refer the reader to Manning & The SuperCLASS Collaboration (2019) (henceforth Paper II) whilst radio and optical shape measurements and weak lensing power spectra constraints are presented in Harrison & The SuperCLASS collaboration (2020) (henceforth Paper III).

## 2 SUPERCLASS SURVEY

The SuperCluster Assisted Shear Survey (SuperCLASS) is a programme using the *e*-MERLIN interferometric array. It was awarded more than 800 h of observations to perform a deep survey of  $\sim 1\text{ deg}^2$  as part of the legacy program. This is the first paper from the survey giving an overview of the project and presenting the first data release. It is a companion to two papers presenting the first science results and serves as a reference for the other papers coming from the survey. The first data release comprises around half the observations in terms of time but due to the nature of the observing strategy the resulting mosaic, which we will refer to as ‘DR1’ throughout this paper, covers  $\approx 0.26\text{ deg}^2$  to the full depth expected for the survey. The rest of the observations – another  $\sim 400$  h – will bring the survey to full depth over the whole field.

### 2.1 Survey design

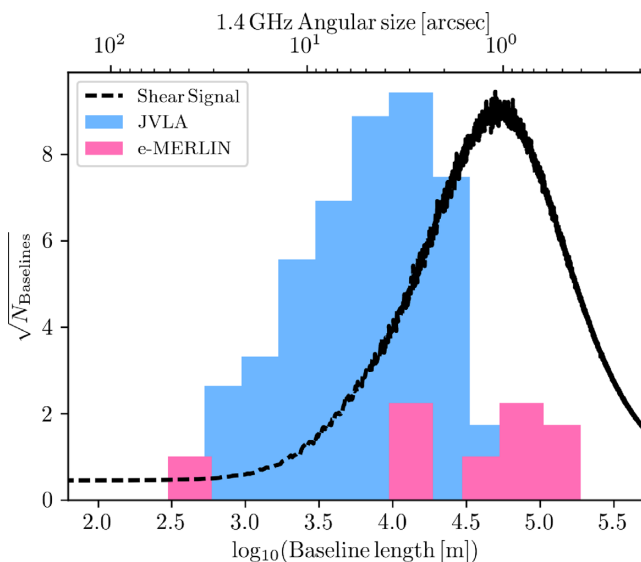
The key criteria used in defining the design of the survey were: (i) to have the most circular beam profile possible for both the *e*-MERLIN and VLA arrays which will allow the best shape measurement and morphological analysis possible; (ii) to allow efficient observing for telescope arrays located at relatively high latitudes in the Northern hemisphere; (iii) have the strongest possible lensing signal over a sufficiently large area to make a meaningful test of shear measurement and to detect the rare counterparts of sub-mm sources. Criteria (i) and (ii) are both met by performing observations at high declinations, while it should be possible to achieve (iii) by observing a region known to have a significant level of large-scale structure, for example, a known supercluster region (Peters et al. 2016). We note that the choice of a supercluster region also facilitates studies of environmental influences on SF/AGN galaxies, as well as in the background population, as in the STAGES project (Heymans et al. 2008; Gray et al. 2009). It was decided that these considerations outweighed the natural desire to choose one of the commonly observed extragalactic fields which have significant multiwavelength coverage and hence we have also embarked on a programme of multiwavelength observations in order to facilitate the science programme. We note that the sensitivity of *e*-MERLIN is  $\sim 30$  per cent higher when observing at high elevation compared to lower declinations necessary to observe more southerly fields.

The choice of depth of the survey and the area covered involves a trade-off between the desire for the survey to be sufficiently deep to detect enough sources to reduce the shot noise in shear measurements and being wide enough to sample the shear correlation mitigating the effects of cosmic/sample variance, and to find rare objects with high SFRs. The overall sensitivity is dictated by the survey speed of *e*-MERLIN which is a sparsely filled interferometer

<sup>1</sup>See <http://www.e-merlin.ac.uk/legacy/projects/> for an overview of these projects

<sup>2</sup><http://www.e-merlin.ac.uk/legacy-emerge.html>





**Figure 1.** An illustration of the complementarity of VLA baselines up to 36 km and *e*-MERLIN baselines up to 217 km. The bars represent the square root of the number of baselines of a given length, whilst the dashed line is the shear signal expected on these Fourier scales when observing at 1.4 GHz. The shear signal is constructed as the difference between sky models of T-RECS sources (see Bonaldi et al. 2019) with and without shape changes due to the simulated effect of weak gravitational lensing expected for a typical supercluster of galaxies.

necessitating significant amounts of integration time. Based on these considerations, and also intending to complement other legacy programs, it was decided to aim for an r.m.s. sensitivity of  $\sim 6 \mu\text{Jy}$  over an area of  $\sim 1 \text{ deg}^2$  which it was calculated would require 832 h of *e*-MERLIN time including that needed for calibration and using the Lovell Telescope (LT) whose 76 m diameter collecting area enables this high level of sensitivity. This should allow high signal-to-noise for detection of sources with flux densities  $S > 40 \mu\text{Jy}$  where we can expect a source density  $\sim 1 \text{ arcmin}^{-2}$ . Unfortunately, the inclusion of the LT brings with it an extra complication in that the primary beam of baselines including the LT is much smaller ( $\sim 10 \text{ arcmin}$ ) than the primary beams of baselines formed from correlating the other 25 m diameter telescopes within *e*-MERLIN, which are typically  $\sim 30 \text{ arcmin}$ .

As with the survey of the HDF-N presented in Muxlow et al. (2005) it is necessary to complement the *e*-MERLIN data with data from the VLA which covers a wider-range of short baselines, but does not have sufficient long baselines to measure ellipticities in relatively small sources. This point is illustrated in Fig. 1 where we have compared the lensing signal, computed by differencing a typical source simulated with and without a shear signal added, to the baseline distribution. Note that the primary beam of the VLA telescopes is similar to the non-LT telescopes of *e*-MERLIN and hence it is possible to construct an observing strategy compatible with both arrays.

## 2.2 Target field selection

In order to search for possible candidate fields, we performed an all sky search for clusters in the NASA Extragalactic Database (NED) with declinations  $> 45^\circ$ ,  $z > 0.15$  and which had previously been studied in order to focus only on actively studied clusters. This list was cross-matched against itself using a matching radius of

0.75 deg. This process turned up a list of five fields containing more than three clusters. We then excluded regions containing strong sources from NVSS (with flux density  $> 100 \text{ mJy}$ ) and where there is a very strong source  $> 1 \text{ Jy}$  within 5 deg. From those which were left, we selected a region at RA  $\sim 10.5 \text{ h}$  and Dec  $\sim 68^\circ \text{ N}$  containing five clusters (A968, A981, A998, A1005, and A1006, see Table 1) which have  $z \approx 0.2$  – we will henceforth call this region the SuperCLASS field, since this appears to be the largest of the currently detected clusters in the field. This region has a typical dust extinction of  $A_V \approx 0.2$  which is within the range that will allow high fidelity optical observations.

The five clusters in the region have been detected by *ROSAT* with luminosities in the range  $(0.3 - 1.7) \times 10^{44} \text{ erg sec}^{-1}$  over the 0.1 to 2.4 keV energy band as tabulated in Table 1. Under the assumption of hydrostatic equilibrium, corresponding masses were calculated in Peters et al. (2016) which suggest that this region of diameter  $\sim 1.5 \text{ deg}$  contains at least  $8 \times 10^{14} M_\odot$  across the five clusters. Using a flat cosmology with Hubble constant  $H_0 = 70 \text{ km sec}^{-1} \text{ Mpc}^{-1}$  and matter density relative to critical  $\Omega_m = 0.3$ , the angular diameter distance to the cluster is  $\approx 680 \text{ Mpc}$  whereas that to typical radio sources with  $z \approx 1$  is  $\approx 1650 \text{ Mpc}$ . Therefore, the region diameter corresponds to 18 Mpc in the supercluster with resolution elements ( $\approx 200 \text{ mas}$ ) corresponding to 0.6 kpc whereas the resolution at typical source redshift is  $\approx 1.6 \text{ kpc}$ .

By choosing a supercluster region it will be possible to study the environmental influences on SF/AGN galaxies as well as in the background population as in the STAGES project (Heymans et al. 2008; Gray et al. 2009).

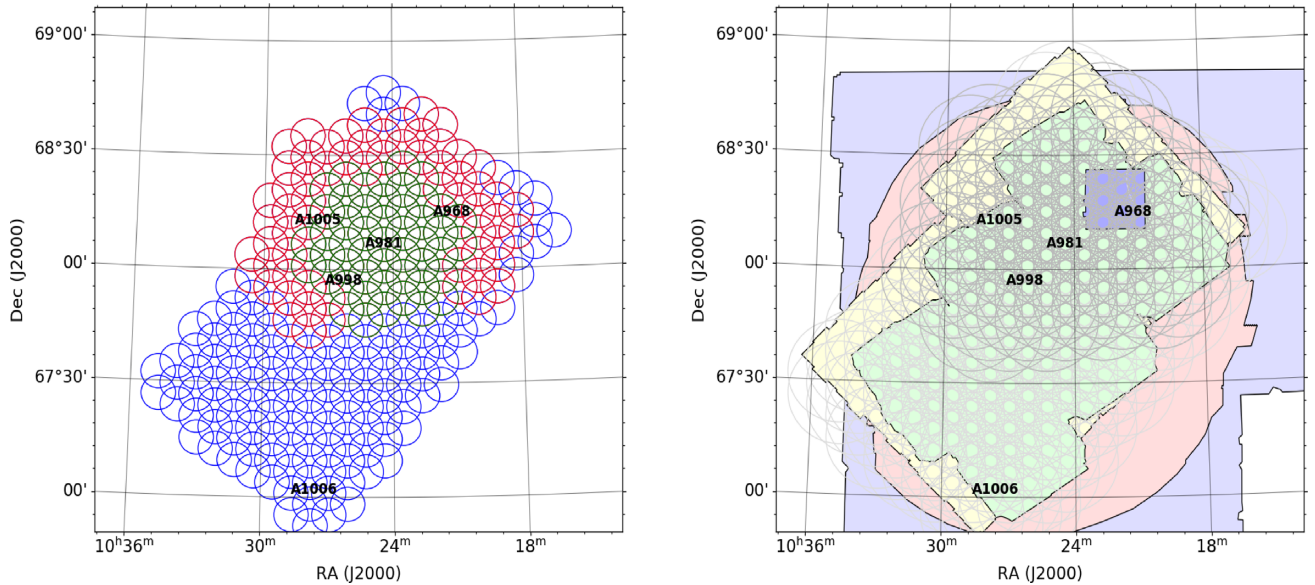
We have already pointed out that the inclusion of the LT complicates our observations since its primary beam is about a factor of 3 smaller than the other telescopes in the *e*-MERLIN array and those which are part of the VLA. In order to cover the SuperCLASS field with approximately uniform noise coverage, it is necessary to use a hexagonally orientated mosaicing strategy whose pointing centres are separated by 5.7 arcmin. The radio pointings observed as part of the project are illustrated in the left-hand panel of Fig. 2 where the coloured circles are indicative of the primary beam of the LT and the grey circles are the equivalent for the smaller *e*-MERLIN telescopes and those in the VLA. The observations of the DR1 region presented in this paper are coloured in green. The red circles indicate the additional SuperCLASS pointings for which the observations are now complete and the data are currently being analysed. The full field covers  $\approx 1 \text{ deg}^2$ , made up of a total of 112 pointings. This region includes four of the clusters (A968, A981, A998, and A1005) which comprise most of the known mass in the region. We have also included in Fig. 2 a proposed extension to the south which would double the area and include the other cluster (A1006). It may be possible to observe this region in a fraction of the time using the phased array feed that will soon be installed on the LT.

## 3 OBSERVATIONS

The field we have chosen to observe with *e*-MERLIN and the VLA is not one of the commonly observed extragalactic fields and therefore, in addition to the radio observations, we have embarked on an extensive programme of optical/NIR observations. These are necessary to obtain photometric redshifts for the detected radio sources. Estimated redshifts are absolutely essential for both of the main science goals of the project: for the lensing observations they are necessary to separate the background and foreground galaxies and to estimate the expected signal from theory, while

**Table 1.** Physical characteristics of the clusters in the SuperCLASS field. The masses are estimated on the basis of hydrostatic mass and a scaling relation between mass and the observed X-ray luminosity  $L_X$  and hence are only indicative as mass measurements. The total mass in the supercluster is estimated to be  $> 8 \times 10^{14} M_\odot$ . For details we refer the reader to Peters et al. (2016) and references therein. Note that Abell 1006 is not in the region of the supercluster field that will be observed as part of the presently awarded observation time.

Name	RA(B1950)	Dec(B1950)	$z$	$L_X/10^{44} \text{ erg sec}^{-1}$	$M_{500}/10^{14} M_\odot$
Abell 968	$10^{\text{h}}17^{\text{m}}44.1^{\text{s}}$	$+68^\circ36'34''$	0.195	0.4	$1.2 \pm 0.3$
Abell 981	$10^{\text{h}}20^{\text{m}}36.0^{\text{s}}$	$+68^\circ20'06''$	0.201	1.7	$2.7 \pm 0.7$
Abell 998	$10^{\text{h}}22^{\text{m}}47.8^{\text{s}}$	$+68^\circ11'13''$	0.203	0.4	$1.2 \pm 0.3$
Abell 1005	$10^{\text{h}}23^{\text{m}}40.0^{\text{s}}$	$+68^\circ27'18''$	0.200	0.3	$1.0 \pm 0.2$
Abell 1006	$10^{\text{h}}24^{\text{m}}10.7^{\text{s}}$	$+68^\circ17'44''$	0.204	1.3	$2.4 \pm 0.6$



**Figure 2.** Left-hand panel: the radio pointings for the SuperCLASS field and the positions of the five clusters. The size of the small circles represents the primary beam size of the Lovell Telescope ( $\sim 10$  arcmin) in a hexagonal pattern with a throw of 5.7 arcmin. Those in green are the DR1 region discussed in this paper, those in red are the additional SuperCLASS data (which has already been collected), and those in blue are a possible extension. The larger grey circles represent the primary beam of the VLA points (and are indicative of the smaller telescopes in the *e*-MERLIN array) which are  $\sim 30$  arcmin in diameter. Right-hand panel: the coverage of the optical/NIR data with the positions and the radio pointings indicated in grey – dark grey for those which have been observed with the present time allocation and light grey for the possible extension. The green region is the observed coverage using Suprime-Cam on the Subaru Telescope in bands  $BVRiz$  and the yellow region was observed using the same telescope using Hyper Suprime-Cam in the  $Y$ -band. The purple regions show the coverage using WIRCAM at  $K_s$ -band using the CFHT and the pink region is that observed at 3.6 and 4.5  $\mu\text{m}$  using the IRAC instrument on the *Spitzer* Space Telescope. The small purple square represents the region which was not observed by WIRCAM.

in the context of galaxy formation/evolution applications they are essential for estimating the size and luminosity of observed sources.

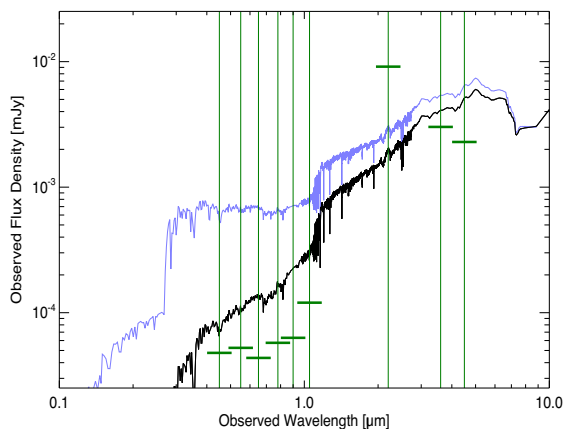
The typical  $r'$ -band magnitude for the sources we expect to detect using radio observations is  $r' \approx 22\text{--}25$  (Cowie et al. 2004). The spectral energy distribution (SED) of a typical DSFG at  $z = 1.5$  is presented in Fig. 3 along with the wavebands we have observed so far and which are discussed in Section 3. These wavebands are the  $BVRiz$  bands observed using the Suprime-Cam on the Subaru Telescope with a goal  $AB$  magnitude limit of 25,  $Y$ -band using Hyper Suprime-Cam again on Subaru to  $AB < 25$ ,  $K_s$  band on the Canada–France–Hawaii Telescope using WIRCAM with a goal of  $AB < 22$ , and observations at wavelengths 3.5 and 4.6  $\mu\text{m}$  using the IRAC instrument on the *Spitzer* Space Telescope. These observing bands straddle the 4000 Å; break in the spectrum over a wide range of redshifts, including that presented in Fig. 3, suggesting that it should be possible to obtain an estimate of the photometric redshift distribution at least in a statistical sense. The footprints of

the optical/NIR observations presented in this paper are shown in the right-hand panel of Fig. 2.

As well as contributing to constraining photometric redshifts, the observations made by *Spitzer* can facilitate other science. First, they will allow estimates of the stellar masses for galaxies at high redshift that should allow the separation of contributions to radio emission due to starburst activity. Moreover, since AGN radio emission emanates from areas  $\ll 1$  kpc, while radio emission from star-forming regions will emanate from the entire galaxy ( $\sim 3\text{--}10$  kpc), it will be possible to use the mid-infrared power law between 3.6 and 4.5  $\mu\text{m}$ , combined with the concentration of the radio emission to separate star-formation from AGN activity.

### 3.1 *e*-MERLIN observations

The *e*-MERLIN observations were performed over the frequency range 1.204 – 1.717 GHz. The bandwidth was divided into eight intermediate frequency (IF) bands, each of 512 channels of width



**Figure 3.** A typical spectral energy distribution for a moderate-age dusty star-forming galaxy at  $z = 1.5$ , without inclusion of nebular line emission. In the blue is the SED that would be observed under the unrealistic assumption that it was not affected by dust obscuration and in black is the same spectrum as in blue, but processed using standard dust attenuation models (Calzetti 2001) combined with synthetic modelling of stellar populations (Bruzual & Charlot 2003). The vertical green lines indicate the positions on the SED of the bands which we have observed. These are  $BVr'i'z'Y$ -bands in the range  $\sim 0.4 - 1.0 \mu\text{m}$  from Subaru,  $K_s \sim 2 \mu\text{m}$  from CFHT and  $3.6/4.5 \mu\text{m}$  from *Spitzer*. The horizontal lines denote the 90 per cent completion depths of the observations presented as part of this first data release which are summarized in Table 2.

125 kHz. This channel width gives a field of view with acceptably low bandwidth smearing of about 30 arcmin, sufficient to map the whole primary beam of the smallest (25 m diameter) telescopes in the *e*-MERLIN array. The integration time used was 1 sec throughout; the choice was again made to maintain acceptably low smearing in the image across 30 arcmin. The data were not averaged further in either time or frequency during the analysis. Observations began on 2014 December 29 and continued with further observing seasons in 2015 and 2016.

Observations were performed in units of 15 h. Each observation consisted of eight passes over seven grid points in a hexagonal configuration. The resulting 56 scans lasted 12.5 min and were interspersed with a 2.5 min observation of a phase calibrator. The phase calibrator used was J1034+6832, a point source of flux density 130 mJy at a distance of about 1 deg selected from the JVAS survey (Patnaik et al. 1992). Each observation was repeated four times, with the starting hour angle of each repetition chosen to ensure as complete  $uv$  coverage as possible in the combined field. Each pointing was therefore observed for 400 min in total. Combination of pointings into mosaics, with the 5.7 arcmin pointing separation, is expected theoretically to increase the signal-to-noise by a factor of 2, and this factor was achieved in practice. The resulting noise levels in the mosaiced images are typically  $7 \mu\text{Jy}$  which is close to the goal sensitivity for the survey.

In addition to the phase calibrator observations every 15 min, the bright flux calibration source 3C 286 was observed once per 15 h observation. This was done to ensure an accurate overall flux density scale (Perley & Butler 2013a). OQ208, a bright point source, was also observed for the purpose of bandpass calibration. Finally, 3C 84, an extremely bright point source (about 20 Jy at  $L$ -band) was observed as an instrumental polarization calibrator, since it has almost zero intrinsic polarization. We defer discussion of the polarization results to a later paper.

Data were reduced in the NRAO AIPS package using a modified version of the pipeline described by Argo (2014). The data were cleaned of Radio Frequency Interference (RFI) using a version of the AOFlogger program (Offringa et al. 2010a,b) whereby binned samples which were above set thresholds in amplitude and noise were flagged. About 15 per cent of the data were typically lost in this way, with the lowest-frequency IF being most affected. Data were then fringe-fitted to remove instrumental delays; although these were usually small, in some epochs delays of 100–200 nsec were seen, which underwent sudden changes during the observations due to reconfigurations within the correlator. Bandpass corrections were then performed, together with an initial phase calibration using the phase calibrator source J1034+6832, and, if necessary, an extra round of self-calibration. Phase corrections were applied to the data and an overall visibility clip at 8 Jy was performed on the target field to exclude aberrant data or data still affected by residual RFI.

Strong sources were imaged and removed from the data by a ‘peeling’ process (Noordam 2004). In the case of moderately strong sources  $< 1 \text{ mJy}$ , the data were phase shifted to the source position, the source was then imaged and the CLEAN components were subtracted from the data. In the case of very strong sources, the source was imaged without phase shifting and the data self-calibrated using the very strong source. This source was then subtracted, and the self-calibration solutions unapplied from the data. About 10 sources needed to be rigorously removed from nearby pointings in order to eliminate their sidelobes from the images of other sources. Data were then exported to FITS files and read into CASA measurement sets and the single pointing imaging was performed using the WSCLEAN algorithm (Offringa et al. 2014). Cotton-Schwab cleaning with  $\text{mgain} = 0.9$  was used with natural weighting for optimal noise and diffuse source sensitivity. Cleaning was stopped when clean components of  $45 \mu\text{Jy}$  were reached. The full mosaic image was compiled by using the AIPS FLAT task to combine the individual pointings using the primary beam correction parameters from the original *e*-MERLIN pipeline.

### 3.2 VLA observations

We used the VLA in its A configuration to observe all 112 pointings over the frequency range 1 – 2 GHz ( $L$ -band). The observations were performed in 2015 August under project code 15A-053. The total observing time was 24 h, divided into  $6 \times 4 \text{ h}$  sessions. These were assigned staggered local sidereal time start times to maximize coverage in the  $uv$  plane. All 112 fields were observed in each session. The correlator was configured to deliver 250 kHz spectral channels (a total of 4000) in full polarization with 1 s time sampling. We observed 3C 286 for the purpose of bandpass, flux density, and position angle calibrations, J0954+7435 for amplitude and phase calibrations, and the unpolarized source J1400+6210 for instrumental leakage calibration. J1400+6210 is less than 0.05 per cent polarized at  $L$ -band.

The data were calibrated using version 4.7.0–1 of the CASA package (McMullin et al. 2007). Hanning smoothing was performed. RFI was identified and flagged using a combination of manual inspection for 3C 286 and automated processing for all other targets using `pieflag` (Hales & Middelberg 2014). The data were flagged to ensure a consistent cross hand phase frame using the task `antintflag` (Hales 2016a). GPS-derived total electron content (TEC) data were obtained from the International GNSS Service and used in combination with the International Geomagnetic Reference Field model of the Earth’s magnetic field to account for atmospheric Faraday rotation when calculating and applying calibration



solutions. Line-of-sight TEC values throughout the observations ranged between 15 to 35 TECU, translating to rotation measures  $< 4 \text{ rad m}^{-2}$ . The flux density scale was referenced to the most recent 2012 value for 3C 286 from Perley & Butler (2013a). The position angle was assumed to be  $-33^\circ$  following Perley & Butler (2013b). Gain solutions were interpolated using `interpgain` (Hales 2016b) to avoid excessive flagging in `applycal` when individual antennas were periodically unavailable due to RFI flagging. Overall, 38 per cent of the target field data were flagged. The quality of the polarization calibration was assessed by imaging the calibrated data for 3C 286 per frequency channel. The recovered spectra for position angle and fractional polarization from each observing session were found to be consistent with Perley & Butler (2013b), modulo a 0.3 per cent increase in fractional polarization due to the higher spatial resolution of our data compared to their D configuration data.

The data were imaged using CASA version 4.7.2. Each of the 112 pointings were imaged individually over a  $1.4 \text{ deg}^2$  field-of-view with  $0.2 \text{ arcsec}$  square pixels and common resolution  $1.9 \text{ arcsec} \times 1.5 \text{ arcsec}$  with position angle  $80^\circ$  east of north. Imaging was performed using a dual Taylor term expansion about a reference frequency of  $1.5 \text{ GHz}$ . Briggs weighting was employed with robust parameter unity to achieve near natural limited sensitivity with an improved, and closer to Gaussian, beam. Widefield corrections were included to account for the non-coplanar nature of the array. A single round of phase self-calibration was applied to each pointing using a manual selection of compact sources. In 49 pointings containing strong extended emission, this required cleaning and subtracting all other sources in the field prior to performing self-calibration using the intended compact sources. This process yielded on-axis r.m.s. noise levels less than  $20 \mu\text{Jy beam}^{-1}$  in all 112 individual pointings. The images were corrected for primary beam attenuation, truncated at the half-power radius, and then linearly mosaiced in slant orthographic (SIN) projection using version 1.5 of the MIRIAD package (Sault, Teuben & Wright 1995). The r.m.s. noise is typically less than  $7 \mu\text{Jy beam}^{-1}$  across the final mosaic, except for small regions around some bright sources (tens of mJy). Polarization imaging and analysis will be described in a later paper.

### 3.3 Subaru optical imaging

Observations of the SuperCLASS field were taken using the Suprime-Cam (SC) and Hyper Suprime-Cam (HSC) instruments on the Subaru 8.4 m Telescope. SC observations covering the  $BVr'i'z'$  filters comprising six fields, each with a field-of-view of  $34 \text{ arcmin} \times 27 \text{ arcmin}$ , were taken over the nights of 2013 February 5 and 6. A significant amount of time on the second night was lost due to poor weather resulting in shorter exposure times in the  $z'$ -band and indeed no coverage for the northern part of the field which has significant overlap with the DR1 region. The total exposure times ranged from 6 min ( $z'$ -band) to 33 min ( $i'$ -band) and the average seeing range from 1.04 to  $1.38 \text{ arcsec}$ . The HSC data with the  $Y$ -band filter were taken on 2015 March 27 in two fields, north and south, whose exposure times were 106 and 96 min respectively and had a seeing of  $1.2 \text{ arcsec}$ .

The individual exposures have been used for optical shape analysis using the best seeing possible as presented in Harrison & The SuperCLASS collaboration (2020) (Paper III) but in order to derive a photometric catalogue for photometric redshift estimation we PSF match all the images removing differences in the seeing and exposure time. In order to do this we use the SC image with the worst seeing ( $1.38 \text{ arcsec}$ ) as our reference image, create a kernel

describing the difference in seeing, and then degrade the PSF of the image for each filter to match the reference image. This convolution smooths noise and signal, so that only contrast and not signal-to-noise is lost.

Flux calibration is performed by comparison to standard stars within the fields. Since the field is at high galactic latitude the Guide Star Catalogue (GSC) is incomplete and therefore we have used a sample of stars which are unsaturated in both the SuperCLASS observations and those from the Data Release 1 (DR1) of the Panoramic Survey Telescope and Rapid Response System (PanSTARRS)  $3\pi$  survey (Chambers et al. 2016) which have similar filter bands. First, stars are separated by comparing the maximum surface brightness with the total magnitude which will have a constant ratio for point sources (stars). The PanSTARRS magnitudes are then ‘colour-corrected’ using stellar templates from the Stellar Flux Library (Pickles 1998) which are convolved with the Subaru filters. Comparison of the magnitudes from SuperCLASS and PanSTARRS then allows a simple zero-point magnitude calibration of the images.

This process yields a photometric catalogue with  $\sim 6 \times 10^5$  sources in the  $BVr'i'$ -bands, around  $2.9 \times 10^5$  in the  $z'$ -band – which was affected by the bad weather on the second night of observations, and  $1.1 \times 10^5$  in the  $Y$ -band once multiple detections from overlapping fields are removed. The limiting  $AB$  magnitudes for 90 per cent and total numbers of detected sources across the 6 Subaru bands are presented in Table 2. We see that they are close to the goal of  $AB = 25$ .

For the science analyses done as part of the DR1 we have performed cuts on the full photometric sample in order to reduce the number of artefacts. In Paper III we define a subsample – the optical lensing sample – containing 106 934 sources corresponding to a source density of  $\sim 19 \text{ arcmin}^{-2}$  in the  $i$ -band used for the optical lensing analysis presented in that paper.

### 3.4 Canada–France–Hawaii telescope observations

Near-IR  $K$ -band observations of the SuperCLASS field were obtained with WIRCAM on the Canada–France–Hawaii Telescope (PI Chapman). Given the field of view of the instrument ( $20 \text{ arcmin} \times 20 \text{ arcmin}$ ), imaging was mosaiced across the field in a variety of conditions spanning multiple nights in queue mode. Images reach moderate depth ( $AB < 19$ ), but unfortunately a key section of the field (encompassing the DR1 coverage area) lacks coverage. Because of this gap and relatively shallow and non-uniform depth,  $K$ -band is not included in the photometric redshift analysis.

### 3.5 Spitzer IRAC observations

Time was awarded as part of *Spitzer*’s Cycle 12 in program #12074 (PI Casey) to survey the SuperCLASS field. The field was observed for 12.7 h with two pointings using the IRAC instrument in  $M$ -band with central wavelength of  $3.6 \mu\text{m}$  and  $N$ -band at  $4.5 \mu\text{m}$ . Flux calibration was completed using astronomical standard stars in *Spitzer*’s continuous viewing zone and were monitored throughout observations. We found the ‘Post-BCD’ (or post-basic calibrated data) software data products sufficient for our scientific use. Sources were extracted in the *Spitzer* bands using SEXTRACTOR (Bertin & Arnouts 1996), accounting for the large point response function (PRF) of the warm mission,  $1.78 \text{ arcsec}$  for the  $M$ -band and  $1.80 \text{ arcsec}$  for the  $N$ -band. No prior positions were used for *Spitzer* source extraction and detections were treated as independent from the source identified in the optical bands. We obtained 90 per cent

**Table 2.** Summary of the optical/NIR observations discussed in sections 3.3, 3.4, and 3.5. We have included an estimate of the centre wavelength of the filter, the  $AB$  magnitude for 90 per cent completion and the number of sources detected. There is relative uniform coverage across  $BVRi'$ -bands close to the goal  $AB$  magnitude limit of 25.  $z'$ -band observations were curtailed by bad weather. Note that there are significantly fewer sources detected in the  $Y$ -,  $K_s$ -,  $N$ -, and  $M$ -bands.

Telescope/Instrument	Band	Wavelength/ $\mu\text{m}$	90 per cent completion limit	Number of sources
Subaru/SC	$B$	0.45	24.8	376 173
Subaru/SC	$V$	0.55	24.6	376 059
Subaru/SC	$r'$	0.65	24.6	375 978
Subaru/SC	$i'$	0.80	24.5	375 728
Subaru/SC	$z'$	0.90	24.5	193 442
Subaru/HSC	$Y$	1.0	23.8	117 318
CFHT/WIRCAM	$K_s$	2.2	19.0	47 819
<i>Spitzer</i>	$N[3.6]$	3.6	20.0	51 364
<i>Spitzer</i>	$M[4.5]$	4.5	20.1	54 137

completion magnitudes of 19.2 and 19.7 for the  $M$ - and  $N$ -bands respectively, finding  $3.7 \times 10^4$  and  $3.8 \times 10^4$  sources in the two filters respectively covering an area of nearly  $2 \text{ deg}^2$ . Sources are matched to the optical catalogue with a nearest neighbour approach up to a maximum separation of 0.1 arcsec.

### 3.6 Ancillary multiwavelength observations of the SuperCLASS field

We have also observed the SuperCLASS field in a number of additional wavebands whose results will be discussed in other papers.

(i) Observations of the field using the Giant Metrewave Radio Telescope (GMRT) at 325 MHz were reported in Riseley et al. (2016). These achieved an r.m.s. noise level of  $34 \mu\text{Jy beam}^{-1}$  over an area of  $\approx 6.5 \text{ deg}^2$  and detected a total of 3257 sources with flux densities between  $183 \mu\text{Jy}$  and  $1.5 \text{ Jy}$ . There are 454 sources in the SCG325 catalogue that lie within the DR1 field. We will consider matches between these sources and those detected using the  $e$ -MERLIN/VLA observations in the DR1 region in Section 5.

(ii) The SuperCLASS field was observed with the Low-Frequency Array (LOFAR; van Haarlem et al. 2013) High-Band Antennas (HBA) during Cycle 0 on 2013 April 3, using the Dutch array (core and remote stations only, no international stations). The field was covered with a single pointing centred on Abell 998, for a total integration time of approximately 8 h. The frequency range covered is 115–162 MHz. We expect to achieve approximately  $\approx 50$ – $100 \mu\text{Jy beam}^{-1}$  sensitivity with LOFAR (based on similar observations by Williams et al. 2016).

(iii) We have taken  $\approx 50 \text{ h}$  of observations using the SCUBA2 instrument on the James Clerk Maxwell Telescope (JCMT) at a wavelength of  $850 \mu\text{m}$ . These observations will be crucial for identifying which of the many sources we have detected are indeed DSFGs. At present maps have achieved an r.m.s. noise level of  $\approx 3 \text{ mJy}$ .

(iv) We have observed part of the SuperCLASS field with the Arcminute Microkelvin Imager (AMI; Zwart et al. 2008) between 2016 August and May, using both the Large Array (LA) and Small Array (SA) which comprise eight 12.8 m (ten 3.7 m) dishes, respectively. AMI operates at a central frequency of  $\sim 15 \text{ GHz}$ , with an effective bandwidth of  $4 \text{ GHz}$ . With the LA, the primary beam FWHM is  $\sim 5.5 \text{ arcmin}$  and therefore the Northern region of the SuperCLASS field was surveyed using twelve raster mosaics of 20 pointings each.

The SA possesses a larger primary beam ( $\sim 20 \text{ arcmin}$ ) allowing the same survey area to be covered by twelve single pointings. We will use these LA and SA observations jointly to (i) classify the source population at high frequencies and low flux densities, and (ii) search for evidence of the Sunyaev-Zel'dovich (SZ) effect from the constituent clusters of the supercluster. For the AMI-LA data, the typical sensitivity is  $32 \mu\text{Jy beam}^{-1}$ , and the resolution is  $50 \times 32 \text{ arcsec}$ . The LA source catalogue is discussed in Riseley et al. (2018).

## 4 RADIO DATA PRODUCTS

Following the independent imaging of  $e$ -MERLIN and VLA data as described in Section 3 above, we perform source finding on each of these images, validate the catalogues produced by the source finder, and add a number of flags in order to define samples useful for weak lensing.

### 4.1 Definition of DR1 region

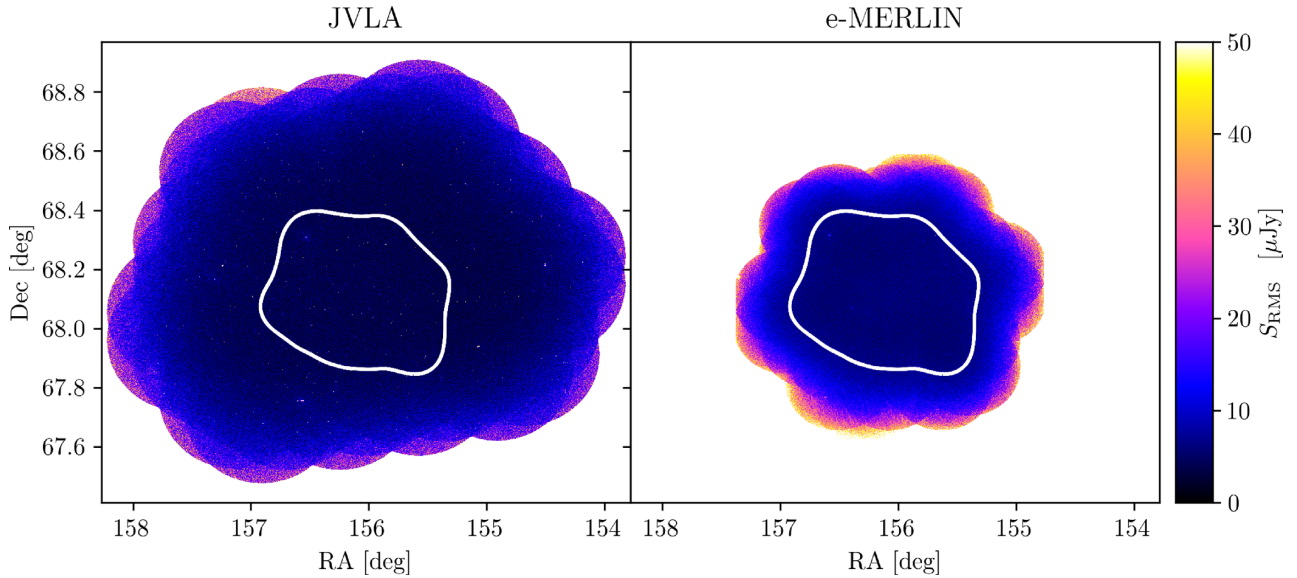
We first define the DR1 region of the SuperCLASS field to be that over which the r.m.s. noise in the mosaiced  $e$ -MERLIN image is approximately constant and below the level of  $7 \mu\text{Jy beam}^{-1}$ . From the image mosaics we create the r.m.s. maps shown in Fig. 4, which show the DR1 region indicated by the solid contour and has an area of  $\approx 0.26 \text{ deg}^2$ .

### 4.2 Released data products

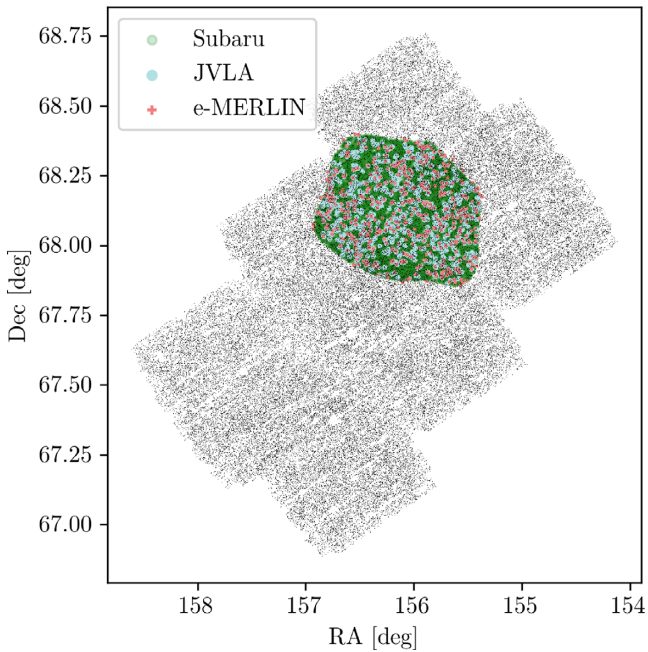
We make catalogues available for the  $e$ -MERLIN and VLA DR1 regions. Both of these catalogues contain the columns from the PYBDSF code as ‘detection’ columns, along with columns marking the sources as resolved and having simple morphology. The VLA catalogue also includes columns relating to shape measurements of the sources as described in Paper III. We also provide the Subaru optical catalogue for the full SuperCLASS region (the black points in Fig. 5). This catalogue contains columns related to the redshift estimation detailed in Paper II and the weak lensing shape measurement detailed in Paper III.<sup>3</sup>

<sup>3</sup><http://cdsarc.u-strasbg.fr/pub/cats/J/MNRAS/XXX/XXXX>





**Figure 4.** Left-hand panel: a map of the r.m.s. noise in the VLA mosaic image for data used for DR1. Right-hand panel: the same for the *e*-MERLIN mosaic image. In each case the solid white line denotes the DR1 region, which is the region with noise  $< 7 \mu\text{Jy}$  in the *e*-MERLIN r.m.s. map.



**Figure 5.** Comparison of the DR1 region (shown by the positions of sources as marked in colours: red for *e*-MERLIN, blue for VLA, and green for Subaru) to the larger full optical region as shown by positions of sources marked by the black points.

#### 4.3 Detection of sources

We carry out source detection using the publicly available PYBDSF code (Mohan & Rafferty 2015). PYBDSF estimates the r.m.s. noise in the image, and finds sources by creating islands of pixels above a specified ‘False Detection Rate’. These islands are then fitted with multiple Gaussians, which may be of different sizes, and grouped into sources. For the *e*-MERLIN image, the initial run of PYBDSF missed a number of resolved, low surface brightness sources which were visible by eye within the image. In order to

account for this we smooth the *e*-MERLIN image to two scales larger than the original 0.4 arcsec beam and run the source finder again (further smoothing with larger scales found no new sources). The union of the catalogues, excluding duplicates, found at the three smoothing scales (0.4, 0.6, 2.0 arcsec) forms our base *e*-MERLIN catalogue. The scale at which each source is found is indicated in the *Scale\_flag* catalogue column. For the VLA image, re-running the source finder at larger smoothing scales was found to produce no new catalogue entries, meaning all sources are from the original mosaic image smoothed with a  $1.9 \times 1.5$  arcsec CLEAN beam, defining the equivalent VLA catalogue.

#### 4.4 Preliminary source classifications

In order to facilitate our science analyses, we add a number of flag columns to the catalogues which pertain to a simple attempt at classification of the sources.

We first classify sources according to the ratio between their integrated and peak fluxes  $S_{\text{int}}/S_{\text{peak}}$ . For unresolved sources unaffected by noise, this should be one. However, noise fluctuations in the image may cause scattering around the one-to-one line. In order to remove sources with unphysical values we find the minimum value of  $S_{\text{int}}/S_{\text{peak}} = r_{\text{min}}$  and remove all sources with  $|1 - S_{\text{int}}/S_{\text{peak}}| < r_{\text{min}}$ . For the *e*-MERLIN catalogue this leads to the removal of 143 sources, but none are excluded from the base VLA catalogue. These sources are not just flagged, but are removed entirely from the catalogues. The number of sources which pass this test are 395 for *e*-MERLIN and 887 for VLA.

For the sources which pass this test, the column *Resolved\_flag* has the value *True* when a source meets all three of the criteria:

$$\begin{aligned} \text{DC\_maj} &> 0, \\ \text{DC\_min} &> 0, \\ \text{Maj} &> \theta_{\text{Maj}}, \end{aligned}$$

where  $\theta_{\text{Maj}}$  is the major axis of the CLEAN beam for the image.

**Table 3.** Numbers of entries in the SuperCLASS DR1 radio catalogues, when sources from flagged columns are successively excluded.

	<i>e</i> -MERLIN	VLA
Full DR1	395	887
Resolved_flag=True	197	789
Simple_morphology_flag=True	144	710

The column `Simple_morphology_flag` has the value `True` for sources which do not show strong visual evidence for having a complex morphology, that is, jet structure in an AGN or a well-resolved irregular star-forming galaxy. We expect sources with a simple morphology to be dominated by star-forming galaxies, and use this set for the shape measurements described in Paper III. In order to create this column we used a private project on the Zooniverse<sup>4</sup> website to provide us with an initial visual classification of identified sources, with the aim of filtering out multisource associations which are likely to be AGN. For each detected source, users were presented with cut-outs of the *e*-MERLIN mosaic image, the VLA mosaic image, and the Subaru *i*-band mosaic image, and asked to classify the source as one of:

- (i) single source (SFG-like);
- (ii) part of multisource association (AGN-like);
- (iii) other / uncertain.

A total of five users performed classifications on all of the available sources, with sources being classed as `Simple_morphology_flag` sources where they were not classified as type (ii) or type (iii) sources by a majority of users.

One of the science goals of the SuperCLASS project is to develop a classification scheme for the detected radio sources based both on morphology and photometric characteristics which is on-going (see Manning & The SuperCLASS Collaboration 2019). For the purposes of this paper we merely attempt a crude preliminary split of the detected radio sources by `Simple_morphology_flag = True`. For weak lensing analyses it is helpful to separate the source population into these two broad categories since they would need to be modelled somewhat differently – note that any future SKA weak lensing survey, being much deeper than any survey done so far, will be dominated by star-forming galaxies with simple morphology (compared to jet-dominated AGN). Therefore, the reader who has a more sophisticated understanding of extragalactic object classification should see this as mainly an exercise to assist the preliminary weak lensing analyses (see Paper III), rather than for it to be used for probing the nature of galaxy evolution. Nonetheless, we would hope that it would not lead to entirely misleading results if used, with caution, for that purpose.

Numbers of sources in the *e*-MERLIN and VLA SuperCLASS DR1 catalogues meeting these criteria are shown in Table 3 and summary plots of the fluxes and sizes are shown in Fig. 6. In addition, we show the behaviour of the peak flux ratios between *e*-MERLIN and VLA in Fig. 7.

From the *e*-MERLIN data we select 56 high-confidence sources, requiring them to have `Resolved_flag = True` in both *e*-MERLIN and VLA catalogues, `Total_flux > 100  $\mu$ Jy` in both *e*-MERLIN and VLA, and `Simple_morphology_flag = True` indicating they have

non-complex morphology (our proxy for being star-forming galaxies). Thumbnails of these sources are shown in Fig. 8, while the associated matching information and Source IDs are provided in Table 4.

#### 4.5 Forced source photometry

As the VLA synthesized beam (the interferometer PSF or ‘dirty beam’) is a better match to the expected source morphology than the *e*-MERLIN one it achieves a better blind detection rate of objects, even though both *e*-MERLIN and VLA surveys have a similar noise level in terms of  $\text{Jy beam}^{-1}$ . For the VLA source positions where no corresponding *e*-MERLIN source was found, a non-blind forced fit can be made in the *e*-MERLIN maps using this extra information of the positions of the VLA sources. All or some of the source flux can be hidden in the *e*-MERLIN images because the surface brightness is close to or below the noise level, but a simple application of aperture photometry should be able to recover the flux. This is also a robust alternative to source fitting for larger diffuse sources based on a model, such as a Sérsic or Gaussian profile, as it is little affected by the source shape.

A simple aperture photometry script was implemented in PYTHON which takes  $256 \times 256$  cut-outs centred on the VLA positions. The image was converted to units of Jy per pixel using the solid angle of the CLEAN beam given in the fits header. To estimate the source flux the signal was integrated within a 1 arcsec radius aperture and subtracted from the median value in an outer annulus of the same area, multiplied by the number of pixels in the central circular aperture. Errors were determined by using the standard deviation of repeated measurements in 99 positions over the cut-out. A stack of the cut-outs is shown in Fig. 9, with the flux of the individually undetected sources now clearly visible and following the expected pattern of the *e*-MERLIN dirty beam PSF.

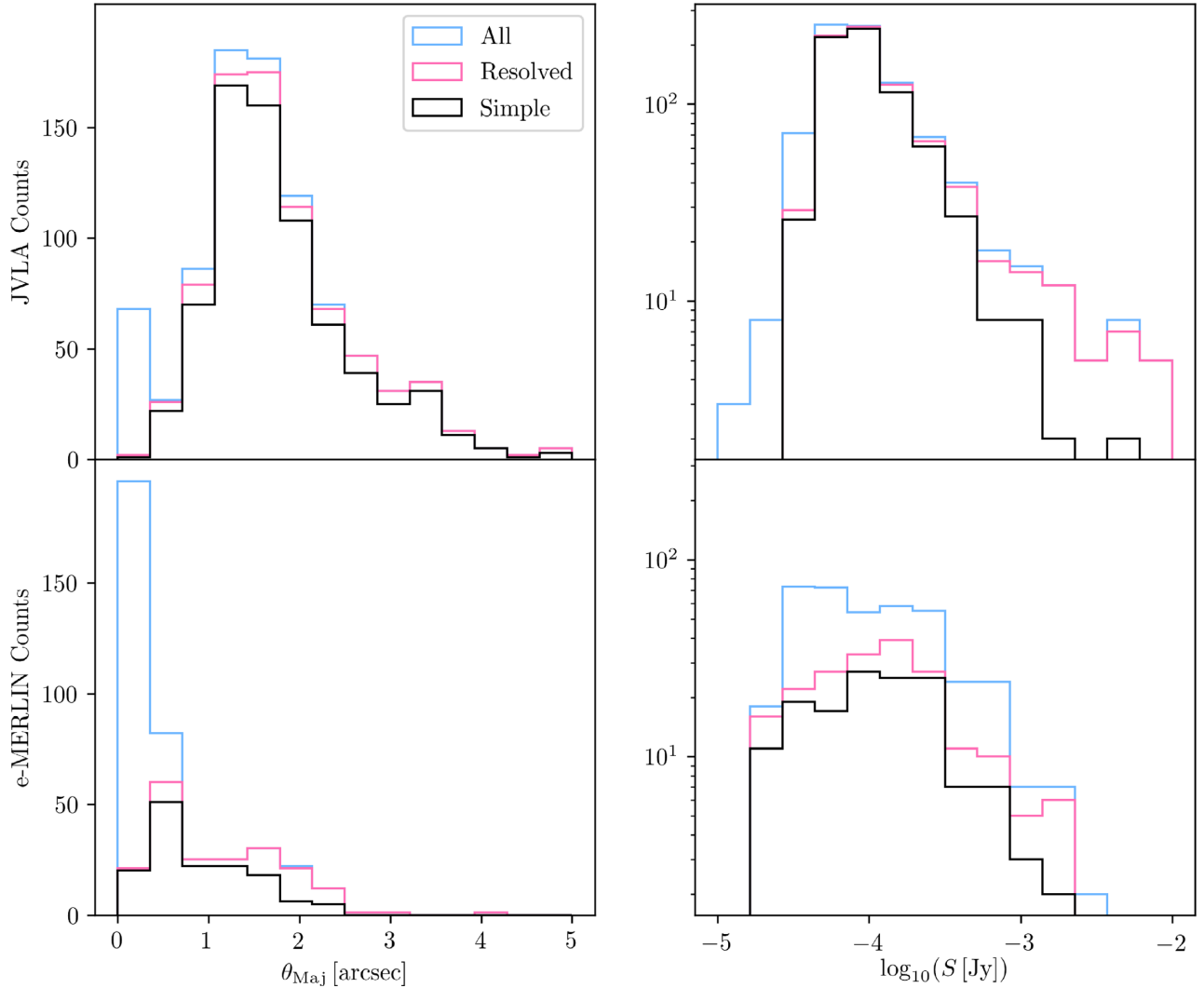
From the PyBDSF catalogues, 246 of the 395 sources detected in the *e*-MERLIN image have matching VLA sources in a 1 arcsec search radius. For the 641 un-matched sources the aperture photometry measurements in the *e*-MERLIN image do in fact show significant positive flux measurements. For the matched sources, a comparison between aperture flux and PyBDSF measurements shows that those with aperture fluxes  $\gtrsim 0.15$  mJy have matches in the modelled catalogue.

As a sanity check the aperture photometry was repeated but using positions offset by 10 arcsec in a random direction from the VLA position. The correlated addition of flux then disappears leaving a symmetric distribution of measured aperture fluxes around zero as expected for just noise. This strongly suggests the bias above zero flux is due to real source flux seen in *e*-MERLIN due to the VLA sources.

#### 4.6 Optical matches

We also cross-match the *e*-MERLIN and VLA catalogues with our Subaru *i*-band weak lensing shape catalogue (as described in Section 3.3), again using a matching radius of 1 arcsec, within the DR1 area. These sources are shown in the context of the full Subaru optical coverage (black points) in Fig. 5. Fig. 10 shows the fraction of radio sources, by flux, which have matches, with this being between  $\sim 1/3$  and  $\sim 2/3$ , dependent on the flux range, for both *e*-MERLIN and VLA sources. This is broadly consistent with the matching fractions of *L*-band sources previously found in studies of the Hubble Deep Field-North by Chapman et al. (2003), Barger, Cowie & Wang (2007).

<sup>4</sup><https://www.zooniverse.org/>



**Figure 6.** Properties of VLA (*upper panels*) and *e*-MERLIN (*lower panels*) detected sources. Left-hand panel: the major axis best-fitting Gaussian FWHM as reported by PYBDSF. Right-hand panel: the flux density of the detected sources. In each case we present three histograms: (i) including all sources (cyan), (ii) including only those sources that are resolved (pink) and (iii) including only those sources that have simple morphologies (black). See the main text for precise definitions of these latter two subsamples. See Section 5.2 for a discussion of the reliability of these size measurements.

## 5 RADIO SOURCE PROPERTIES

### 5.1 Source counts

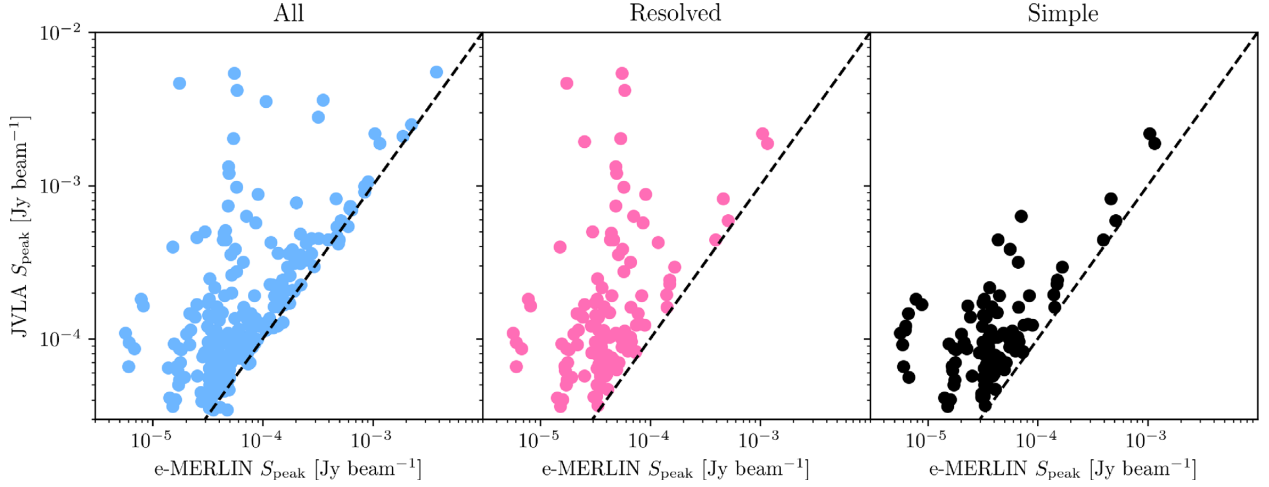
As well as the pure flux counts shown in Fig. 6, we also convert the VLA catalogue source counts to a Euclidean-normalized differential source count in Fig. 11 and compare to counts from the literature. We take differential counts of the sources (all those detected) in logarithmic bins and multiply with the value of the flux bin centre to the power of 5/2 and divide by the area of the DR1 region in Steradians and the difference in bin central flux.

We see from Fig. 11 that our estimates of the differential source count are broadly compatible with those already in the literature over the relevant range of flux density. However, we caution against taking seriously any discrepancies since we have not attempted to model the completeness of our survey in this analysis. For this reason we have not published specific values for the differential source counts.

### 5.2 Source sizes

We also present the source sizes (in terms of the FWHM of the best-fitting Gaussian, including deconvolution of the CLEAN restoring beam, as determined by PYBDSF) for the VLA catalogue in Fig. 6. However, we have reason to believe these measured source sizes are as much a function of the instruments involved as the intrinsic source sizes. Distributions for both telescopes peak around the sizes of the CLEAN restoring beam, even where the ‘resolved’ criteria of Section 4.4 are applied. The procedure of fitting Gaussian profiles to CLEAN image reconstructions – where Gaussians of the size of the restoring beam are added back to the residual image – is not one which can be expected to return correct morphological information for true sources with intensity profiles which are different to that of a Gaussian. Of the two true dirty beam PSFs involved, the VLA PSF is large ( $1.9 \times 1.5$  arcsec) compared to the expected  $\sim 1$  arcsec size of the star-forming galaxy sources (see e.g. Biggs & Ivison 2008), and the *e*-MERLIN PSF, which has significant negative sidelobes,





**Figure 7.** Comparison between the peak *e*-MERLIN fluxes and those from the VLA for the cross-matched detections, those which are resolved in both and those deemed ‘simple’ under our preliminary source classification. Once we concentrate only on sources which are resolved the VLA flux is larger than the *e*-MERLIN flux as one would expect – the longer baselines of *e*-MERLIN resolve out much of the extended flux on  $> 1$  arcsec scales.

will ‘resolve out’ much of the emission from diffuse profiles with shallow wings.

To further quantify this effect we investigate a robust, model-independent quantity related to the extent of the detected sources: the ratio of total measured source fluxes between *e*-MERLIN and VLA images which we refer to as  $R_S$ . This quantity is constructed by measuring the total flux within apertures as described in Section 4.5 at the locations of sources detected in the VLA images. Using the *uv*-coverage of both telescopes, we simulate the flux recovered by each for sources parametrized with a Sérsic intensity profile as a function of radius:

$$I(\theta) = I_{\text{HLR}} \exp \left\{ -b_n \left[ \left( \frac{\theta}{\theta_{\text{HLR}}} \right)^{1/n} - 1 \right] \right\}, \quad (1)$$

where  $\theta_{\text{HLR}}$  is the half-light radius of the profile,  $n$  is the Sérsic index parameter which sets how steeply the intensity profile decreases, and  $b_n \approx 2n - 1/3$ . Fig. 12 shows this quantity from the simple simulations as the solid lines, one for an  $n = 1$  Sérsic profile (corresponding to an exponential profile) and one for an  $n = 4$  Sérsic profile (corresponding to the much ‘peakier’ de Vaucouleurs profile, shown as an example for more compact emission, though we do not expect radio star-forming galaxies to have this profile). The *e*-MERLIN to VLA total flux ratio is shown as a function of the half-light radius of the simulated source. As can be seen,  $R_S$  goes to one as expected for sources with no spatial extent ( $\theta_{\text{HLR}} \rightarrow 0$ ) which are unresolved by both telescopes. However, as the  $\theta_{\text{HLR}}$  of both source profiles is increased  $R_S$  decreases significantly, mostly due to the ‘resolving out’ effect of the interferometers, whereby diffuse flux is not captured due to a lack of smaller spacings between the interferometer elements. For model sources with these sizes and Sérsic indices, this ‘resolving out’ effect is more important for *e*-MERLIN than it is for the VLA.

To confirm these simple simulations we also run our full simulation pipeline (as created for the weak lensing science and described in Paper III), consisting of a source flux-size population model from the T-RECS simulation (Bonaldi et al. 2019), *uv*-data creation and imaging, and measurement of aperture fluxes. We do this with two sky models, one in which all sources are point sources (Fig. 12, left-hand panel) and one in which sources follow the input T-RECS size distribution but have exponential intensity profiles (Fig. 12, right-

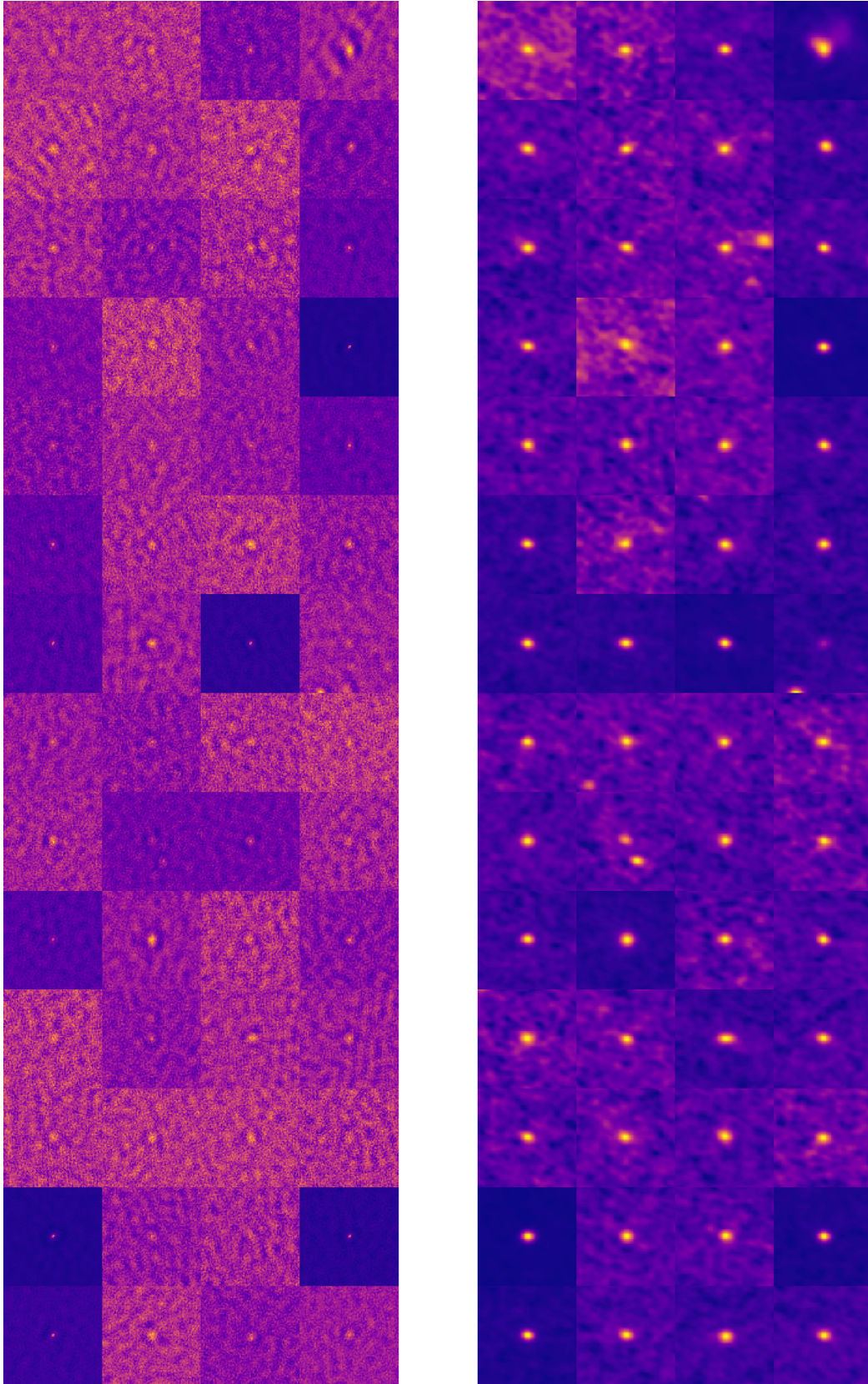
hand panel). Histograms of  $R_S$  measured in these two simulations are shown as the greyscale histogram bars, and the median is shown by the white dashed line. As can be seen, the values of  $R_S$  measured for the exponential sources is consistent with the true distribution of  $n = 1$  profiles with a median half-light radius  $\sim 1$  arcsec, and the point sources (which are more affected by correlated noise fluctuations in the imaging process) are marginally consistent with the expected  $R_S$  of one.

We compare these results to the same quantity  $R_S$  measured in apertures around VLA positions in our DR1 images in Fig. 13, both for sources which have matches between the *e*-MERLIN and VLA images on the left and those which do not on the right. The measured values are again shown as the greyscale histograms with median indicated by the white dashed lines. As can be seen these measurements constrain the degenerate combination of Sérsic index and half-light radius for this population, with the profiles consistent with the  $\sim 1$  arcsec exponentials expected for star-forming galaxies at these flux densities, as well as de Vaucouleurs profiles with sizes up to 5 arcsec and exponential sources with smaller sizes between 0.5 and 1 arcsec.

### 5.3 Spectral properties

An initial investigation of the spectral properties of radio sources in this field was performed by Riseley et al. (2016). With the new VLA data presented in this work, we now have two highly sensitive catalogues with which to further investigate the spectral index distribution.

The 325 MHz SCG325 catalogue contains approximately 213 sources in the region covered by our VLA catalogue. Of these, we find 205 GMRT sources with counterparts at 1.5 GHz. Given the large difference in resolution (13 arcsec at 325 MHz compared to  $\approx 2$  arcsec at 1.5 GHz), many GMRT sources become resolved into multiple sources in our VLA catalogue. We summed the flux densities of all VLA sources corresponding to each GMRT source when deriving our spectral index distribution, and applied a correction factor 0.972 to ensure consistent flux density scaling between the VLA catalogue (which uses the flux scale of Perley & Butler 2013a) and the SCG325 catalogue (which uses the Scaife & Heald 2012 scale).

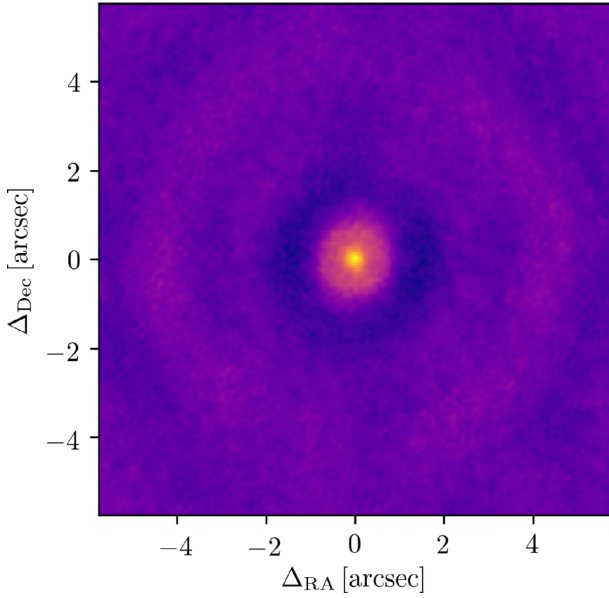


**Figure 8.** The 56 sources with flux  $> 100 \mu\text{Jy}$  which are detected and resolved by both *e*-MERLIN and VLA, showing the data quality available from our observations and the scales probed by the two different observatories. On the *left* are the *e*-MERLIN images and on the *right* are the VLA images. Each thumbnail is 10 arcsec across and colour scales are normalized individually for each one to emphasize the morphology of sources.

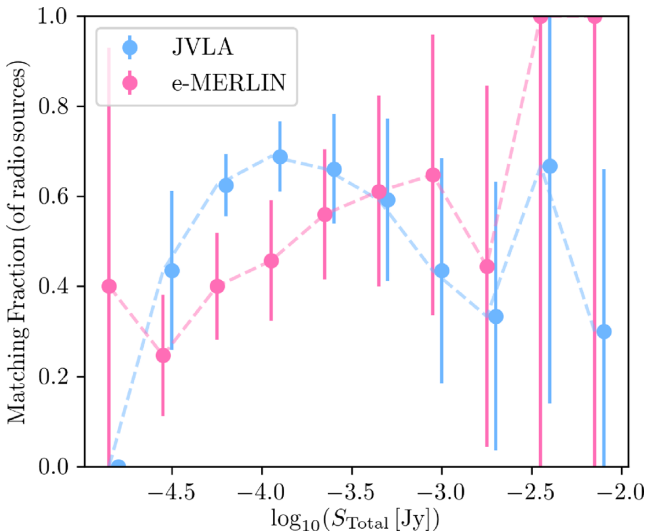


**Table 4.** Fluxes and separations in *e*-MERLIN and VLA images of the 56 sources shown in thumbnails in Fig. 8. The full catalogues containing these sources is available at the CDS VizieR service <http://cdsarc.u-strasbg.fr/pub/cats/J/MNRAS/{vol}/{firstpage}>

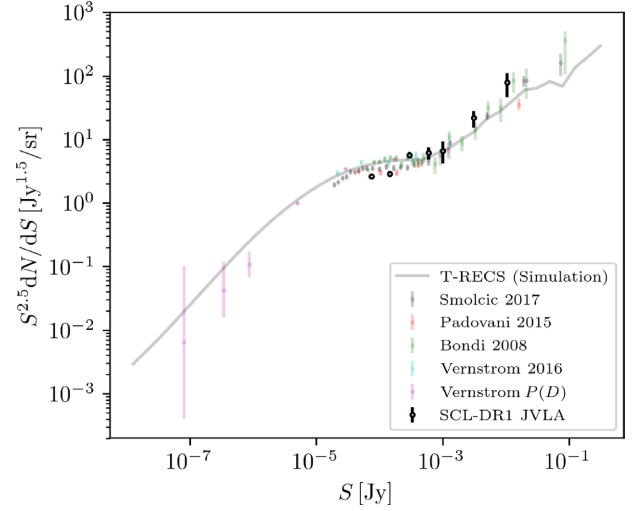
Source ID <i>e</i> -MERLIN	Peak flux <i>e</i> -MERLIN (Jy beam <sup>-1</sup> )	Peak flux VLA (Jy beam <sup>-1</sup> )	Total flux <i>e</i> -MERLIN (Jy)	Total flux VLA [Jy]	Separation (arcsec)
SCL-EM-J102458.43+675300.37	0.05	0.06	0.26	0.10	0.24
SCL-EM-J102416.01+675237.81	0.04	0.07	0.22	0.11	0.15
SCL-EM-J102409.08+675247.19	0.08	0.19	0.18	0.26	0.09
SCL-EM-J102226.97+675447.27	0.07	0.63	1.62	1.83	0.46
SCL-EM-J102545.40+675636.41	0.04	0.11	0.28	0.20	0.04
⋮	⋮	⋮	⋮	⋮	⋮
SCL-EM-J102314.77+682004.35	0.07	0.16	0.22	0.23	0.06



**Figure 9.** *e*-MERLIN image data, stacked at locations where a detection of a source was made in the VLA image, but not the *e*-MERLIN image. Undetected flux from these sources adds coherently with the shape of the dirty beam, showing that information is contained within the *e*-MERLIN data from the sources which were undetected in the *e*-MERLIN-only image.



**Figure 10.** Matching fractions of radio sources in VLA and *e*-MERLIN catalogues which have a Subaru source within a matching radius of 1 arcsec, as a function of their radio flux.



**Figure 11.** Euclidean-normalized source counts for the SuperCLASS VLA DR1 region, shown alongside other *L*-band flux counts from the literature (Bondi et al. 2008; Vernstrom et al. 2014; Padovani et al. 2015; Vernstrom et al. 2016; Smolčić et al. 2017), and the T-RECS simulation source counts at the same frequency.

We present a histogram of the spectral index distribution in Fig. 14. Our distribution appears reasonably well-described by a single Gaussian centred on  $\alpha = -0.62$ , although the population exhibits a relatively high standard deviation of  $\sigma = 0.28$ .

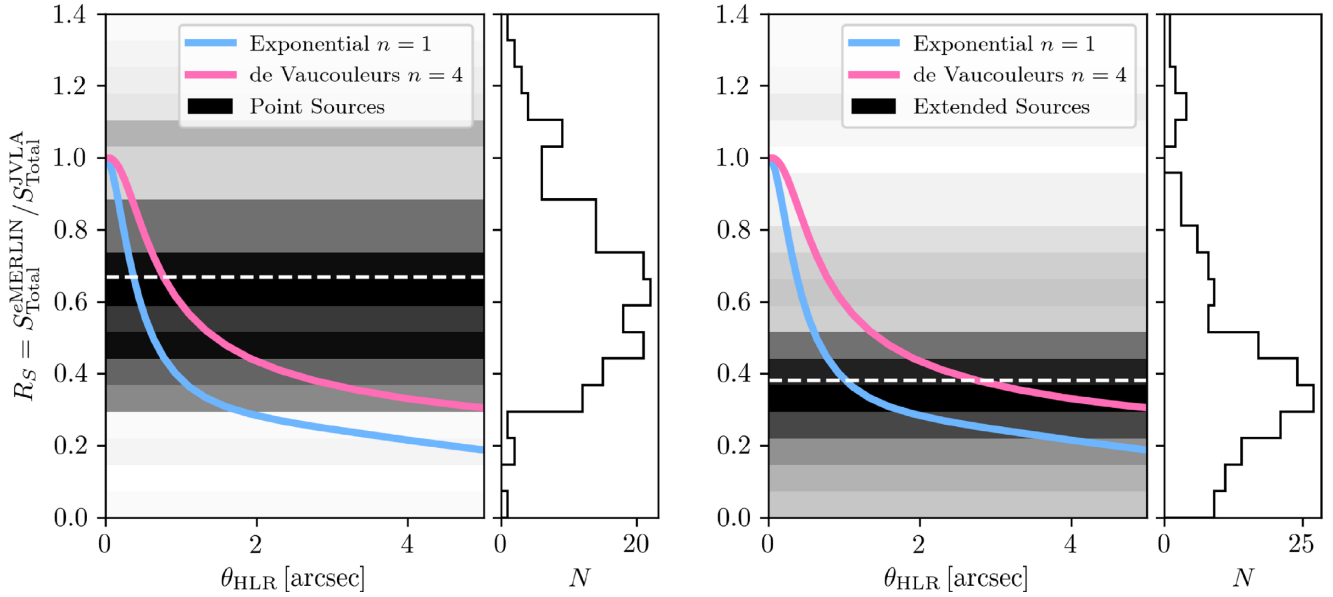
Four sources in our cross-matched catalogue<sup>5</sup> exhibit positive spectra between 325 MHz and 1.5 GHz. All are compact at the resolution of both the GMRT and the VLA.

## 6 SUMMARY AND CONCLUSIONS

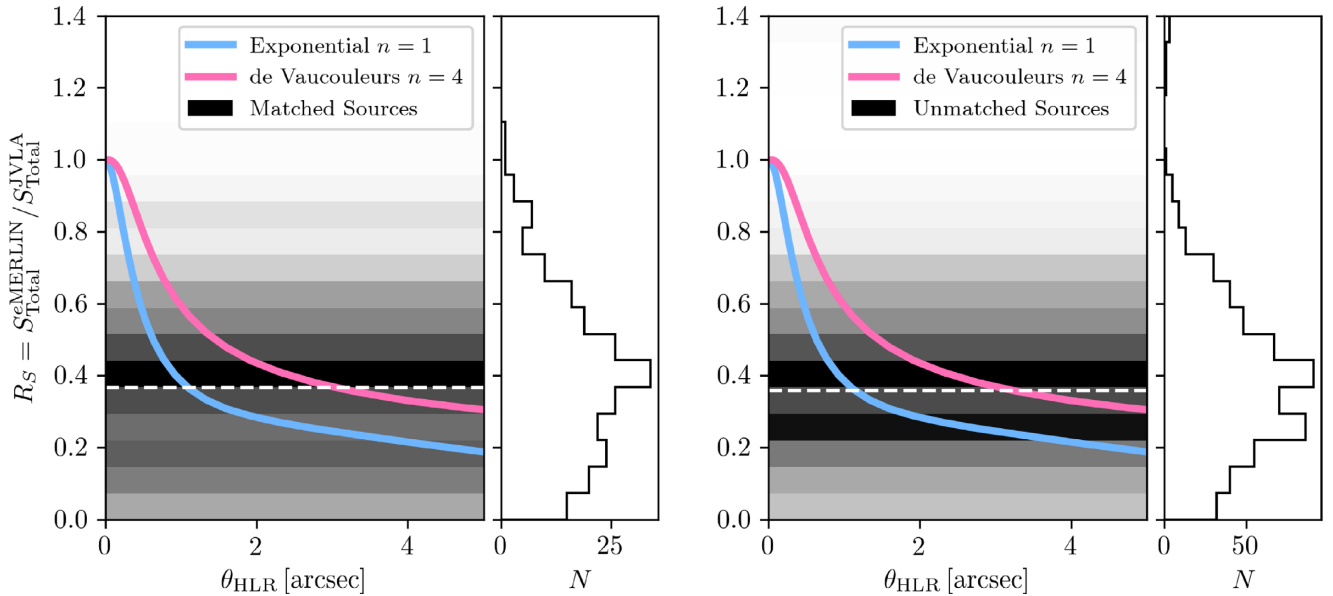
We have presented the Data Release 1 (DR1) data set from the SuperCLASS survey. This data set consists of multiwavelength observations including in the optical, near-infrared, sub-mm, and at high (GHz) and low (MHz) radio frequencies, all covering a  $\sim 1 \text{ deg}^2$  field in the Northern sky containing a massive supercluster at a redshift of  $\sim 0.2$ . The aim of these observations is to detect and characterize, in terms of their spectra and morphology,  $10^4$ – $10^5$  star-forming galaxies covering a redshift range up to  $z \sim 1$ . This will allow us to: use the statistical properties of their observed shapes

<sup>5</sup>These are SCL-JV-J102330.18+681211.47, SCL-JV-J102501.59+681918.63, SCL-JV-J102531.59+680853.13 and SCL-JV-J102612.19+680827.02





**Figure 12.** *e*-MERLIN to VLA flux ratios as measured for simulated point sources (left-hand panel) and extended sources (Exponential profiles, right-hand panel). The red and blue lines represent the expectation for different Sérsic profiles of differing half light radii. The flux ratio constrains the degenerate combination of source size and Sérsic index. The horizontal bars represent histogram counts of measured source flux ratios, with the median shown by the white dashed line. These plots should be compared with those below in Fig. 13, which show the degenerate combination of Sérsic index and half light radii of the observed sources is more consistent with them being extended sources than point sources.

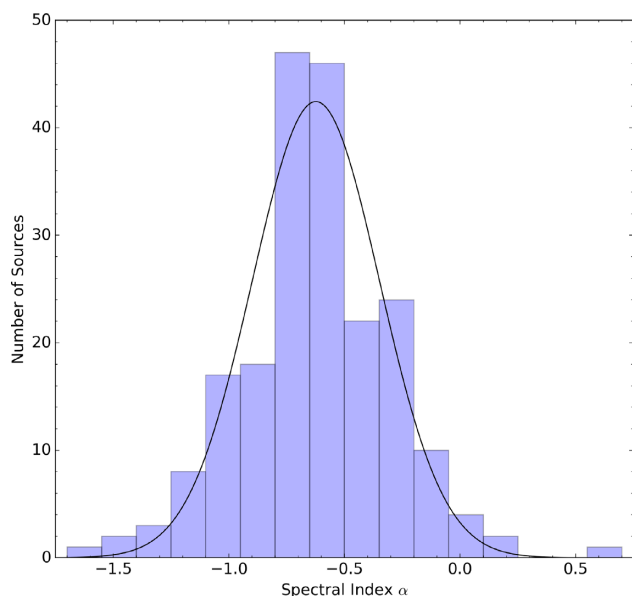


**Figure 13.** Same as Fig. 12 but for the real DR1 catalogue. Left-hand panel: the measured ratios for sources which are detected independently in both the *e*-MERLIN and VLA images. Right-hand panel: measured ratios for those sources that are detected in the VLA image only, with forced photometry at the corresponding location in the *e*-MERLIN image. The measured  $R_S$  for both un-matched and matched sources is consistent with them being extended, rather than point sources but cannot measure independently their Sérsic index or radius.

to measure a weak lensing shear signal from the intervening dark and baryonic mass along the line of sight; and to understand their formation and evolution through access to their star formation and AGN activity unobscured by dust absorption.

In this paper we have described the acquisition, reduction, and initial source catalogue creation of the DR1 data set. With these catalogues in hand we refer the readers to Manning & The SuperCLASS Collaboration (2019) (Paper II) for further details

on studies of the sources' redshifts and matched properties between radio, optical, and near-infrared observations. Similarly, Harrison & The SuperCLASS collaboration (2020) (Paper III) describes the use of the source catalogues and images to measure simple morphological properties – the ellipticities – of the sources in order to infer the gravitational lensing effect of the supercluster mass, both in the optical and radio data alone and together in cross-correlation.



**Figure 14.** Histogram of the spectral index ( $\alpha$ ) distribution for the 205 sources matched between the SCG325 catalogue and our VLA catalogue. The curve denotes the best-fitting Gaussian, centred on  $\alpha = -0.62$ , with  $\sigma = 0.28$ .

We have shown that the flux distribution of our source catalogues, constructed using the VLA mosaic image as a detection image, is consistent with previous studies at similar  $\mu\text{Jy}$  depths at 1.4 GHz radio  $L$ -band. We have presented catalogues of the properties of these sources as measured by the PyBDSF source finding algorithm. Cross-matching the positions of these sources (and those of sources detected in the  $e$ -MERLIN image) with the positions of sources detected in the all-band Subaru HSC image yielded matching fractions between 1/3 and 2/3 of the radio sources, again consistent with expectations of previous studies. Paper II uses this data to fit SED templates to these matched sources and classify them according to AGN and star forming activity. We have also measured spectral indices between the radio  $L$ -band sources and their matches in the lower radio frequency catalogue from our GMRT data, again finding good agreement with the literature.

As well as the fluxes of the detected sources, we have also conducted preliminary investigations into their morphology, in particular constraining the  $e$ -MERLIN to VLA total flux ratio  $R_S$  of the sources. We have argued that this quantity is more robust to the observational effects of radio interferometers than fitting parametrized Gaussian profiles, in particular where lack of short spacings ‘resolves out’ diffuse flux. By comparing with simulations we have used  $R_S$  to constrain the degenerate combination of source Sérsic index and size across the population of sources in our  $e$ -MERLIN and VLA catalogues, finding consistency with (but not strong constraints on) the sizes of  $\sim 1$  arcsec with an exponential profile slope. Better constraints on the source morphology are a key goal of the final data release of SuperCLASS, and we intend to enable this through combination of the data from both  $e$ -MERLIN and VLA together in the native  $uv$  plane (see the extended discussion in Section 6 of Paper III).

We have not discussed the redshift distribution of sources here, instead deferring to the detailed study of Paper II, which finds a prominent over density of sources at  $z \sim 0.2$  as expected from historical observations of Abell clusters in the region.

The data and catalogues presented here will be extended in the full SuperCLASS data release to include approximately three times the area of  $L$ -band radio imaging, greatly increasing the statistical power available to us in constraining both the weak lensing signal and galaxy formation and evolution histories. This data set will be unique in the pre-SKA era in allowing access to small angular scales  $\lesssim 1$  arcsec through its inclusion of data from  $e$ -MERLIN. We therefore expect it to be highly valuable in studying the source populations for large scale cosmology and galaxy evolution surveys with SKA in the next decade. Surveys with the full SKA will be large enough to be competitive with the premier surveys in other wavelengths (e.g. LSST and the *Euclid* satellite in the optical and near-infrared), but maximizing their potential will require designing surveys and observational techniques informed by real data as well as theoretical models. SuperCLASS aims to provide that data to enable this important future science, allowing radio surveys to fulfil their potential in working alongside optical surveys to mitigate each other’s observational systematics. Such robustness to systematics is essential in order to converge on the correct model of the Universe, and to understand how structures within it form on both large (cosmological) and small (galaxy) scales.

## ACKNOWLEDGEMENTS

This publication uses data generated via the Zooniverse.org platform, development of which is funded by generous support, including a Global Impact Award from Google, and by a grant from the Alfred P. Sloan Foundation.

IH, MLB, SC, and BT acknowledge the support of an ERC Starting Grant (grant no. 280127). TH is supported by a Science and Technology Facilities Council studentship. IH acknowledges support from the European Research Council in the form of a Consolidator Grant with number 681431, and from the Beecroft Trust.

CAH acknowledges financial support from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 705332.

CMC thanks the National Science Foundation for support through grants AST-1714528 and AST-1814034, and additionally CMC and SMM thank the University of Texas at Austin College of Natural Sciences for support. In addition, CMC acknowledges support from the Research Corporation for Science Advancement from a 2019 Cottrell Scholar Award sponsored by IF/THEN, an initiative of Lyda Hill Philanthropies. SMM also thanks the NSF for a Graduate Student Research Fellowship Program (GRFP) Award number DGE-1610403.

The authors also acknowledge the contribution of Paul Howson to the  $K$ -band CFHT data reduction.

This research has made use of the VizieR catalogue access tool, CDS, Strasbourg, France (DOI: 10.26093/cds/vizieR). The original description of the VizieR service was published in 2000, A&AS 143, 23.

## REFERENCES

- Argo M., 2014, *e*-MERLIN data reduction pipeline. Astrophysics Source Code Library.
- Bacon D. et al., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14). Proceedings of Science, p. 145
- Bacon D. J., Refregier A. R., Ellis R. S., 2000, *MNRAS*, 318, 625
- Barger A. J., Cowie L. L., Wang W.-H., 2007, *ApJ*, 654, 764
- Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291

- Becker R. H., White R. L., Helfand D. J., 1995, *ApJ*, 450, 559
- Bertin E., Arnouts S., 1996, *A&AS*, 117, 393
- Biggs A. D., Ivison R. J., 2008, *MNRAS*, 385, 893
- Bonaldi A., Harrison I., Camera S., Brown M. L., 2016, *MNRAS*, 463, 3686
- Bonaldi A., Bonato M., Galluzzi V., Harrison I., Massardi M., Kay S., De Zotti G., Brown M. L., 2019, *MNRAS*, 482, 2
- Bondi M., Ciliegi P., Schinnerer E., Smolčić V., Jahnke K., Carilli C., Zamorani G., 2008, *ApJ*, 681, 1129
- Bonzini M., Padovani P., Mainieri V., Kellermann K. I., Miller N., Rosati P., Tozzi P., Vattakunnel S., 2013, *MNRAS*, 436, 3759
- Braun R., Bourke T., Green J. A., Keane E., Wagg J., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14). Proceedings of Science, p. 174
- Brown M. L., Battye R. A., 2011a, *ApJ*, 735, L23
- Brown M. L., Battye R. A., 2011b, *MNRAS*, 410, 2057
- Brown M. et al., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14). Proceedings of Science, p. 23
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Calzetti D., 2001, *PASP*, 113, 1449
- Camera S. et al., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14). Proceedings of Science, p. 25
- Camera S., Harrison I., Bonaldi A., Brown M. L., 2017, *MNRAS*, 464, 4747
- Casey C. M. et al., 2009a, *MNRAS*, 399, 121
- Casey C. M., Chapman S. C., Muxlow T. W. B., Beswick R. J., Alexander D. M., Conselice C. J., 2009b, *MNRAS*, 395, 1249
- Casey C. M., Narayanan D., Cooray A., 2014, *Phys. Rep.*, 541, 45
- Chambers K. C. et al., 2016, preprint ([arXiv:1612.05560](https://arxiv.org/abs/1612.05560))
- Chang T.-C., Refregier A., Helfand D. J., 2004, *ApJ*, 617, 794
- Chapman S. C. et al., 2003, *ApJ*, 585, 57
- Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, *ApJ*, 622, 772
- Clifton T., Ferreira P. G., Padilla A., Skordis C., 2012, *Phys. Rep.*, 513, 1
- 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
- Copeland E. J., Sami M., Tsujikawa S., 2006, *Int. J. Mod. Phys. D*, 15, 1753
- Cowie L. L., Barger A. J., Fomalont E. B., Capak P., 2004, *ApJ*, 603, L69
- Demetroullas C., Brown M. L., 2016, *MNRAS*, 456, 3100
- Demetroullas C., Brown M. L., 2018, *MNRAS*, 473, 937
- Gray M. E. et al., 2009, *MNRAS*, 393, 1275
- Hales C. A., 2016a, antinflag v1.0
- Hales C. A., 2016b, interpgain v1.1
- Hales C. A., Middelberg E., 2014, pieflag: CASA task to efficiently flag bad data. Astrophysics Source Code Library
- Harrison I., The SuperCLASS collaboration, 2020, *MNRAS*, 495, 1737, (Paper III)
- Harrison I., Camera S., Zuntz J., Brown M. L., 2016, *MNRAS*, 463, 3674
- Harrison I., Lochner M., Brown M. L., 2017, preprint ([arXiv:1704.08278](https://arxiv.org/abs/1704.08278))
- Heymans C. et al., 2008, *MNRAS*, 385, 1431
- Hikage C. et al., 2019, *PASJ*, 71, 43
- Hildebrandt H. et al., 2017, *MNRAS*, 465, 1454
- Hillier T., Brown M. L., Harrison I., Whittaker L., 2019, *MNRAS*, 488, 5420
- Ibar E., Ivison R. J., Biggs A. D., Lal D. V., Best P. N., Green D. A., 2009, *MNRAS*, 397, 281
- Jarvis M., Bacon D., Blake C., Brown M., Lindsay S., Raccanelli A., Santos M., Schwarz D. J., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14). Proceedings of Science, p. 18
- Jarvis M. J. et al., 2016, in MeerKAT Science: On the Pathway to the SKA, PoS, p. 6
- Kaiser N., Wilson G., Luppino G. A., 2000, preprint (astro-ph/0003338)
- Laureijs R. et al., 2011, preprint ([arXiv:1110.3193](https://arxiv.org/abs/1110.3193))
- Mainieri V. et al., 2008, *ApJS*, 179, 95
- Manning S. M., The SuperCLASS Collaboration, 2020, *MNRAS*, 495, 1724, (Paper II)
- Massey R. et al., 2013, *MNRAS*, 429, 661
- McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, ASP Conf. Ser. Vol. 376, Astronomical Data Analysis Software and Systems XVI. Astron. Soc. Pac., San Francisco, p. 127
- Mohan N., Rafferty D., 2015, PyBDSF: Python Blob Detection and Source Finder. Astrophysics Source Code Library
- Mould J., Blandford R., Villumsen J., Brainerd T., Smail I., Small T., Kells W., 1994, *MNRAS*, 271, 31
- Muxlow T. W. B. et al., 2005, *MNRAS*, 358, 1159
- Noordam J. E., 2004, in Oschmann J. M., Jr, ed., Proc. SPIE Conf. Ser., Vol. 5489, Ground-based Telescopes. SPIE, Bellingham, p. 817
- Norris R. P. et al., 2011, *Publ. Astron. Soc. Aust.*, 28, 215
- Offringa A. R., de Bruyn A. G., Biehl M., Zaroubi S., Bernardi G., Pandey V. N., 2010a, *MNRAS*, 405, 155
- Offringa A. R., de Bruyn A. G., Zaroubi S., Biehl M., 2010b, in RFI Mitigation Workshop, PoS, p. 36
- Offringa A. R. et al., 2014, *MNRAS*, 444, 606
- Owen F. N., Morrison G. E., 2008, *AJ*, 136, 1889
- Padovani P., Mainieri V., Tozzi P., Kellermann K. I., Fomalont E. B., Miller N., Rosati P., Shaver P., 2009, *ApJ*, 694, 235
- Padovani P., Bonzini M., Kellermann K. I., Miller N., Mainieri V., Tozzi P., 2015, *MNRAS*, 452, 1263
- Patel P., Bacon D. J., Beswick R. J., Muxlow T. W. B., Hoyle B., 2010, *MNRAS*, 401, 2572
- Patnaik A. R., Browne I. W. A., Wilkinson P. N., Wrobel J. M., 1992, *MNRAS*, 254, 655
- Perley R. A., Butler B. J., 2013a, *ApJS*, 204, 19
- Perley R. A., Butler B. J., 2013b, *ApJS*, 206, 16
- Peters A., Brown M. L., Kay S. T., Barnes D. J., 2016, *MNRAS*, 474, 3173
- Pickles A. J., 1998, *PASP*, 110, 863
- Richards E. A., Fomalont E. B., Kellermann K. I., Windhorst R. A., Partridge R. B., Cowie L. L., Barger A. J., 1999, *ApJ*, 526, L73
- Riseley C. J. et al., 2016, *MNRAS*, 462, 917
- Riseley C. J. et al., 2018, *MNRAS*, 474, 5598
- Rivi M., Miller L., 2018, *MNRAS*, 476, 2053
- Rivi M., Miller L., Makhathini S., Abdalla F. B., 2016, *MNRAS*, 463, 1881
- Rivi M., Lochner M., Balan S. T., Harrison I., Abdalla F. B., 2019, *MNRAS*, 482, 1096
- Sault R. J., Teuben P. J., Wright M. C. H., 1995, in Shaw R. A., Payne H. E., Hayes J. J. E., eds, ASP Conf. Ser. Vol. 77, Astronomical Data Analysis Software and Systems IV. Astron. Soc. Pac., San Francisco, p. 433
- Scaife A. M. M., Heald G. H., 2012, *MNRAS*, 423, L30
- Schinnerer E. et al., 2010, *ApJS*, 188, 384
- Smail I., Ivison R. J., Blain A. W., 1997, *ApJ*, 490, L5
- Smolčić V. et al., 2017, *A&A*, 602, A2
- Square Kilometre Array Cosmology Science Working Group et al., 2018, Publications of the Astronomical Society of Australia, 37, article id. e007,
- The LSST Dark Energy Science Collaboration et al., 2018, preprint ([arXiv:1809.01669](https://arxiv.org/abs/1809.01669))
- Thomasson P., 1986, *Q. J. R. Astron. Soc.*, 27, 413
- Thomson A. P. et al., 2019, *ApJ*, 883, 204
- Troxel M. A. et al., 2018, *Phys. Rev. D*, 98, 043528
- Tunbridge B., Harrison I., Brown M. L., 2016, *MNRAS*, 463, 3339
- Tyson J. A., Valdes F., Wenk R. A., 1990, *ApJ*, 349, L1
- van Haarlem M. P. et al., 2013, *A&A*, 556, A2
- Van Waerbeke L. et al., 2000, *A&A*, 358, 30
- Vernstrom T. et al., 2014, *MNRAS*, 440, 2791
- Vernstrom T., Scott D., Wall J. V., Condon J. J., Cotton W. D., Kellermann K. I., Perley R. A., 2016, *MNRAS*, 462, 2934
- Whittaker L., Brown M. L., Battye R. A., 2015, *MNRAS*, 451, 383
- Williams W. L. et al., 2016, *MNRAS*, 460, 2385
- Wittman D. M., Tyson J. A., Kirkman D., Dell'Antonio I., Bernstein G., 2000, *Nature*, 405, 143
- Zwart J. T. L. et al., 2008, *MNRAS*, 391, 1545

<sup>1</sup>Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, UK

<sup>2</sup>Department of Astronomy, University of Texas at Austin, 2515 Speedway Blvd, Stop C1400, TX78712 Austin, TX, USA



<sup>3</sup>*Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*

<sup>4</sup>*Centre for Extragalactic Astronomy, Durham University, South Road, Durham DH1 3LE, UK*

<sup>5</sup>*National Radio Astronomy Observatory, PO Box 0, Socorro, NM 87801, USA*

<sup>6</sup>*School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne NE1 7RU, UK*

<sup>7</sup>*Dipartimento di Fisica e Astronomia, Università degli Studi di Bologna, via P. Gobetti 93/2, I-40129 Bologna, Italy*

<sup>8</sup>*INAF – Istituto di Radioastronomia, via P. Gobetti 101, I-40129 Bologna, Italy*

<sup>9</sup>*CSIRO Astronomy and Space Science, PO Box 1130, Bentley, WA 6102, Australia*

<sup>10</sup>*Department of Physics and Astronomy, University College London, Gower Place, London WC1E 6BT, UK*

<sup>11</sup>*H.H. Willis Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, UK*

<sup>12</sup>*Cyprus University of Technology, Archiepiskopou Kyprianou 30, Limassol 3036, Cyprus*

<sup>13</sup>*Department of Physics and Atmospheric Science, Dalhousie, Halifax, NS B3H 4R2, Canada*

<sup>14</sup>*SKA Organization, Lower Withington, Macclesfield, Cheshire SK11 9DL, UK*

<sup>15</sup>*Dipartimento di Fisica, Università degli Studi di Torino, Via P. Giuria 1, I-10125 Torino, Italy*

<sup>16</sup>*INFN – Istituto Nazionale di Fisica Nucleare, Sezione di Torino, Via P. Giuria 1, I-10125 Torino, Italy*

<sup>17</sup>*Institute for Astronomy, 2680 Woodlawn Drive, University of Hawaii, Honolulu, HI 96822, USA*

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.