Research Article

EMU/GAMA: A new approach to characterising radio luminosity functions

Jahang Prathap[1,2], Andrew Hopkins[1,2], José Afonso[3,4], Maciej Bilicki[5], Michael Cowley[6,7], Scott Croom[8], Yjan Gordon[9], Steven Phillipps[10], Elaine Sadler[8,11], Stanislav Shabala[12], Ummee Tania Ahmed[2,13,14], Stergios Amarantidis[15], Michael Brown[16], Rodrigo Carvajal[3,4], Denis Leahy[17], Joshua Marvil[18], Tamal Mukherjee[1,2], Jayde Willingham[1,2], and Tayyaba Zafar[1,2]

[1]School of Mathematical and Physical Sciences, Macquarie University, Sydney, NSW, Australia, [2]Macquarie University Astrophysics and Space Technologies Research Centre, Sydney, NSW, Australia, [3]Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, OAL, Lisbon, Portugal, [4]Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal, [5]Centre for Theoretical Physics, Polish Academy of Sciences, Warsaw, Poland, [6]School of Chemistry and Physics, Queensland University of Technology, Brisbane, QLD, Australia, [7]Centre for Astrophysics, University of Southern Queensland, Toowoomba, QLD, Australia, [8]Sydney Institute for Astronomy (SIfA), School of Physics, A28, The University of Sydney, Camperdown, NSW, Australia, [9]Physics Department, University of Wisconsin-Madison, Madison, WI, USA, [10]School of Physics, University of Bristol, Bristol, UK, [11]CSIRO Space and Astronomy, Epping, NSW, Australia, [12]School of Natural Sciences, University of Tasmania, Hobart, TAS, Australia, [13]Australian Astronomical Optics, Macquarie University, North Ryde, NSW, Australia, [14]Centre for Astrophysics, University of Southern Queensland, Springfield Central, QLD, Australia, [15]Institut de Radioastronomie Millimétrique (IRAM), Granada, Spain, [16]School of Physics, Monash University, Clayton, VIC, Australia, [17]Department of Physics and Astronomy, University of Calgary, Calgary, AB, Canada and [18]National Radio Astronomy Observatory, Socorro, NM, USA

Abstract

This study characterises the radio luminosity functions (RLFs) for star forming galaxies (SFGs) and active galactic nuclei (AGN) using statistical redshift estimation in the absence of comprehensive spectroscopic data. Sensitive radio surveys over large areas detect many sources with faint optical and infrared counterparts, for which redshifts and spectra are unavailable. This challenges our attempt to understand the population of radio sources. Statistical tools are often used to model parameters (such as redshift) as an alternative to observational data. Using the data from GAMA G23 and EMU early science observations, we explore simple statistical techniques to estimate the redshifts in order to measure the RLFs of the G23 radio sources as a whole and for SFGs and AGN separately. Redshifts and AGN/SFG classifications are assigned statistically for those radio sources without spectroscopic data. The calculated RLFs are compared with existing studies, and the results suggest that the RLFs match remarkably well for low redshift galaxies with an optical counterpart. We use a more realistic high redshift distribution to model the redshifts of (most likely) high redshift radio sources and find that the LFs from our approach match well with measured LFs. We also look at strategies to compare the RLFs of radio sources without an optical counterpart to existing studies.

1. Introduction

Mapping the distribution of galaxies is fundamental in understanding the large scale matter distribution of the Universe, for which luminosity functions (LFs) have been instrumental throughout. The galaxy LF is the comoving source density with luminosity and is a statistical quantity used to understand the galaxy distribution. Redshift evolution of the LFs quantifies the cosmic evolution of the statistical properties of galaxy population (e.g. Felten 1977; Gunawardhana et al. 2015; Smolčić et al. 2017; Novak et al. 2017; Peters 2019). The population of star forming galaxies (SFGs) and active galactic nuclei (AGN) exhibit different characteristics as a function of luminosity and redshift.

Consequently, separate analyses are important in disentangling the cosmic evolution of these populations (e.g. Mauch & Sadler 2007; Novak et al. 2017; Smolčić et al. 2017).

To characterise the evolution of these distinct populations, advancements in multiwavelength astronomy have been critical, enabling LFs to be measured across the entire electromagnetic spectrum. Ultraviolet (UV) and optical LFs, owing to a large number of space and ground-based surveys, benefit from having large samples, which results in a better constrained measurement (e.g. Croom et al. 2004; Bouwens et al. 2015). However, both need to account for dust obscuration. The infrared (IR) regime can trace the obscured stellar/AGN light but is not sensitive to emission that freely escapes from the galaxy. Radio emission, on the other hand, is not affected by dust and, hence, is a valuable alternative as a probe of the galaxy population independent of dust obscuration (e.g. Mauch & Sadler 2007; Novak et al. 2017; Smolčić et al. 2017). Luminous radio emission typically arises from AGN, while low-luminosity radio emission is typically from star formation in nearby galaxies (Longair 1966). The faint radio source population

was originally probed through single pointing radio surveys (e.g. Windhorst, van Heerde, & Katgert 1984; Windhorst et al. 1985). In the past few decades, this has been significantly expanded through the advent of deep and wide surveys (Sadler et al. 2002; Mauch & Sadler 2007; Pracy et al. 2016; Peters 2019).

Because of cosmic variance, many factors go into calculating LFs when accounting for the Universe as a whole. No single survey can detect all the sources in a given sky area to provide a complete, unbiased sample. The detection probability of an object is affected by various factors, such as the survey sensitivity, the input catalogue of the survey, its target and spectroscopic completeness, area and depth, and the observing conditions. The intrinsic properties of a source, such as its surface brightness and size, can also influence whether it gets detected. The luminosity of a source, together with these factors, governs its likelihood of detection. Consequently, low-luminosity sources are typically the fewest in any magnitude limited survey (e.g. Geller et al. 2012; Gunawardhana et al. 2015). Although they do not contribute much to the overall luminosity in proportion, faint galaxies greatly outnumber their bright counterparts. Standard methods are often used to account for these biases and incompleteness (Schmidt 1968, also see Section 5.2).

Once the LF is corrected for sample incompleteness and any biases, it is typical to represent the function using an analytical form. A common representation is that of Schechter (1976):

$$\phi(L)\, dL = \frac{\phi_*}{L_*} \left(\frac{L}{L_*} \right)^a e^{-L/L_*}\, dL, \qquad (1)$$

where ϕ_* is the normalisation factor that defines the overall density of galaxies, L_* is the characteristic galaxy luminosity (an L_* galaxy is comparable to the Milky Way, whereas a galaxy of $L < 0.1\,L_*$ is a dwarf), and a defines the faint end slope. This functional form behaves as a power law with slope a for luminosities $L < L_*$ and as an exponential for $L > L_*$. While the Schechter function provides a useful framework for representing LFs across various wavelengths, RLFs offer a unique advantage by being insensitive to dust obscuration.

1.1. Radio luminosity functions

The RLF represents the source density of radio galaxies per unit comoving volume distributed in bins of radio luminosity. The radio source population has been extensively explored using RLFs (e.g. Sadler et al. 2007; Mauch & Sadler 2007; Best & Heckman 2012; Novak et al. 2017; Smolčić et al. 2017; Ceraj et al. 2018; Condon, Matthews, & Broderick 2019; Jose et al. 2024; Wang et al. 2024; Morabito et al. 2025). In addition, RLFs have also been used in probing radio source clustering (Magliocchetti et al. 1998; Overzier et al. 2003). The population of galaxies with radio emission is a mix of sources powered by AGN and SFGs (e.g. Cowley et al. 2016). Classical radio galaxies, typically ellipticals, are either compact or multicomponent (Shabala et al. 2008; Hardcastle et al. 2019), and are driven by AGN (Condon 1989; Condon 1992; Mobasher et al. 1999; Miller et al. 2009). This is usually the case for radio galaxies with $L_{1.4\,GHz} \geq 10^{23}$ W Hz^{-1} as evident from local 1.4 GHz RLFs (Condon, Cotton, & Broderick 2002; Mauch & Sadler 2007; Best & Heckman 2012). Below this threshold, an increasing number of radio-emitting SFGs are found (Condon 1989; Seymour et al. 2008).

For the current generation of deep radio surveys, a low redshift sample contains a significant number of star formers with AGN dominating at higher redshifts (Willott et al. 2001). Consequently, the shape and evolution of the RLF with redshift constrain the nature of activity in massive galaxies and their evolution. Just as UV and optical LFs are commonly parameterised using the Schechter function (Equation 1), IR and radio LFs are commonly parameterised using the Saunders function (Saunders et al. 1990; Mauch & Sadler 2007):

$$\phi(L)dL = \frac{\phi_*}{L_*} \left(\frac{L}{L_*} \right)^a \exp \left[- \left(\frac{\log\,(1 + L/L_*)}{\sqrt{2}\sigma} \right)^2 \right]\, dL. \qquad (2)$$

It behaves in the same fashion as the Schechter function for $L < L_*$ and as a Gaussian (rather than exponential) in $\log L/L_*$ with a width σ for $L > L_*$. Radio AGN, on the other hand, are commonly parameterised using a two power-law function (Dunlop & Peacock 1990; Brown, Webster, & Boyle 2001, also see Section 6).

Longair (1966), one of the first to attempt to determine the evolution of radio source population, found that the models where the most powerful radio sources undergo greater comoving space density evolution than less powerful sources best fitted the data. Evolution is generally incorporated into the estimated RLF by two models: pure density evolution (PDE) and pure luminosity evolution (PLE) (e.g. Pozzetti, Bruzual, & Zamorani 1996; Fontanot et al. 2007; Novak et al. 2017, also see luminosity dependent density evolution (LDDE), Schmidt & Green 1983).

In PDE, the number density of galaxies is free to evolve with redshift, and their luminosities are presumed to remain the same. Physically, this can be achieved through merger events (not generally true for classical radio sources). If $\phi_0(L)$ is the local RLF, PDE can be incorporated into it as:

$$\phi_z(L, z) = (1 + z)^{k_D} \phi_0(L), \qquad (3)$$

where k_D is the free parameter that determines the rate of evolution (Rowan-Robinson et al. 1993; Hopkins 2004; Peters 2019).

PLE, on the other hand, assumes no evolution in the galaxy demography. Rather, it includes changes in galaxy luminosities with redshift. Galaxies become fainter due to star formation quenching, such as through AGN feedback or stellar winds. A sudden influx of cold gas or AGN feedback can boost star formation and increase the luminosity of the galaxy as a whole. The redshift dependence of the LF is governed by the evolution in luminosity parameterised as:

$$\phi_z(L, z) = \phi_0 \left(\frac{L}{(1 + z)^{k_L}} \right), \qquad (4)$$

where k_L is the free parameter that determines the rate of evolution based on the changes in luminosity, as opposed to density in PDE.

1.2. The need for redshift

Spectroscopic redshifts of sources are crucial in measuring the LF and in any analysis requiring AGN classification (e.g. Condon et al. 2002; Sadler et al. 2002; Best et al. 2005; Mauch & Sadler 2007). In the context of radio surveys, however, the tens of millions of radio sources being discovered are rapidly outstripping the capabilities of existing spectroscopic facilities (e.g. Luken et al. 2023).

Photometric template fitting (Baum 1957; Loh & Spillar 1986) has been extensively used in the absence of spectroscopic

redshifts, often approaching comparable accuracy for typical galaxy populations, especially at lower redshifts ($z \lesssim 1$, see e.g. Ilbert et al. 2009; Straatman et al. 2016; Li et al. 2024; Siudek et al. 2024). With the new generation of radio surveys, (e.g. Evolutionary Map of the Universe (EMU), Norris et al. 2011, LOw Frequency ARray (LOFAR) Two-metre Sky Survey (LoTSS), Shimwell et al. 2017, GaLactic and Extragalactic All-sky Murchison Widefield Array survey eXtended (GLEAM-X), Hurley-Walker et al. 2022) the required photometric coverage is now the limiting factor for a large number of radio sources. The difficulty in disentangling star forming and AGN emission as well as the lack of specialised templates modelling AGN further complicates the template fitting approach for radio sources (Salvato, Ilbert, & Hoyle 2019; Luken et al. 2023).

When photometric or spectroscopic redshifts are not feasible for all sources, one can leverage statistical techniques to model redshifts of radio sources informed by available data from similar sources. This is done by looking for correlations between observed properties of the radio source, such as its redshift and flux density, with some underlying assumptions and then extending the relation to the source with unknown redshift.

Similar to its application in various other fields, machine learning (ML) techniques have been used in astronomy for redshift estimation (Smith & Geach 2023; Webb & Goode 2023). For instance, algorithms such as k-nearest neighbours (kNN; Cover & Hart 1967) in Cavuoti et al. (2017) and Luken et al. (2023), random forest (RF; Ho 1995) in Cavuoti et al. (2017) and Carvajal et al. (2021, see also Carvajal et al., in preparation), neural networks (NNs) in Curran (2022) and Luken et al. (2023), and Gaussian processes (GPs) in Duncan et al. (2018a,b, 2021) and Luken et al. (2023) demonstrate the capability of ML techniques in redshift estimation. Karademir et al. (2022) used a different approach in which the evolving galaxy LF was measured with a clustering-based redshift inference technique using spatial cross-correlation statistics.

In this work, we explore simple statistical techniques that draw on minimal available information in order to derive RLFs, anticipating future large datasets where the large numbers of deep multiwavelength photometric measurements needed by other techniques may not be available. Here we implement such an approach to model the redshift distributions of radio sources from EMU early science data (Gürkan et al. 2022) and derive their LFs. This data is cross-matched with the Galaxy and Mass Assembly (GAMA; Driver et al. 2011; Bellstedt et al. 2020) survey data release 4 (DR4; Driver et al. 2022) for spectroscopic redshifts and r-band apparent magnitudes (m_r). Radio sources with spectroscopic counterparts serve as priors for redshift estimation, while the m_r, where available, supplements this estimation by helping to constrain redshifts for sources lacking spectroscopic data.

The structure of the paper is as follows: The data are described in Section 2. Section 3 details the statistical approach in estimating the redshifts of radio sources with and without optical counterparts. The radio sources with spectroscopic redshifts are classified into SFGs and AGN in Section 4. This section also extends the classification to sources without spectroscopic redshifts. The RLFs are derived in Section 5 and the initial results are described in Section 6. Section 7 discusses the main findings. Section 8 summarises the work and concludes the paper. Throughout the paper we assume the Planck Collaboration et al. (2016) cosmology with $H_0 = 67.8 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$, $\Omega_M = 0.308$, and $\Omega_\Lambda = 0.692$.

2. Data

2.1. GAMA

GAMA is a multiband imaging and spectroscopic survey undertaken at the Anglo Australian Telescope (AAT) using the two-degree field (2dF) fibre feed and the AAOmega spectrograph. The survey provides high quality spectroscopy covering five separate fields (G02, G09, G12, G15, and G23) with a limiting magnitude of $m_r < 19.8$ mag ($m_i < 19.2$ mag for the G23 field), collectively spanning an extensive area of 286 square degrees.

Four of the GAMA fields (G09, G12, G15, and G23) benefit from the overlap with the footprints of the European Southern Observatory (ESO) VST Kilo Degree Survey (KiDS, de Jong et al. 2015), ESO Visible and Infrared Survey Telescope for Astronomy (VISTA) Kilo-degree Infrared Galaxy Public Survey (VIKING, Edge et al. 2013), Galaxy Evolution Explorer (GALEX, Martin et al. 2005), Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010), and Herschel (Pilbratt et al. 2010) surveys, ensuring extensive multiwavelength imaging and photometric coverage ranging from FUV to FIR. The GAMA photometric catalogue has a specific flux cut applied on the KiDS r-band magnitude (< 20 mag, Bellstedt et al. 2020), whereas the underlying KiDS and VIKING data go much deeper (Wright et al. 2024). The analysis in this work thus can be extended to include current and upcoming deep photometric data (e.g. LSST Science Collaboration et al. 2017).

The latest data release of GAMA, DR4 (Driver et al. 2022) combines photometric data from the aforementioned surveys together with the spectroscopy from the AAOmega spectrograph. We are interested in the G23 field ($338.1° < \mathrm{RA} < 351.9°$, $-35° < \delta < -30°$) since it overlaps with the EMU early science observation (Gürkan et al. 2022).

In this paper, we use m_r from the gkvScienceCatv02 table (Bellstedt et al. 2020) in the gkvInputCatv02 data management unit (DMU), with 458 844 sources. The SpecLineSFRv05 DMU contains the GaussFitSimplev05 table ((Gordon et al. 2017, 49 623 sources)) from which spectroscopic redshifts are retrieved. We use these two tables to perform multiple crossmatches with the radio sample defined in the next section.

2.2. EMU

The Australian Square Kilometre Array Pathfinder (ASKAP, Johnston et al. 2007; McConnell et al. 2016; Hotan et al. 2021), a precursor of the Square Kilometre Array (SKA), consists of 36 x 12-m antennae each of which is equipped with a phased array feed (PAF) resulting in a field of view of 30 deg^2 and high survey speeds. ASKAP is designed to observe in the 800–1 800 MHz frequency range with an instantaneous bandwidth of 288 MHz and offers rapid observations with good resolution and sensitivity.

EMU (Norris et al. 2011, 2021; Hopkins et al. 2025) is an ongoing wide-field radio continuum survey conducted using ASKAP, providing a deep radio continuum map of the entire southern sky, extending up to $\delta = 0°$ at 943 MHz. The survey achieves a sensitivity of approximately 20 μ Jy beam^{-1} and a resolution of 15″ and is currently around 20% complete in terms of observations. Upon completion, it is anticipated that EMU will detect and catalogue roughly 20 million galaxies, including SFGs up to $z \sim 1$, powerful starbursts reaching even greater redshifts, and AGN to the edge of the visible Universe (Hopkins et al. 2025).

Table 1. The number of sources in different samples. GEM I sources have both spectroscopic redshifts (z), and m_r; GEM II has only m_r; GEM III has neither. EMU represents the entire radio sample.

Sample	Number of sources	Optical data
GEM I	6 425	z, m_r
GEM II	4 566	m_r
GEM III	28 821	—
EMU	39 812	—

The multiwavelength FUV-FIR photometric coverage of the GAMA G23 region is augmented by radio data through the ASKAP commissioning (Leahy et al. 2019) and later as part of the EMU early science program (Gürkan et al. 2022). In this paper, we use the data from the early science program that covers an area of 82.7 deg² and achieves a sensitivity of 0.038 m Jy beam⁻¹ at a frequency of 887.5 MHz. The survey detected 55 247 sources, out of which 39 812 have signal-to-noise (S/N) ratios $\geq$ 5. The brightest radio galaxies are, in general, depending on the angular resolution of the survey, multi component. In the Gürkan et al. (2022) catalogue, however, we have < 1% multi component sources, also their radio luminosities are of the order of single component sources. Because of this, we do not expect any potential biases due to the existence of multi component sources in the sample. We use the parameter Total_flux from Gürkan et al. (2022) corresponding to the measured integrated flux from the radio source at 887.5 MHz. This radio sample is referred to as 'EMU' throughout the paper.

2.3. Final sample

To facilitate our analysis, we first implement two cross-matches between the radio and optical samples using Topcat (Taylor 2005). A cross-match between EMU and the gkvScienceCatv02 table with a cross-match radius of 5″ (Norris et al. 2021; Ahmed et al. 2024) results in 10 991 radio sources with m_r. This table is then cross-matched with the GaussFitSimplev05 spectroscopic data, resulting in 6 425 radio sources with both spectroscopic redshifts and m_r, and 4 566 radio sources with only m_r.

Consequently, we have three sets of catalogues: (*i*) radio sources with m_r and spectroscopic redshifts (6 425 sources), (*ii*) radio sources with m_r but not spectroscopic redshifts (4 566 sources), and (*iii*) radio sources with no optical counterparts (28 821 sources). These samples are referred to as GEM I, GEM II, and GEM III, respectively (for GAMA-EMU, see Table 1). GEM I facilitates modelling the redshifts for both GEM II and GEM III as detailed in Section 3. GEM II has the advantage of having m_r, which further constrains the estimated redshifts. Figure 1 illustrates the radio flux density distributions of these samples and Figure 2 shows the redshift distribution of GEM I sources.

As discussed in Section 1, survey completeness plays a crucial role in LF measurements. Figure 7 in Gürkan et al. (2022) shows the completeness of the EMU sample as a function of source flux density. The EMU sample is more than 95% complete above a flux density of 1 mJy. Since we are estimating the redshifts of the entire EMU sample, the completeness values from Gürkan et al. (2022) can be applied directly to our calculations.

3. Redshift estimation

To accurately construct RLFs for AGN and SFGs, this study applies simple statistical techniques to model redshift distributions
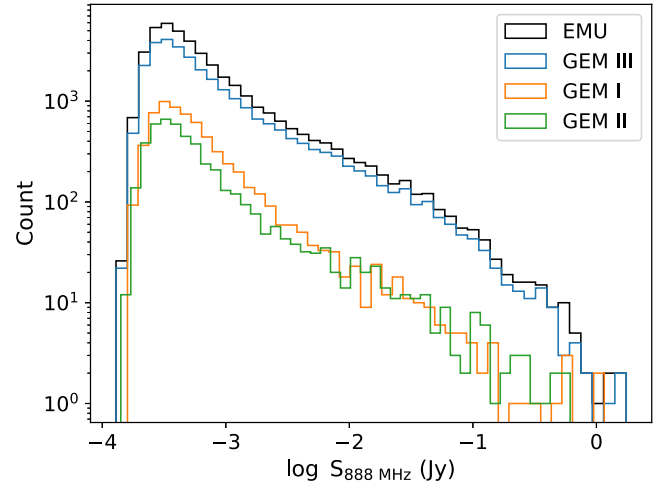
Figure 1. The 888 MHz radio flux density distributions of EMU (black), GEM III (blue), GEM I (orange), and GEM II (green) samples.
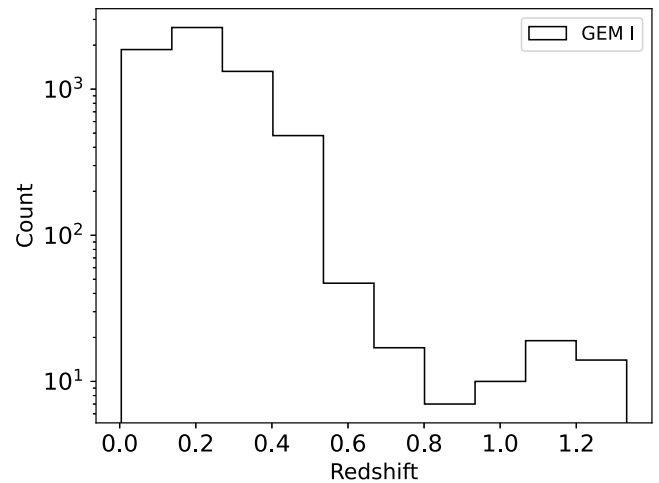
Figure 2. The redshift distribution of GEM I radio sources. The redshifts of these sources are the references for those without redshifts.

of radio sources. We start with the assumption that radio sources with similar radio flux densities and, when available, similar m_r, share comparable redshift distributions. We refine our assumptions in order to make them more realistic in later sections (see Sections 6.1 and 7).

It is worth noting here that this work anticipates future large datasets where large numbers of multiwavelength counterparts may not be available. This is the reason for using an observed parameter like radio flux density to perform the analysis. It is important to gauge the potential value in analyses that draw only on the fewest observed parameters.

From Figure 2, it is evident that most of the radio sources with spectroscopic redshifts lie at $z < 0.5$, which is expected since the targets are selected from GAMA with a limiting magnitude of $m_r < 19.8$ ($m_i < 19.2$; Driver et al. 2011, 2022). The following subsections describe our statistical redshift estimation procedure. GEM II and GEM III sources are discussed separately for clarity.

3.1. GEM II: Sources with r-band magnitude

Figure 3 shows the radio flux density and m_r distributions, as 2D histograms, of GEM I (left panel) and GEM II (right panel)
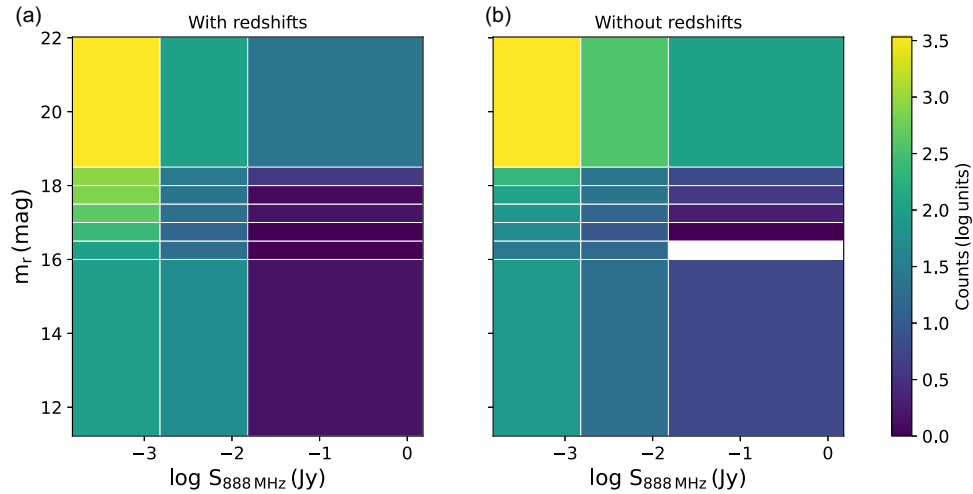
Figure 3. 2D histograms of radio flux densities and r-band magnitudes, (a): GEM I, (b): GEM II. Each bin is colour-coded based on the number counts. In both panels, radio sources are preferentially distributed towards lower radio flux densities (below $\log S_{888\,MHz} = -2.5$ Jy) and fainter r-band magnitudes (above $m_r = 16$). One white-coloured bin in (b) corresponds to zero sources in that bin.
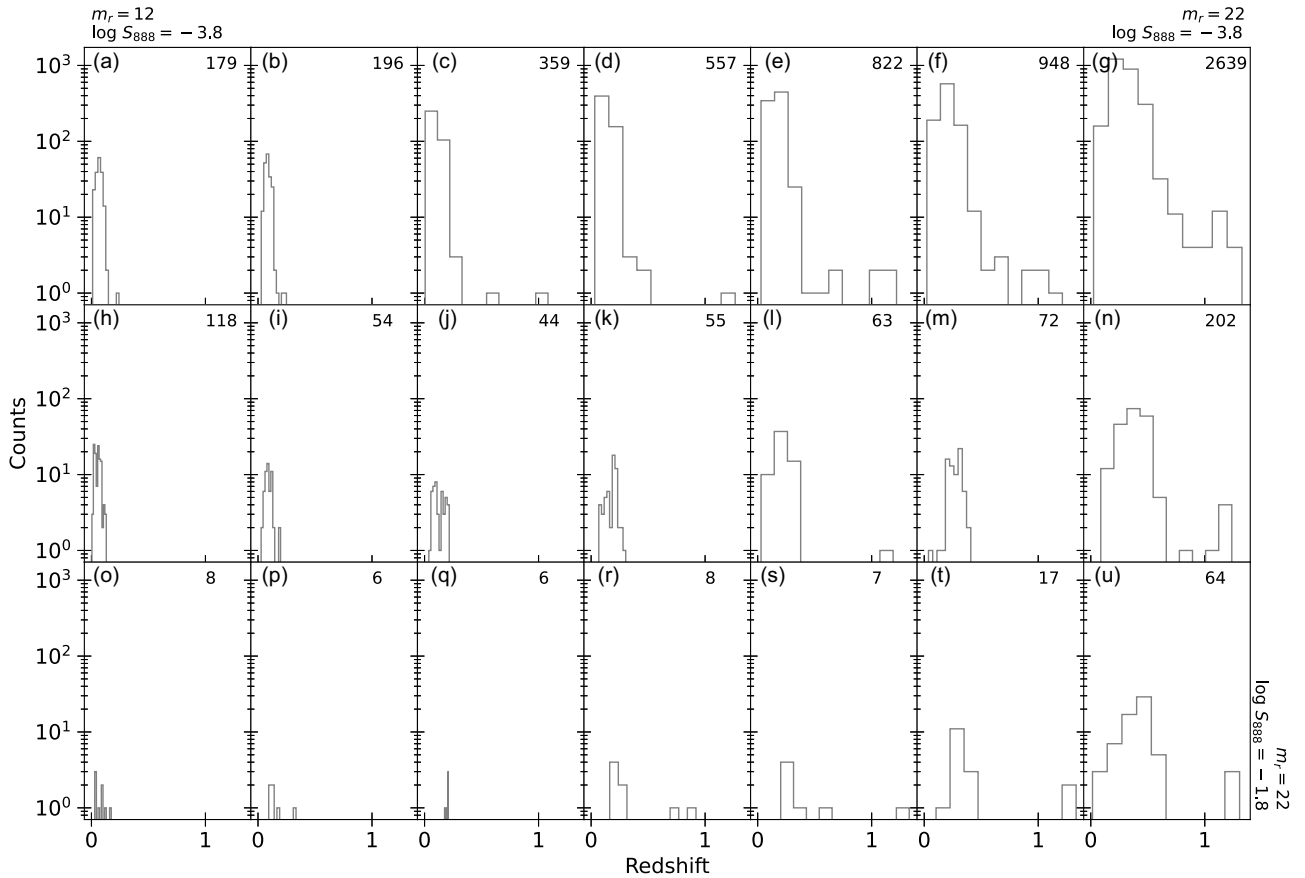
Figure 4. The redshift distribution of the GEM I sources corresponding to Figure 3a. The x-axis shows the redshift, and the counts are shown in the y-axis. The number of sources in each bin is shown in the upper right corner. The m_r and S_{888} values shown outside the panel correspond to the respective bin positions in Figure 3a.

samples. The bins are colour coded according to the number counts as shown in the colour bar.

Each bin in Figure 3a has a counterpart in Figure 3b and a corresponding redshift distribution. The bins are non-uniformly spaced so that at least six sources exist in any given bin to facilitate the fitting procedure described below to model the redshift distribution. Figure 4 shows the redshift distribution and source counts corresponding to each bin in Figure 3a. In Figure 4, m_r increases
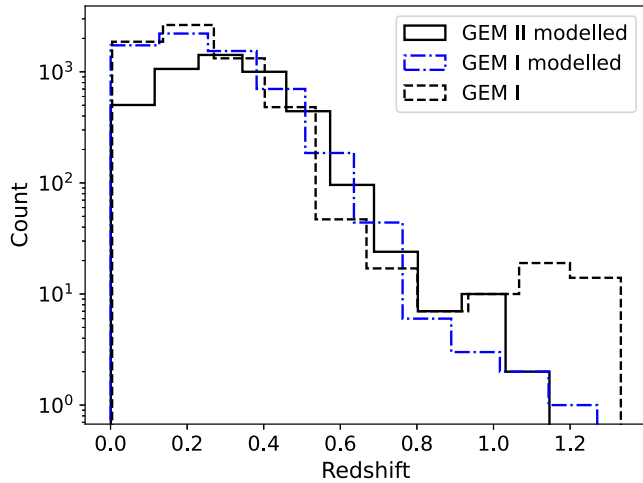
Figure 5. The modelled redshift distribution of the GEM II sample (solid line). The redshift distribution of the GEM I sample used for modelling is also shown (dashed line). The blue dot-dashed line shows the GEM I redshift distribution modelled using GEM I spectroscopic redshifts. The model reproduces the actual GEM I distribution, except for $z > 1$, details of which are given in the text (Section 3.1).

Figure 6. The 888 MHz radio flux density distributions of GEM I (dashed line) and GEM III (solid line) samples. The number counts increase to fainter radio flux densities in the same way as GEM II. However, a noticeable number of GEM III sources exist at higher radio flux densities.

across the x-axis (redshift) from the origin, and radio flux density decreases along the y-axis (count) from the origin.

We use the fitter[a] package for fitting normal distributions to each of the bins in Figure 4. The package uses the fit method of SciPy to extract the parameters of the distribution that best describes the data. The redshift distributions in each bin were tested for their normality using the quantile-quantile (Q-Q) plot. In most cases, the data are consistent with a Gaussian shape. This justifies the choice of a normal distribution in the fitting procedure. If, however, instead of choosing to fit a Gaussian, we assign redshifts randomly from the GEM I redshift distribution in a given bin, it yields comparable results for the derived RLFs. We make sure to avoid any negative redshifts given out by the model.

For a given bin, numbers are drawn randomly from the resulting fit using inbuilt random number generators in Python and are assigned as the redshift to each source in that bin of GEM II (those sources without spectroscopic redshifts). This process is repeated for all the bins, and redshifts are generated for the entire GEM II sample. Figure 5 shows how the resulting estimated GEM II redshifts are distributed (solid histogram).

We tested our statistical approach using the GEM I redshift distribution, shown in Figure 5 (blue dot-dashed histogram), which it accurately reproduces, except for $z \geq 1$. This difference for $z \geq 1$ is clear when compared to the GEM I spectroscopic redshift distribution (black dashed histogram). The underestimate here is a result of our simple single Gaussian model for the redshift distributions, which is limited when the reference set has only a few objects. This could be improved (and to include the radio sources with $z > 1$) using multiple Gaussians or Gaussian mixture models (GMMs). When these methods are applied to GEM II and GEM III samples, the modelled redshift distributions reliably reproduce that of GEM I, as expected. This is not actually desirable, however, since we do not want the GEM I redshift distribution to be duplicated in fine detail. The choice to adopt single Gaussians in modelling the redshifts has the effect of smoothing over such
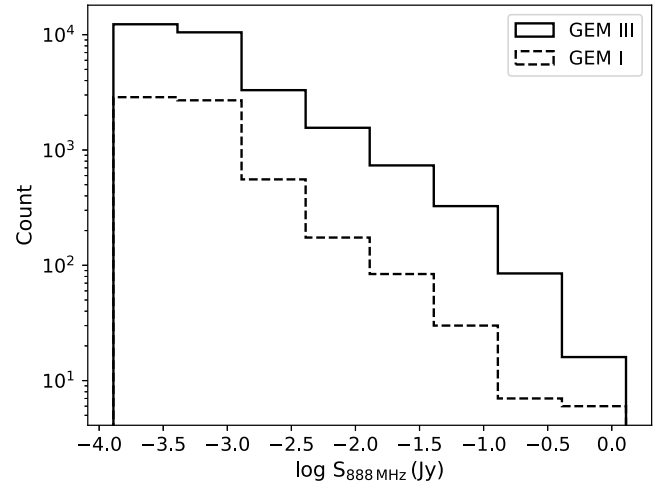
fine features, that would correspond to large scale structures or potentially rare sources at $z \geq 1$. Smoothing over these features ensures that the inferred redshift distribution follows the broad overall trend, rather than reproducing fine features that are not appropriate for the entire population.

3.2. GEM III: Sources with no r-band magnitude

Figure 6 shows the radio flux density distributions of GEM I (dashed histogram) and GEM III (solid histogram) samples. GEM III is significantly larger in size than GEM I, reflecting the fact that the full sample of radio sources far outstrips those with spectroscopic counterparts. Most of these radio sources are expected to reside at a redshift beyond the maximum limit probed by GAMA. So, we treat any results with caution while performing this initial analysis using these predictions for GEM III.

Figure 7 shows the redshift distribution in each bin of GEM I as seen in Figure 6. Here, the radio flux density increases along the x-axis. The bins corresponding to the lowest radio flux densities contain many sources with redshifts. We follow the same process here as for GEM II to assign redshifts.

Figure 8 shows the distribution of modelled GEM III redshifts (solid histogram). It is evident that GEM II and GEM III sources have similar ranges of modelled redshift distribution. GEM II sources have the advantage of having m_r, putting an additional physical constraint on the estimation procedure. In the case of GEM III sources, however, only the radio flux densities constrain the possible redshifts assigned to the sources. The redshifts allotted for GEM III sources have a distribution that mimics the input GEM I sample, but as we later show, this is not likely to be accurate. We revisit this when deriving the RLFs in Section 5. We now have the EMU sample with 39 812 sources, all of which either have spectroscopic or modelled redshifts.

4. AGN diagnostics

A plethora of techniques exist for performing AGN classification. Since the number of ionising photons from AGN and SFGs are

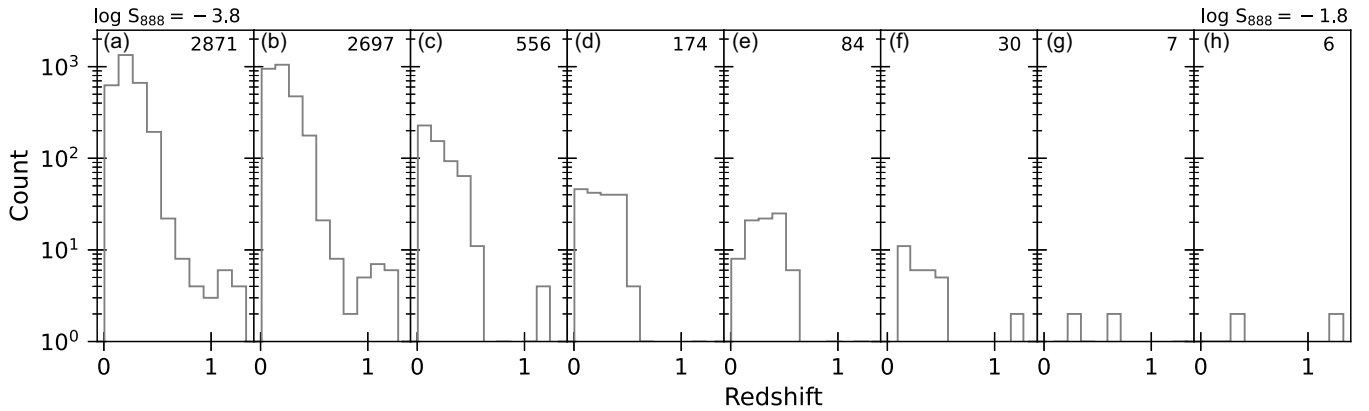[a]https://github.com/cokelaer/fitter.

Figure 7. The redshift distribution of GEM I sources, used as the template for assigning redshifts to GEM III sources, arranged according to the bins in Figure 6. The *x*-axis in each panel corresponds to redshift, and the *y*-axis shows the counts. The radio flux density increases along the *x*-axis. The number of sources is shown in the upper right corner of each panel.
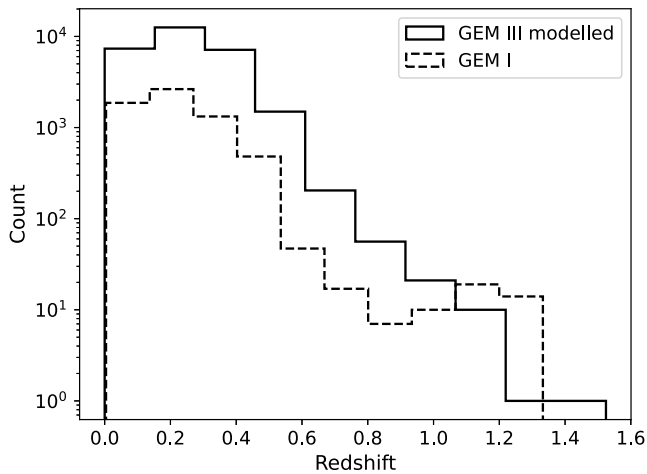

Figure 8. The modelled redshift distribution of GEM III sources (solid line). The redshift distribution of GEM I sources used for modelling is also shown (dashed line).

Figure 9. The BPT diagram: emission line classification scheme using the line ratios [NII/Hα] and [OIII]/Hβ. The diagram classifies GEM I sources into SFGs, composite galaxies, and AGN. The purple stars represent SFGs, the orange stars represent composite galaxies, and the blue triangles represent AGN. The black solid line corresponds to the Kewley et al. (2001), and the black dotted line corresponds to the Kauffmann et al. (2003) separation.

different and they are distinct in their features across the electromagnetic spectrum (Ferland & Netzer 1983; Stasińska 1984), classification can be done using optical emission lines (e.g. Baldwin, Phillips, & Terlevich 1981; Veilleux & Osterbrock 1987; Juneau et al. 2011), photometric colours (optical; e.g. Trouille et al. 2011, IR; e.g. Stern et al. 2012), or spectral energy distributions (e.g. Thorne et al. 2022; Prathap et al. 2024). It is worth noting here that the AGN discussed here belong to the radiatively efficient type, omitting the low-excitation radio galaxies (LERGs) from this approach (Best & Heckman 2012; Ching et al. 2017; Prathap et al. 2024).

Here we use the emission line classification proposed by Baldwin et al. (1981), the BPT diagram. Using the emission line ratios [NII]/Hα and [OIII]/Hβ, the BPT diagram classifies the galaxies based on their dominant ionising radiation source: star formation or black hole accretion. The classification scheme was extended by Kewley et al. (2001) with a theoretical maximum on the starburst population and Kauffmann et al. (2003) with a semi-empirical demarcation between SFGs and AGN. This revised scheme separates a given galaxy population into three: star forming, composites, and AGN. The emission lines used in the BPT diagram for classification have been shown to evolve as a function of redshift (Kewley et al. 2013). We do not need to address

this explicitly in our AGN classification, as we subsequently redefine our AGN criterion, as described below in Section 6.1. The sample used here does not have any S/N cut-off applied to the relevant spectral lines in order that the entire sample can be retained. Repetition of the analysis with an S/N>2 limit applied on the Balmer lines reduces our sample size by 30% of the current size, but does not noticeably change our final derived RLFs. We choose the emission line based method for classification here, as it is the most common in other studies that estimate LFs (e.g. Mauch & Sadler 2007; Pracy et al. 2016).

GEM I sources, having optical spectra available, are classified using the BPT diagram in Figure 9, where blue triangles represent AGN, orange stars represent composite galaxies, and purple stars represent SFGs. Of the 6 425 GEM I sources, 4 229 are SFGs, 1 569 are composites, and the remaining 627 are AGN. Since spectral lines are required to perform this classification, we cannot apply the method directly to GEM II and GEM III. The choice of an emission line based classification, however, again constrains our

Figure 10. The redshift distributions of (a) the EMU sample (solid line) and the GEM I sample (dashed line), (b) EMU SFGs (solid line) and GEM I SFGs (dashed line), and (c) EMU AGN (solid line) and GEM I AGN (dashed line),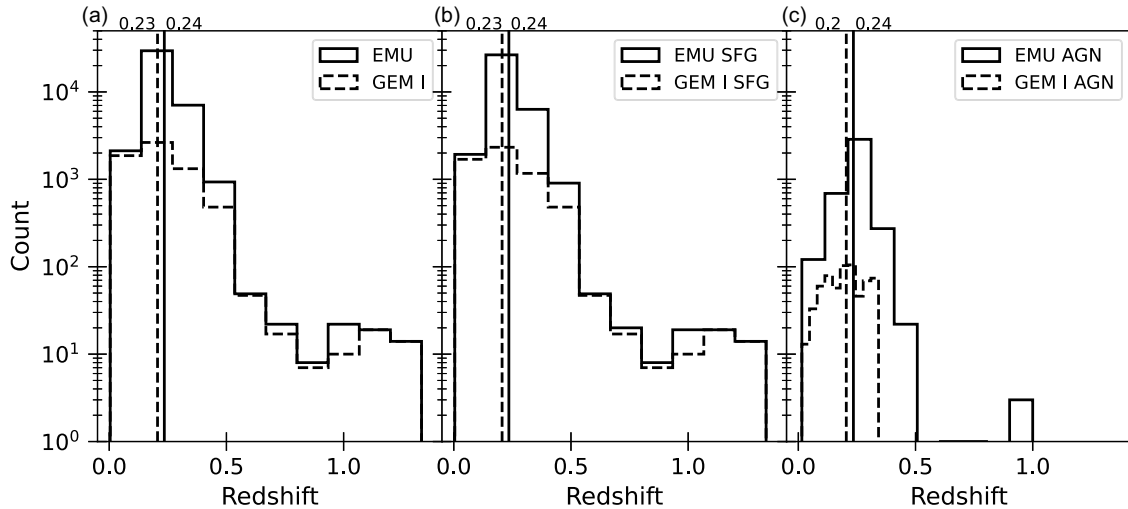 before reassigning redshifts (see Section 6.1) and AGN classification (see Section 6.1). The vertical dashed line in each panel corresponds to the median redshift of GEM I, and the vertical solid line shows the median redshift of the EMU sample. The median values of each class are shown on top of each panel.

ability to probe a larger dataset. To address this and assign classifications to the GEM II+GEM III samples, we again follow a statistical approach where classifications from GEM I sources are extended to both GEM II and GEM III.

GEM I sources are divided into bins of radio flux densities. Galaxies in each of these bins have a BPT classification. Composite galaxies are included among SFGs here because the primary study that we compare the results against (Mauch & Sadler 2007) treats composites in this way (see also, Jackson & Londish 2000; Sadler et al. 2002). The combined GEM II+GEM III sample is also divided into the same bins of radio flux densities. Probability mass functions (PMFs) are constructed for each bin where the probabilities of AGN/SFG come from the fraction of AGN/SFG over the total number of sources in that bin. For instance, if a given bin contains 565 AGN and 5 194 SFGs, the constructed PMF would correspond to ~10% AGN and ~90% SFG.

PMFs were constructed for the four bins considered (-3.89, -2.78, -1.67, -0.57, and 0.54 Jy, in log units), and the probabilities of AGN are 9.81%, 11.7%, 3.89%, and 42.8%, respectively. The bin sizes are chosen such that there are enough AGN/SFG classifications in each bin so that the newly assigned classifications are statistically robust. For a given bin, a number was chosen randomly according to the PMF of that bin to assign a classification to each source. The process was repeated, assigning a class to all sources in both GEM II and GEM III samples. This results in the entire EMU sample having a classification with 4 047 AGN and 35 765 SFGs.

5. The 888 MHz radio luminosity function

5.1. Sample properties

The EMU sample now has radio sources with their redshifts and AGN classification. Figure 10a shows the redshift distribution of the EMU sample (solid histogram) and the GEM I sub-sample (dashed histogram). The solid and dashed vertical lines correspond to the median redshifts of the EMU and GEM I samples, respectively. Figure 10b and Figure 10c show the same

for star formers in EMU and GEM I and AGN in EMU and GEM I, respectively. EMU AGN have similar median redshift (0.24), comparable to the EMU star formers (0.24). This is striking since the studies in the local universe report AGN with a higher median redshift than star formers (Mauch & Sadler 2007; Pracy et al. 2016; Prescott et al. 2016). This directly impacts the LFs derived below and is discussed in detail in Section 6. Figure 11 shows the 888 MHz radio luminosities as a function of redshift, where the EMU sample (blue circles) and the GEM I sample (orange circles) are shown in Figure 11a, the respective star formers in Figure 11b, and AGN in Figure 11c. The lower envelope results from the flux limit (200 μJy) of the EMU early science survey (Gürkan et al. 2022). It is evident from the figures that radio luminosities of EMU AGN are distributed in the same range as that of EMU SFGs. Typically, AGN are the dominant radio sources above 1.4 GHz luminosities of 10^{23} W Hz^{-1} (Condon 1989; Condon 1992; Prescott et al. 2016). However, we see more star formers in the EMU sample in this regime, which we attribute to our statistical approach to redshift estimation. AGN that are actually at a higher redshift, if assigned a lower redshift value, will appear in the sample as a low-luminosity source. We address this in Section 6 when discussing the RLFs.

5.2. Volume corrections: The classical $1/V_{max}$ method

It is typical for flux-limited samples to become progressively more incomplete as the flux limit is approached. The completeness of a catalogue can be improved by adding missing objects from other sources (Kiang 1961) or each entry of the catalogue can be statistically weighted by the magnitude-dependent incompleteness function (Sandage, Tammann, & Yahil 1979). Several methods exist, but we focus on the $1/V_{max}$ method.

Schmidt (1968) devised the V/V_{max} technique to test and to correct for incompleteness. Here, V is the sample volume between the galaxy and the observer, and V_{max} represents the maximum volume in which the galaxy can exist without dropping below the flux limit (in this case, the radio flux density limit, 200 μJy and the optical magnitude, $m_r = 19.8$). Felten (1976) showed that instead of having the number of galaxies in a given luminosity bin divided
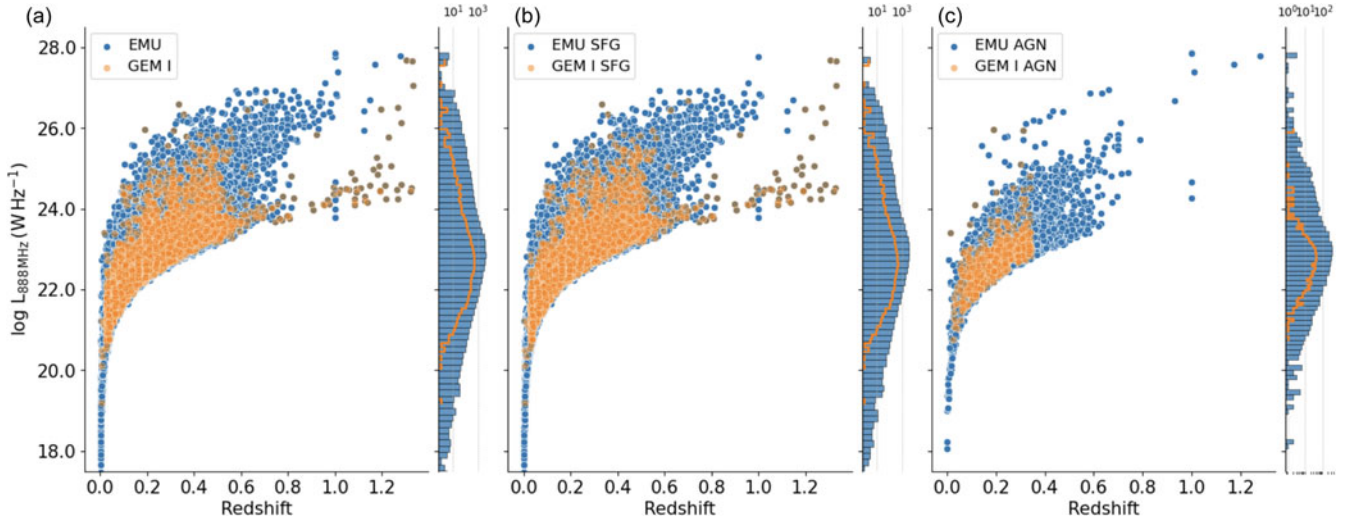
Figure 11. The figure shows the variation of the 888 MHz luminosities with redshift of (a) the entire EMU sample (blue circles) and GEM I sample (orange circles), (b) EMU star formers (blue circles) and GEM I star formers (orange circles), and (c) EMU AGN (blue circles) and GEM I AGN (orange circles), before reassigning redshifts (see Section 6.1) and AGN classification (see Section 6.1). The marginal histogram on the y-axis shows the distribution of 888 MHz luminosities in log units with the same colours as the scatter plots.

by V_{max}, the LF can be estimated by the sum, $\sum (1/V_{max})$ over all galaxies in that bin.

To calculate the LF, the EMU sample is divided into equally spaced redshift bins ($\Delta z = 0.1$). Each redshift bin is further divided into equally spaced luminosity bins ($\Delta \log L = 0.25$). Construction of the LF requires, along with volume corrections, the flux-dependent incompleteness as discussed in Section 2.3, and the area correction, since the EMU early science observation covers a patch of the sky, the G23 field, totalling 83 deg^2 (Gürkan et al. 2022). Including all these contributions, the RLF is calculated as,

$$\phi(L) = \frac{1}{\Delta \log L} \sum_{i=1}^{N} \frac{1}{V_{max,i}}, \quad (5)$$

where i runs through each galaxy in a given redshift and luminosity bin. $V_{max,i}$ is calculated as

$$V_{max,i} = [V(z_{max}) - V(z_{min})] \; C_i, \quad (6)$$

where z_{min} is the lower bound of a given redshift bin and z_{max} is the upper bound of that redshift bin, or it is the redshift at which the galaxy falls below the flux limit of the survey, whichever is the minimum. C_i is the geometrical and statistical correction factor that takes into account the observed area and flux-dependent completeness:

$$C_i = \frac{A}{41\,253 \text{ deg}^2} \; C_{S_{888\,MHz}}, \quad (7)$$

where $A = 83$ deg^2 and $C_{S_{888\,MHz}}$ is the flux-dependent completeness factor (Section 2.3). Rest frame luminosities (Hogg et al. 2002) are calculated by assuming a ΛCDM cosmology following the prescription in Novak et al. (2017) for K-correction with $\alpha = -0.7$, where α is the radio spectral index (we adopt the convention, $S_\nu \propto \nu^\alpha$).

6. Initial outcomes

Figure 12 shows our resulting 888 MHz RLFs in eight redshift bins. In each panel, the blue circles represent the EMU sample, the stars and triangles correspond to the SFGs and AGN, respectively, in the combined GEM I and GEM II (hereafter GEM I+II) sample. The error bars are derived by assuming a Poisson distribution for the sources in each luminosity bin. The solid and dashed lines represent the parametric fits to the SFG and AGN RLFs, respectively, derived by Mauch & Sadler (2007) (labelled as MS+07) using the analytical functions from Saunders et al. (1990, Equation 2) for SFGs and Dunlop & Peacock (1990) and Brown et al. (2001) for AGN, converted from 1.4 GHz to 888 MHz assuming $\alpha = -0.7$. For higher redshifts, the fitted lines are allowed to evolve assuming PLE (Eq. 4), where k_L is taken to be 3 (Sadler et al. 2007; Padovani et al. 2011; McAlpine, Jarvis, & Bonfield 2013; Prescott et al. 2016), and the redshift is the mean of the bin. We do not fit our derived LFs with an analytical form; rather, we compare them with previous results.

It is clear in Figure 12 that the RLF constructed using the modelled redshifts (the blue circles) results in an overestimate with respect to the Mauch & Sadler (2007) result. This discrepancy arises due to the use of the GAMA redshift distribution to sample radio source redshifts, since most are actually likely to be lying at higher redshifts.

GEM I+II sources (stars for SFGs and triangles for AGN) have redshifts that are both spectroscopic and modelled (from the estimation procedure described in Section 3), with $\approx 36\%$ of them being modelled. It is remarkable to see that the total RLF for the GEM I+II sample is in reasonably good agreement with the Mauch & Sadler (2007) result up to a redshift of $z = 0.5$. The LFs for the AGN are severely underestimated for $z \geq 0.2$ (see panels c – h of Figure 12). Although, the GEM I+II SFGs and AGN show agreement with the respective Mauch & Sadler (2007) lines only upto $z = 0.2$. This is most likely associated with the selection of only optically active AGN, as described in Section 4. We address this issue in Section 6.1. Since we are looking at a subsample that may not be complete, the SFGs and AGN in the high redshift bins
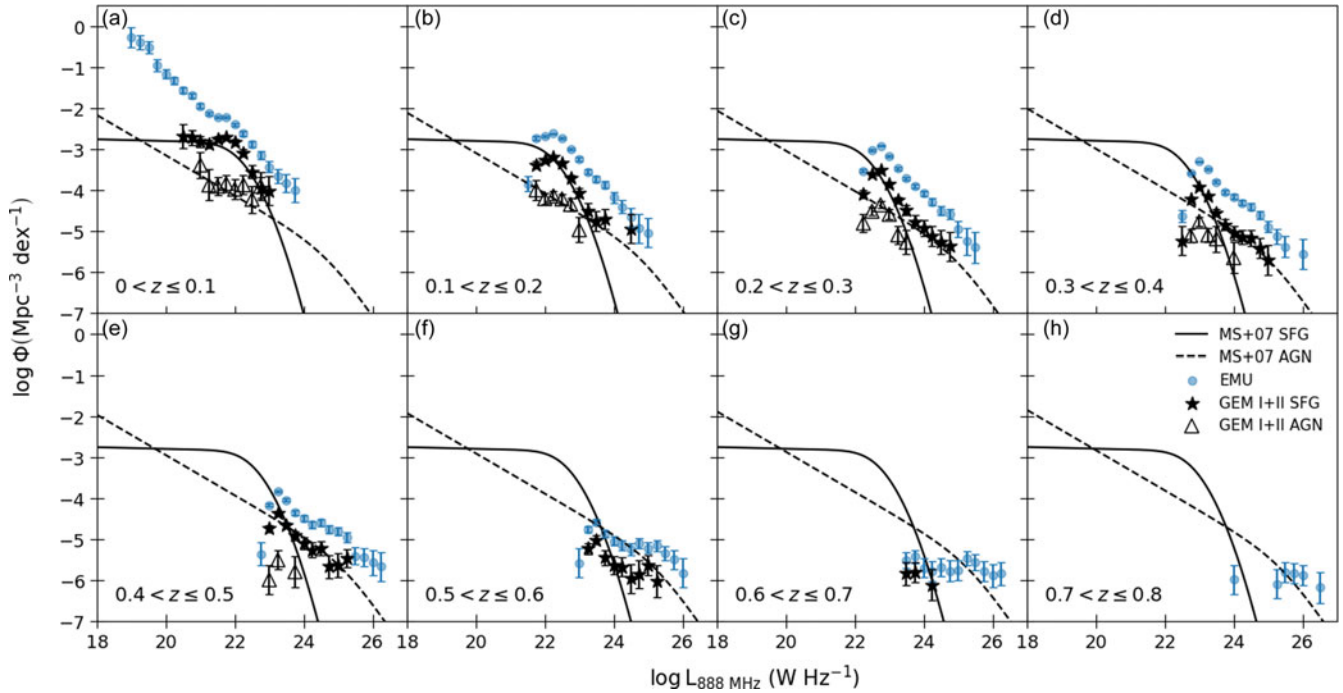
Figure 12. The 888 MHz RLFs of the EMU sample (blue circles) compared with GEM I+II SFGs (the stars) and GEM I+II AGN (the triangles). The error bars assume that the numbers are Poisson distributed. The solid and dashed lines correspond to the parametric fits of Mauch & Sadler (2007) for SFGs and AGN, respectively, extrapolated to 888 MHz assuming $\alpha = -0.7$. The bins $a - h$ correspond to the RLFs at progressively increasing redshifts where the Mauch & Sadler (2007) lines are allowed to evolve in a pure luminosity fashion as described in the text. It is worth noting that how well the total GEM I+II RLF follow the Mauch & Sadler (2007) lines, up to $z = 0.5$ (see text for details). It is also crucial to note that the redshifts of GEM II sources are modelled (close to 36% of the GEM I+II sample). The low-luminosity end starts to drop off as we move to higher redshifts, which is a consequence of the flux limit of the survey. It is also interesting to note that high redshift bins lack AGN, and there are star formers at high radio luminosities ($> 10^{24}$ W Hz^{-1}) in these bins.

are expected to fall in number, as can be seen in Figure 12. The GEM I+II star formers and AGN fall progressively below the faint end of the Mauch & Sadler (2007) lines as we move to higher redshifts, which is a consequence of the flux limit of the survey and the resulting incompleteness of the faint sources.

The mismatch between our GEM III LFs and the parametric fits by Mauch & Sadler (2007) can be attributed to the fact that a large portion of our modelled redshifts (close to 70%) have the radio flux density as the sole constraint, most likely realistically lying at a larger redshift. We explore this further in the next section.

6.1. Redistributing the redshifts and AGN classification

The GEM I+II sample, ignoring the AGN underestimate above $z = 0.2$, is a good match for the Mauch & Sadler (2007) results at low redshifts ($0 < z \le 0.5$) and is likely to accurately account for the entirety of that population given the high degree of spectroscopic completeness in the GAMA measurements. To confirm, we checked the agreement between the RLFs for GEM I and the Mauch & Sadler (2007) results and found that the total RLFs agree only upto $z = 0.3$. This means, both GEM I and GEM II samples are required to accurately account for the RLFs up to $z = 0.5$. To account for the overestimation in LF for radio sources without optical counterparts (GEM III, Figure 12), we adopt a different strategy.

The GEM III sources are unlikely to be well represented by the GAMA redshift distribution ($z_{max} \approx 0.5$). To account for this, we choose a more realistic high redshift distribution for these radio galaxies. Smolčić et al. (2017) derived the 1.4 GHz RLF for
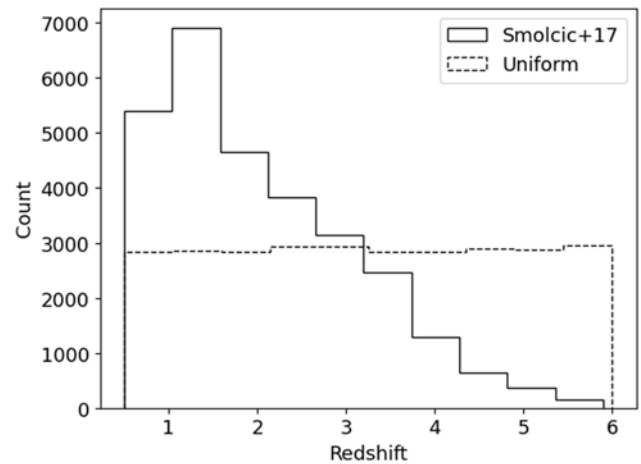
Figure 13. Solid line: the redshift distribution of the high redshift universe probed by Smolčić et al. (2017). This scaled up version in the range $0.5 < z < 6$ is a more realistic representation that can be used to model the GEM III source redshifts. Dashed line: the uniform redshift distribution sampling the high redshift Universe in the absence of a realistic high redshift distribution (Section 7.1).

AGN out to $z \approx 5.5$ using the VLA-COSMOS observations and the COSMOS mutliwavelength data. We use this distribution and rederive the redshifts for the GEM III sources (as described in Section 3.2). Figure 13 shows the Smolčić et al. (2017) distribution scaled up to match the GEM III sample size (solid line), along with a uniform distribution (dashed line) that we explore further in Section 7.1 below.

The GEM III sources from the low redshift bins ($z \leq 0.5$) have now been redistributed to redshifts above $z = 0.5$ and the LFs are recalculated. The LFs for the AGN, however, are still severely underestimated for $z \geq 0.3$. This is a consequence of neglecting radio loudness (or strong radio emission; $L_{888\,MHz} \geq 10^{23.5}$ W Hz^{-1}) as a selection criterion for AGN. It has been established that luminous radio sources are typically AGN (White et al. 2000). We have around 10% of the sources currently classified as star formers based solely on the BPT inspired classification, which also have high radio luminosities ($L_{888\,MHz} \geq 10^{23.5}$ W Hz^{-1}) from the method above. Converting these radio luminosities to star formation rates (Hopkins et al. 2003) would result in an SFR ≥ 100 M$_\odot$yr^{-1}, which is not typical for a large sample of SFGs. Erroneously classifying this population as AGN would have a negligible effect on the inferred luminosity functions. This approach has the inherent assumption that the GEM III sources are dominated by AGN. This argument is supported by the small fraction ($\approx 0\%$) of SFGs with $L_{1.4\,GHz} \geq 10^{23.5}$ W Hz^{-1} (see Figure 8 of Mauch & Sadler 2007) and also that the GEM III sources are redistributed in to $z \geq 0.5$.

To resolve this, we modify our classification of GEM II+GEM III to include radio luminosity as a criterion. The sources in the EMU sample with $L_{888\,MHz} > 10^{23.5}$ W Hz^{-1} are now labelled AGN, and the RLFs are recalculated. In addition, to understand the uncertainties in RLFs for different estimated redshifts, we also implement multiple realisations of the RLF. Table 2 contains the 888 MHz RLFs of the SFGs and AGN in the EMU sample.

These LFs are estimated as the median of one hundred realisations, with redshifts assigned randomly, following the procedures detailed above each time. The uncertainties are calculated as the 1σ standard deviations of the one hundred realisations (superscript) and lie between 0.01 and 0.1. The actual errors are dominated by that from Poisson counting (subscript). We discuss these rederived LFs in the next section.

7. Discussion

Figure 14 shows the result of calculating one hundred realisations of the 888 MHz RLFs with the error bars representing the Poisson errors. The symbols and lines follow Figure 12. These new RLFs exhibit significant improvement over Figure 12 in how the star formers and AGN follow the Mauch & Sadler (2007) lines. Over the entire redshift range considered, with the applied PLE evolution, the star formers and AGN follow the reference LFs very well.

Another interesting feature in Figure 14 is the low-luminosity end of the lowest redshift bin as it rises considerably above the extrapolation from the Mauch & Sadler (2007) results. This is a consequence of the multiple realisations of redshifts sampling the low-luminosity end, which are otherwise not likely to be detected. This is probably because of the sensitivity limits of the survey and the multiple realisations of the RLF most likely increases the likelihood of detection of these faint sources. Physically, this implies a larger number of faint radio sources at low redshifts, compared to the power-law tail of the Saunders function fit of Mauch & Sadler (2007). The deviation begins close to $L_{888\,MHz} \approx 10^{20}$ W Hz^{-1} corresponding to an SFR of ≈ 0.07 M$_\odot$ yr^{-1} (Hopkins et al. 2003). Such an SFR value implies that the faint galaxies are of low stellar masses with $\log M_*/M_\odot \leq 9.5$ (Noeske et al. 2007a;

Table 2. The 888 MHz RLFs ($\log \phi$) of the SFGs and AGN in the EMU sample and corresponding numbers (N) in each bin. The values here correspond to the logarithm of the median of the LFs calculated with one hundred estimated redshifts. The reported uncertainties are the 1σ deviations from the median (superscript) and Poisson errors (subscript). The luminosities shown here are the mean of each luminosity bin.

Redshift	$\log L$	SFGs		AGN	
		N	$\log \phi$	N	$\log \phi$
	(W Hz^{-1})		(Mpc^{-3} dex^{-1})		(Mpc^{-3} dex^{-1})
$0.0 < z \leq 0.1$	19.75	3	$-1.59^{\pm0.01}_{\pm0.58}$	–	–
	20.0	6	$-1.95^{\pm0.10}_{\pm0.41}$	–	–
	20.25	9	$-2.27^{\pm0.11}_{\pm0.33}$	–	–
	20.5	15	$-2.54^{\pm0.09}_{\pm0.26}$	–	–
	20.75	32	$-2.60^{\pm0.06}_{\pm0.18}$	–	–
	21	64	$-2.72^{\pm0.04}_{\pm0.13}$	8	$-3.42^{\pm0.04}_{\pm0.35}$
	21.25	113	$-2.87^{\pm0.03}_{\pm0.09}$	10	$-3.68^{\pm0.05}_{\pm0.32}$
	21.5	304	$-2.76^{\pm0.01}_{\pm0.06}$	25	$-3.83^{\pm0.04}_{\pm0.20}$
	21.75	373	$-2.68^{\pm0.01}_{\pm0.05}$	22	$-3.89^{\pm0.06}_{\pm0.21}$
	22	247	$-2.82^{\pm0.01}_{\pm0.06}$	20	$-3.89^{\pm0.05}_{\pm0.22}$
	22.25	122	$-3.10^{\pm0.01}_{\pm0.09}$	19	$-3.84^{\pm0.05}_{\pm0.23}$
	22.5	48	$-3.49^{\pm0.02}_{\pm0.14}$	12	$-4.04^{\pm0.06}_{\pm0.29}$
	22.75	16	$-3.92^{\pm0.04}_{\pm0.25}$	11	$-3.96^{\pm0.07}_{\pm0.30}$
	23	8	$-3.92^{\pm0.04}_{\pm0.35}$	–	–
$0.1 < z \leq 0.2$	21.75	136	$-3.37^{\pm0.02}_{\pm0.09}$	22	$-4.00^{\pm0.07}_{\pm0.23}$
	22	498	$-3.23^{\pm0.02}_{\pm0.05}$	56	$-4.21^{\pm0.03}_{\pm0.13}$
	22.25	786	$-3.20^{\pm0.01}_{\pm0.04}$	85	$-4.14^{\pm0.02}_{\pm0.11}$
	22.5	597	$-3.29^{\pm0.01}_{\pm0.04}$	78	$-4.17^{\pm0.03}_{\pm0.11}$
	22.75	219	$-3.68^{\pm0.01}_{\pm0.07}$	51	$-4.29^{\pm0.04}_{\pm0.14}$
	23	67	$-4.18^{\pm0.02}_{\pm0.12}$	30	$-4.50^{\pm0.06}_{\pm0.18}$
	23.25	25	$-4.56^{\pm0.03}_{\pm0.20}$	13	$-4.79^{\pm0.08}_{\pm0.28}$
	23.5	–	–	22	$-4.62^{\pm0.06}_{\pm0.21}$
	23.75	–	–	15	$-4.77^{\pm0.07}_{\pm0.26}$
	24	–	–	11	$-4.90^{\pm0.06}_{\pm0.30}$
$0.2 < z \leq 0.3$	22.25	90	$-4.09^{\pm0.06}_{\pm0.11}$	17	$-4.83^{\pm0.04}_{\pm0.24}$
	22.5	786	$-3.61^{\pm0.01}_{\pm0.04}$	85	$-4.54^{\pm0.02}_{\pm0.11}$
	22.75	1010	$-3.48^{\pm0.01}_{\pm0.03}$	129	$-4.36^{\pm0.02}_{\pm0.09}$
	23	437	$-3.80^{\pm0.01}_{\pm0.05}$	80	$-4.52^{\pm0.03}_{\pm0.11}$
	23.25	145	$-4.26^{\pm0.02}_{\pm0.08}$	57	$-4.62^{\pm0.05}_{\pm0.13}$
	23.5	–	–	94	$-4.41^{\pm0.02}_{\pm0.10}$
	23.75	–	–	49	$-4.68^{\pm0.04}_{\pm0.14}$
	24	–	–	29	$-4.90^{\pm0.06}_{\pm0.19}$
	24.25	–	–	24	$-5.00^{\pm0.05}_{\pm0.20}$
	24.5	–	–	16	$-5.12^{\pm0.06}_{\pm0.25}$
	24.75	–	–	10	$-5.27^{\pm0.06}_{\pm0.32}$
$0.3 < z \leq 0.4$	22.75	294	$-4.29^{\pm0.02}_{\pm0.06}$	32	$-5.03^{\pm0.12}_{\pm0.18}$
	23	718	$-3.88^{\pm0.01}_{\pm0.04}$	86	$-4.77^{\pm0.03}_{\pm0.11}$
	23.25	431	$-4.06^{\pm0.02}_{\pm0.05}$	45	$-5.02^{\pm0.05}_{\pm0.15}$
	23.5	64	$-4.84^{\pm0.05}_{\pm0.13}$	110	$-4.62^{\pm0.02}_{\pm0.10}$
	23.75	–	–	82	$-4.72^{\pm0.03}_{\pm0.11}$
	24	–	–	50	$-4.90^{\pm0.04}_{\pm0.14}$

Table 2. Continued.

Redshift	log L	SFGs		AGN	
		N	log ϕ	N	log ϕ
	(W Hz^{-1})		(Mpc^{-3} dex^{-1})		(Mpc^{-3} dex^{-1})
	24.25	–	–	36	$-5.06^{\pm0.04}_{\pm0.14}$
	24.5	–	–	32	$-5.09^{\pm0.04}_{\pm0.18}$
	24.75	–	–	25	$-5.22^{\pm0.04}_{\pm0.32}$
	25	–	–	10	$-5.56^{\pm0.05}_{\pm0.32}$
$0.4<z\leq0.5$	23	117	$-4.85^{\pm0.03}_{\pm0.09}$	8	$-5.66^{\pm0.17}_{\pm0.37}$
	23.25	384	$-4.33^{\pm0.02}_{\pm0.05}$	24	$-5.49^{\pm0.07}_{\pm0.20}$
	23.5	128	$-4.75^{\pm0.03}_{\pm0.09}$	96	$-4.87^{\pm0.02}_{\pm0.10}$
	23.75	32	$-5.29^{\pm0.07}_{\pm0.18}$	68	$-4.99^{\pm0.04}_{\pm0.12}$
	24	–	–	58	$-5.03^{\pm0.03}_{\pm0.13}$
	24.25	–	–	48	$-5.09^{\pm0.04}_{\pm0.14}$
	24.5	–	–	37	$-5.19^{\pm0.04}_{\pm0.16}$
	24.75	–	–	26	$-5.35^{\pm0.05}_{\pm0.20}$
	25	–	–	17	$-5.51^{\pm0.07}_{\pm0.24}$
	25.25	–	–	18	$-5.47^{\pm0.06}_{\pm0.24}$
$0.5<z\leq0.6$	23.25	196	$-4.77^{\pm0.01}_{\pm0.07}$	25	$-5.58^{\pm0.03}_{\pm0.20}$
	23.5	72	$-5.18^{\pm0.05}_{\pm0.12}$	304	$-4.54^{\pm0.01}_{\pm0.06}$
	23.75	33	$-5.43^{\pm0.07}_{\pm0.17}$	138	$-4.82^{\pm0.01}_{\pm0.09}$
	24	6	$-5.83^{\pm0.07}_{\pm0.41}$	100	$-4.95^{\pm0.01}_{\pm0.10}$
	24.25	–	–	60	$-5.12^{\pm0.04}_{\pm0.13}$
	24.5	–	–	44	$-5.28^{\pm0.04}_{\pm0.15}$
	24.75	–	–	35	$-5.37^{\pm0.04}_{\pm0.17}$
	25	–	–	20	$-5.53^{\pm0.06}_{\pm0.22}$
	25.25	–	–	28	$-5.46^{\pm0.04}_{\pm0.19}$
	25.5	–	–	9	$-5.78^{\pm0.06}_{\pm0.33}$
$0.6<z\leq0.7$	23.25	10	$-6.16^{\pm0.03}_{\pm0.32}$	–	–
	23.5	–	–	215	$-4.82^{\pm0.01}_{\pm0.07}$
	23.75	–	–	209	$-4.78^{\pm0.01}_{\pm0.07}$
	24	–	–	101	$-5.06^{\pm0.01}_{\pm0.10}$
	24.25	–	–	60	$-5.26^{\pm0.01}_{\pm0.13}$
	24.5	–	–	38	$-5.43^{\pm0.03}_{\pm0.16}$
	24.75	–	–	30	$-5.52^{\pm0.03}_{\pm0.18}$
	25	–	–	27	$-5.55^{\pm0.03}_{\pm0.19}$
	25.25	–	–	16	$-5.73^{\pm0.06}_{\pm0.25}$
	25.5	–	–	16	$-5.74^{\pm0.05}_{\pm0.25}$
	25.75	–	–	10	$-5.92^{\pm0.05}_{\pm0.32}$
$0.7<z\leq0.8$	23.5	–	–	175	$-5.00^{\pm0.01}_{\pm0.08}$
	23.75	–	–	381	$-4.62^{\pm0.01}_{\pm0.05}$
	24	–	–	201	$-4.85^{\pm0.01}_{\pm0.07}$
	24.25	–	–	118	$-5.05^{\pm0.01}_{\pm0.09}$
	24.5	–	–	55	$-5.35^{\pm0.01}_{\pm0.13}$
	24.75	–	–	54	$-5.33^{\pm0.01}_{\pm0.14}$
	25	–	–	23	$-5.66^{\pm0.03}_{\pm0.21}$
	25.25	–	–	28	$-5.62^{\pm0.02}_{\pm0.19}$
	25.5	–	–	20	$-5.73^{\pm0.03}_{\pm0.22}$
	25.75	–	–	13	$-5.77^{\pm0.06}_{\pm0.28}$

Noeske et al. 2007b) and they are potentially dwarf systems. Such upturns have also been seen in optical LFs estimated using GAMA data (e.g. Loveday et al. 2012; Kelvin et al. 2014; Vázquez-Mata et al. 2020). The new generation of radio surveys probes large sky areas with high sensitivity, enabling the identification of numerous faint galaxies, and we expect that this feature may start to be detected as well at low radio luminosities.

Surprisingly, perhaps, random redshift assignments based on the radio flux densities and m_r do a remarkable job in RLF estimation. This is in part a coincidence arising from the GAMA magnitude limit ($m_i < 19.2$; Driver et al. 2011), together with the high level of completeness of the GAMA survey. Consequently, radio sources with a GAMA counterpart will also be complete for the low redshift ranges ($z \leq 0.5$) that GAMA is primarily sensitive to. Conversely, radio sources without a GAMA counterpart must lie at higher redshifts than those probed by GAMA. Some sources could be of low stellar mass ($\log M_*/M_\odot \leq 7.5 - 8$) and hence faint in radio luminosity, lying below the radio luminosities probed here.

Since the RLFs below $z \approx 0.5$ match remarkably well with the Mauch & Sadler (2007) results, it is apparent that radio sources having optically faint counterparts ($m_r > 19.8$) most likely lie at redshifts above 0.5. Figure 14 also implies that radio sources brighter than $L_{888\,MHz} > 10^{23.5}$ W Hz^{-1} are typically AGN.

7.1. A scenario with no reference distribution

We now look at a scenario where we do not have a realistic high redshift distribution to sample redshifts from. To explore from a very basic level, we start with the simplest possible form: a uniform distribution in the range $0.5 < z < 6$. The redshifts for radio sources in GEM III are now sampled from this uniform distribution (Figure 13, dashed line) and the entire procedure is repeated to derive the new RLFs.

Figure 15 compares the RLFs derived using both the Smolčić et al. (2017) redshift distribution (blue triangles) and uniform redshift distribution (red triangles) with the radio AGN LFs of Smolčić et al. (2017) (magenta dots). The magenta dashed line represents the analytical fit derived by Smolčić et al. (2017) with an incorporated PLE. These RLFs are also compared against a bunch of recent works (Ceraj et al. 2018; Butler et al. 2019; Condon et al. 2019; Wang et al. 2024), all measured at 1.4 GHz and converted to 888 MHz assuming $\alpha = -0.7$, as shown in the figure. The redshift bins in this plot are chosen to match Smolčić et al. (2017). The residual plots quantify the logarithmic separation between the derived LFs (following the same colours as the LFs) with the Smolčić et al. (2017) analytical fit.

The RLFs we derived using the Smolčić distribution (blue triangles) and the Smolčić et al. (2017) LFs (magenta circles) follow each other well in most of the redshift ranges probed here. In addition, our derived LFs are more densely sampled and extend to the high luminosity end (close to 1 dex brighter) towards progressively higher redshifts.

The LFs derived in this work using the Smolčić (blue triangles) and the uniform redshift distribution (red triangles) agree with each other well in most redshift ranges. These LFs exhibit a drop in the lowest few luminosity bins which can be attributed to various reasons that still persist even after the completeness corrections, including the sensitivity limits of the surveys used. The separation
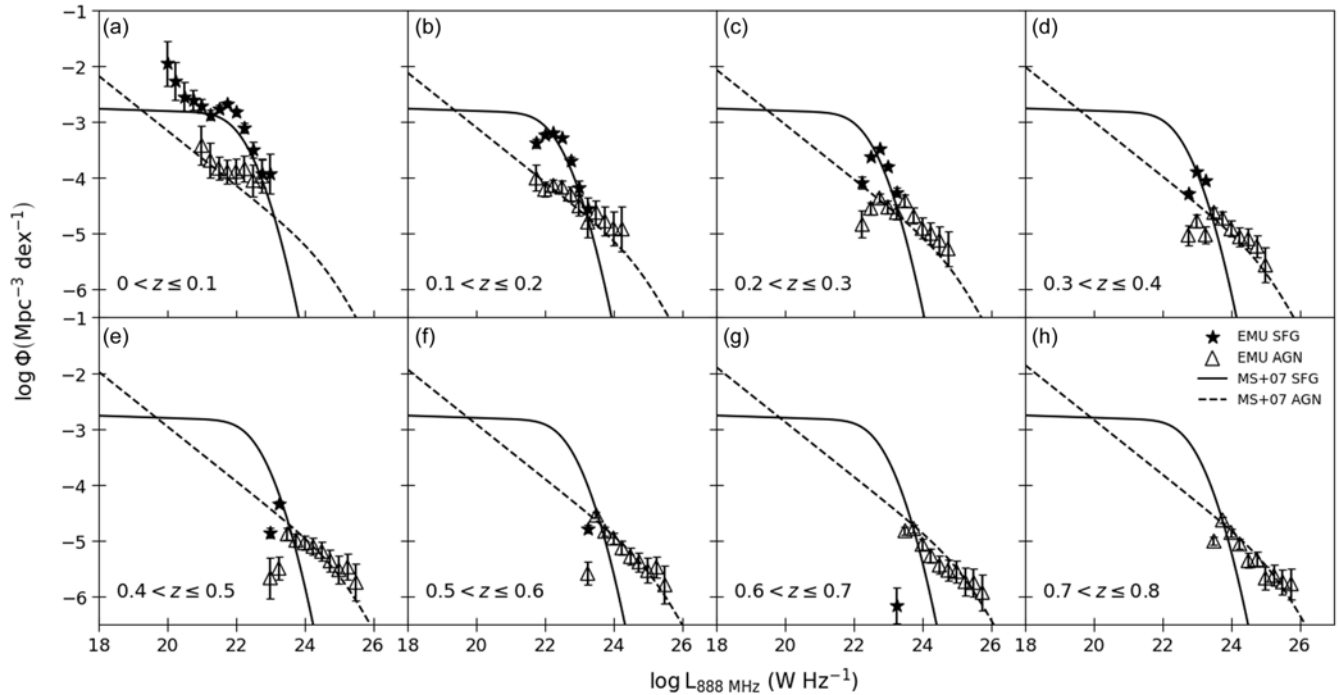
Figure 14. The 888 MHz RLFs derived from the one hundred realisations. The symbols and lines follow the previous figures. The LFs plotted here are the median of the one hundred estimated LFs, and the error bars correspond to the Poisson errors. It is remarkable to see that the SFG and AGN LFs follow the Mauch & Sadler (2007) lines in every redshift bin. The upturn in the faint end of the low redshift LF, as discussed in the text, is also worth noting.

between the two approaches a maximum of ≈ 0.6 dex in the highest redshift bin (at $L_{888\,MHz} = 10^{26}$ W Hz^{-1}). These results point to two realisations, (*i*) our statistical approach to measuring RLFs does surprisingly well even at $z \approx 5.5$, and (*ii*) even adopting a uniform distribution for radio galaxy redshifts up to $z = 6$ leads to results broadly comparable, at levels notably better than an order of magnitude, in deriving the LFs. This agreement is also reflected in their respective residuals.

The RLF is a statistical description of how common radio galaxies of various luminosities are in the Universe. The lack of redshifts for a large fraction of the sample requires us to look for alternatives in calculating the RLFs. Our approach is the application of a simple statistical method to measure the RLF in the absence of redshfits. Rather than the properties of a single radio galaxy, we focus on the characteristics of a population and are successful in generating LFs that compare well with previous studies. The application of the uniform distribution demonstrates another viable option in measuring the RLFs in the absence of any additional multiwavelength information and a robustly constrained redshift distribution.

8. Conclusion

We have used a sample of 39 812 radio galaxies from the EMU early science observations to derive their RLFs at 888 MHz. Since only $\approx 16\%$ of the radio galaxies have optical counterparts and hence redshifts, we modelled the redshift distributions of the radio sources using simple statistical tools, which is then used to derive RLFs. Since the redshift distribution of GAMA sources was not a realistic representation of high redshift radio galaxies, we used the Smolčić et al. (2017) redshift distribution to derive the RLFs of

those high redshift radio galaxies. Considering a scenario where the high redshift radio sources do not have a multiwavelength counterpart, we assumed a uniform distribution for the high redshift radio galaxy redshifts and rederived the RLFs. The main conclusions are the following:

- In the absence of observational counterparts and hence individual redshifts, modelling a redshift distribution for the entire sample has turned out to be a surprisingly effective tool in calculating the LFs.
- A probabilistic approach can help detect population that are less likely to be identified observationally. The faint end population seen in Figure 14 serves as an example.
- Given a sparsely sampled dataset spanning any redshift range, our method can be used to create redshift distributions, hence densely populated samples. The RLFs derived using this distribution probes the high luminosity end, especially at higher redshifts. This feature is evident in Figure 15 while comparing the blue and magenta data points.

In the absence of accurate redshift information, it is clearly still possible to reproduce reliable properties for broad galaxy population statistics, using this kind of method. This has the advantage of leveraging large numbers of sources to improve the sampling of such distributions compared to small samples with precise redshift measurements. The limitations, of course, are that rare populations and accurate redshifts are not able to be reproduced.

In a follow up work we plan to use the LFs derived in this way to explore the star formation rate and AGN luminosity density
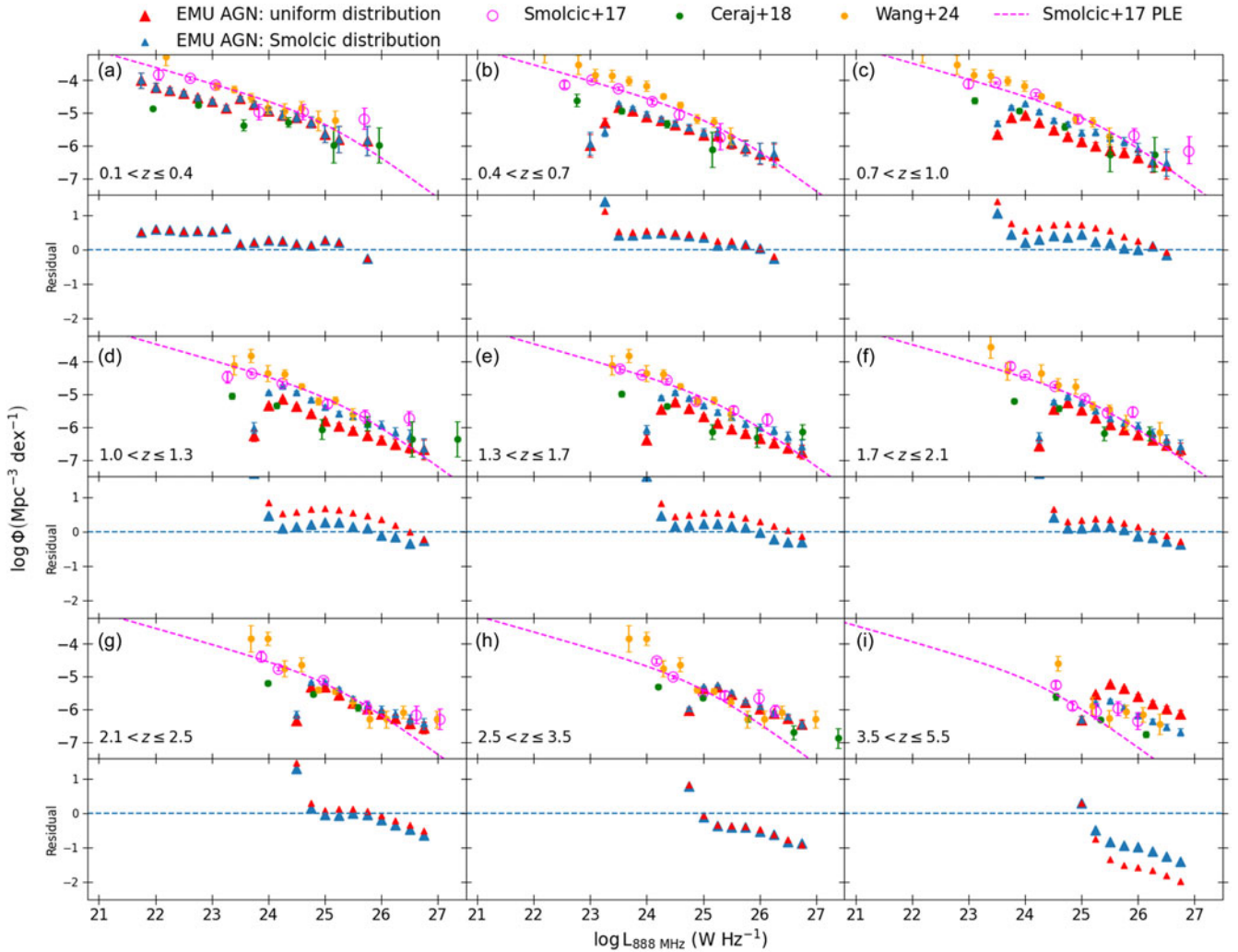
Figure 15. The 888 MHz RLFs derived using Smolčić et al. (2017) distribution (blue triangles), uniform distribution (red triangles), and 1.4 GHz RLFs derived by Smolčić et al. (2017) (magenta dots). The magenta dashed line shows the analytical fit to Smolčić et al. (2017) data which undergoes a pure luminosity evolution (PLE). The rows labelled 'Residual' show the difference between the Smolčić et al. (2017) fit to the LFs derived in this work (both red and blue points). Various results from the literature, converted from 1.4 GHz to 888 MHz assuming $\alpha = -0.7$ are also shown, as noted in the legend, with similar redshift ranges.

evolution, drawing on the large numbers available to investigate dependencies on other galaxy properties, such as stellar mass.

References

Ahmed, U. T., et al. 2024, PASA 41, e021
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP 93, 5
Baum, W. A. 1957, AJ 62, 6
Bellstedt, S., et al. 2020, MNRAS 496, 3235

Best, P. N., & Heckman, T. M. 2012, MNRAS 421, 1569

Best, P. N., Kauffmann, G., Heckman, T. M., & Ivezi, Ž. 2005, MNRAS 362, 9

Bouwens, R. J., et al. 2015, ApJ 803, 34

Brown, M. J. I., Webster, R. L., & Boyle, B. J. 2001, AJ 121, 2381

Butler, A., et al. 2019, A&A 625, A111

Carvajal, R., et al. 2021, Galaxies 9, 86

Cavuoti, S., et al. 2017, MNRAS 465, 1959

Ceraj, L., et al. 2018, A&A 620, A192

Ching, J. H. Y., et al. 2017, MNRAS 464, 1306

Condon, J. J. 1989, ApJ 338, 13

Condon, J. J. 1992, ARA&A 30, 575

Condon, J. J., Cotton, W. D., & Broderick, J. J. 2002, AJ 124, 675

Condon, J. J., Matthews, A. M., & Broderick, J. J. 2019, ApJ 872, 148

Cover, T., & Hart, P. 1967, IEEE TIT 13, 21

Cowley, M. J., et al. 2016, MNRAS 457, 629

Croom, S. M., et al. 2004, MNRAS 349, 1397

Curran, S. J. 2022, MNRAS 512, 2099

de Jong, J. T. A., et al. 2015, A&A 582, A62

Driver, S. P., et al. 2011, MNRAS 413, 971

Driver, S. P., et al. 2022, MNRAS 513, 439

Duncan, K. J., Jarvis, M. J., Brown, M. J. I., & Röttgering, H. J. A. 2018a, MNRAS, 477, 5177

Duncan, K. J., et al. 2018b, MNRAS, 473, 2655

Duncan, K. J., et al. 2021, A&A 648, A4

Dunlop, J. S., & Peacock, J. A. 1990, MNRAS 247, 19

Edge, A., et al. 2013, Msngr 154, 32

Felten, J. E. 1976, ApJ 207, 700

Felten, J. E. 1977, AJ 82, 861

Ferland, G. J., & Netzer, H. 1983, ApJ 264, 105

Fontanot, F., et al. 2007, A&A 461, 39

Geller, M. J., Diaferio, A., Kurtz, M. J., Dell'Antonio, I. P., & Fabricant, D. G. 2012, AJ 143, 102

Gordon, Y. A., et al. 2017, MNRAS 465, 2671

Gunawardhana, M. L. P., et al. 2015, MNRAS 447, 875

Gürkan, G., et al. 2022, MNRAS 512, 6104

Hardcastle, M. J., et al. 2019, A&A 622, A12

Ho, T. K. 1995, in Proceedings of 3rd International Conference on Document Analysis and Recognition Vol. 1, 278

Hogg, D. W., Baldry, I. K., Blanton, M. R., & Eisenstein, D. J. 2002, arXiv e-prints, astro

Hopkins, A. M. 2004, ApJ 615, 209

Hopkins, A. M., et al. 2003, ApJ 599, 971

Hopkins, A. M., et al. 2025, arXiv e-prints, arXiv:2505.08271

Hotan, A. W., et al. 2021, PASA 38, e009

Hurley-Walker, N., et al. 2022, PASA 39, e035

Ilbert, O., et al. 2009, ApJ 690, 1236

Jackson, C. A., & Londish, D. M. 2000, PASA 17, 234

Johnston, S., et al. 2007, PASA 24, 174

Jose, C., et al. 2024, MNRAS 532, 1504

Juneau, S., Dickinson, M., Alexander, D. M., & Salim, S. 2011, ApJ 736, 104

Karademir, G. S., et al. 2022, MNRAS 509, 5467

Kauffmann, G., et al. 2003, MNRAS 346, 1055

Kelvin, L. S., et al. 2014, MNRAS 439, 1245

Kewley, L. J., et al. 2013, ApJ 774, 100

Kewley, L. J., Heisler, C. A., Dopita, M. A., & Lumsden, S. 2001, ApJS 132, 37

Kiang, T. 1961, MNRAS 122, 263

Leahy, D. A., et al. 2019, PASA 36, e024

Li, C., et al. 2024, AJ 168, 233

Loh, E. D., & Spillar, E. J. 1986, ApJ 303, 154

Longair, M. S. 1966, MNRAS 133, 421

Loveday, J., et al. 2012, MNRAS 420, 1239

LSST Science Collaboration, et al. 2017, arXiv e-prints, arXiv:1708.04058

Luken, K. J., et al. 2023, PASA 40, e039

Magliocchetti, M., Maddox, S. J., Lahav, O., & Wall, J. V. 1998, MNRAS 300, 257

Martin, D. C., et al. 2005, ApJ 619, L1

Mauch, T., & Sadler, E. M. 2007, MNRAS 375, 931

McAlpine, K., Jarvis, M. J., & Bonfield, D. G. 2013, MNRAS 436, 1084

McConnell, D., et al. 2016, PASA 33, e042

Miller, N. A., et al. 2009, AJ 137, 4450

Mobasher, B., Cram, L., Georgakakis, A., & Hopkins, A. 1999, MNRAS 308, 45

Morabito, L. K., et al. 2025, MNRAS 536, L32

Noeske, K. G., et al. 2007a, ApJ, 660, L47

Noeske, K. G., et al. 2007b, ApJ, 660, L43

Norris, R. P., et al. 2011, PASA 28, 215

Norris, R. P., et al. 2021, PASA 38, e046

Novak, M., et al. 2017, A&A 602, A5

Overzier, R. A., Röttgering, H. J. A., Rengelink, R. B., & Wilman, R. J. 2003, A&A 405, 53

Padovani, P., et al. 2011, ApJ 740, 20

Peters, J. 2019, PhD thesis, University of Oxford

Pilbratt, G. L., et al. 2010, A&A 518, L1

PLANCK Collaboration, et al. 2016, A&A 594, A13

Pozzetti, L., Bruzual A., G., & Zamorani, G. 1996, MNRAS 281, 953

Pracy, M. B., et al. 2016, MNRAS 460, 2

Prathap, J., et al. 2024, PASA 41, e016

Prescott, M., et al. 2016, MNRAS 457, 730

Rowan-Robinson, M., Benn, C. R., Lawrence, A., McMahon, R. G., & Broadhurst, T. J. 1993, MNRAS 263, 123

Sadler, E. M., et al. 2002, MNRAS 329, 227

Sadler, E. M., et al. 2007, MNRAS 381, 211

Salvato, M., Ilbert, O., & Hoyle, B. 2019, NatAs 3, 212

Sandage, A., Tammann, G. A., & Yahil, A. 1979, ApJ 232, 352

Saunders, W., et al. 1990, MNRAS 242, 318

Schechter, P. 1976, ApJ 203, 297

Schmidt, M. 1968, ApJ 151, 393

Schmidt, M., & Green, R. F. 1983, ApJ 269, 352

Seymour, N., et al. 2008, MNRAS 386, 1695

Shabala, S. S., Ash, S., Alexander, P., & Riley, J. M. 2008, MNRAS 388, 625

Shimwell, T. W., et al. 2017, A&A 598, A104

Siudek, M., et al. 2024, A&A 691, A308

Smith, M. J., & Geach, J. E. 2023, RSOS 10, 221454

Smolčić, V., et al. 2017, A&A 602, A6

Stasińska, G. 1984, A&AS 55, 15

Stern, D., et al. 2012, ApJ 753, 30

Straatman, C. M. S., et al. 2016, ApJ 830, 51

Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29

Thorne, J. E., et al. 2022, MNRAS 509, 4940

Trouille, L., et al. 2011, ApJ 742, 46

Vázquez-Mata, J. A., et al. 2020, MNRAS 499, 631

Veilleux, S., & Osterbrock, D. E. 1987, ApJS 63, 295

Wang, Y., et al. 2024, A&A 685, A79

Webb, S. A., & Goode, S. R. 2023, arXiv e-prints, arXiv:2304.00512

White, R. L., et al. 2000, ApJS 126, 133

Willott, C. J., Rawlings, S., Blundell, K. M., Lacy, M., & Eales, S. A. 2001, MNRAS 322, 536

Windhorst, R. A., Miley, G. K., Owen, F. N., Kron, R. G., & Koo, D. C. 1985, ApJ 289, 494

Windhorst, R. A., van Heerde, G. M., & Katgert, P. 1984, A&AS 58, 1

Wright, A. H., et al. 2024, A&A 686, A170

Wright, E. L., et al. 2010, AJ 140, 1868