

Research Article

The Evolutionary Map of the Universe: A new radio atlas for the southern hemisphere sky

Andrew Hopkins¹, Anna Kapinska², Joshua Marvil², Tessa Vernstrom^{3,4}, Jordan Collier^{5,6,7}, Ray Norris^{8,9}, Yjan Gordon¹⁰, Stefan Duchesne³, Lawrence Rudnick¹¹, Nikhel Gupta³, Ettore Carretti¹², Craig Anderson¹³, Shi Dai⁸, Gulay Gürkan^{14,15}, David Parkinson¹⁶, Isabella Prandoni¹², Simone Riggi¹⁷, Chandra Shekhar Saraf¹⁶, Yik Ki Ma¹³, Miroslav D. Filipović⁷, Grazia Umana¹⁷, Benedict Bahr-Kalus^{18,19,20}, Bärbel Silvia Koribalski^{7,8}, Emil Lenc⁸, Adriano Ingallinera¹⁷, José Afonso^{21,22}, Adeel Ahmad⁷, Ummee Tania Ahmed²³, Emma Alexander²⁴, Heinz Andernach^{25,26}, Jacobo Asorey²⁷, Andrew J. Battisti^{4,13,28}, Maciej Bilicki²⁹, Andrea Botteon¹², Michael Brown³⁰, Marcus Brüggen³¹, Michael Cowley^{32,33}, Kristen Dage³⁴, Catherine Laura Hale³⁵, Martin Hardcastle³⁶, Roland Kothes^{37,38}, Sanja Lazarević^{7,8,39}, Yen-Ting Lin⁴⁰, Kieran Luken⁷, Jeremy Moss⁴¹, P K Jahang Prathap^{1,28,42}, Syed Faisal ur Rahman^{43,44}, Thomas Reiprich⁴⁵, Christopher Riseley^{12,46,47}, Mara Salvato⁴⁸, Nicholas Seymour³⁴, Stanislav Shabala⁴⁹, Daniel Smith³⁶, Mattia Vaccari^{6,12,50,51}, Jacco Th. van Loon⁵², O. Ivy Ivy Wong^{3,4}, Rami Zainal Alsaberi^{7,53}, Albany Asher^{7,8}, Brianna Ball³⁸, Davi Barbosa^{22,54}, Nadia Biava^{12,25}, Aaron Bradley⁷, Rodrigo Carvajal^{22,54}, Evan J. Crawford⁷, Timothy James Galvin³, Minh Huynh^{3,4}, Denis Leahy⁵⁵, Israel Matute^{21,22}, Vanessa Moss⁸, Ciro Pappalardo^{22,54}, Zachary Smeaton⁷, Velibor Velović⁷ and Tayyaba Zafar¹

¹School of Mathematical and Physical Sciences, 12 Wally's Walk, Macquarie University, Sydney, NSW, Australia, ²National Radio Astronomy Observatory, Socorro, NM, USA, ³Australia Telescope National Facility, CSIRO, Space and Astronomy, Bentley, WA, Australia, ⁴International Centre for Radio Astronomy Research (ICRAR), The University of Western Australia, Crawley, WA, Australia, ⁵Australian SKA Regional Centre, Curtin Institute of Radio Astronomy (CIRA), Bentley, WA, Australia, ⁶Inter-University Institute for Data Intensive Astronomy, Department of Astronomy, University of Cape Town, Cape Town, South Africa, ⁷Western Sydney University, Penrith South DC, NSW, Australia, ⁸Australia Telescope National Facility, CSIRO, Space and Astronomy, Epping, NSW, Australia, ⁹Western Sydney University, Locked Bag 1797, Penrith South DC, NSW, Australia, ¹⁰Department of Physics, University of Wisconsin-Madison, Madison, WI, USA, ¹¹University of Minnesota, Minneapolis, MN, USA, ¹²INAF – Istituto di Radioastronomia, Bologna, Italy, ¹³Research School of Astronomy & Astrophysics, Australian National University, Canberra, ACT, Australia, ¹⁴Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield, UK, ¹⁵Australia Telescope National Facility, CSIRO Space and Astronomy, Bentley, WA, Australia, ¹⁶Korea Astronomy and Space Science Institute, Yuseong-gu, Daejeon, Republic of Korea, ¹⁷INAF – Osservatorio Astrofisico di Catania, Catania, Italy, ¹⁸INAF – Osservatorio Astrofisico di Torino, Pino Torinese, Italy, ¹⁹INFN – Sezione di Torino, Torino, Italy, ²⁰Dipartimento di Fisica, Università degli Studi di Torino, Torino, Italy, ²¹Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, OAL, Lisbon, Portugal, ²²Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal, ²³Australian Astronomical Optics, Macquarie University, Sydney, NSW, Australia, ²⁴School of Physics & Astronomy, University of Leeds, Leeds, UK, ²⁵Thüringer Landessternwarte, Tautenburg, Germany, ²⁶Departamento de Astronomía, Universidad de Guanajuato, DCNE, Guanajuato, GTO, Mexico, ²⁷Departamento de Física Teórica, Centro de Astropartículas y Física de Altas Energías, Universidad de Zaragoza, Zaragoza, Spain, ²⁸ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Canberra, Australia, ²⁹Center for Theoretical Physics, Polish Academy of Sciences, Warsaw, Poland, ³⁰School of Physics & Astronomy, Monash University, Clayton, VIC, Australia, ³¹Hamburger Sternwarte, University of Hamburg, Hamburg, Germany, ³²School of Chemistry & Physics, Faculty of Science, Queensland University of Technology, Brisbane, QLD, Australia, ³³University of Southern Queensland, Centre for Astrophysics, Toowoomba, QLD, Australia, ³⁴International Centre for Radio Astronomy Research (ICRAR) – Curtin University, Perth, WA, Australia, ³⁵Department of Physics, University of Oxford, Oxford, UK, ³⁶Centre for Astrophysics Research, University of Hertfordshire, Hatfield, UK, ³⁷Dominion Radio Astrophysical Observatory, Herzberg Research Centre for Astronomy and Astrophysics, National Research Council Canada, Penticton, BC, Canada, ³⁸Department of Physics, University of Alberta, Edmonton, AB, Canada, ³⁹Astronomical Observatory, Belgrade, Serbia, ⁴⁰Institute of Astronomy and Astrophysics, Academia Sinica (ASIAA), Taipei, Taiwan, ⁴¹School of Chemical and Physical Sciences, Victoria University of Wellington, Kelburn, Wellington,

Corresponding author: Andrew Hopkins; Email: andrew.hopkins@mq.edu.au

Cite this article: Hopkins A, Kapinska A, Marvil J, Vernstrom T, Collier J, Norris R, Gordon Y, Duchesne S, Rudnick L, Gupta N, Carretti E, Anderson C, Dai S, Gürkan G, Parkinson D, Prandoni I, Riggi S, Shekhar Saraf C, Ma YK, Filipović MD, Umana G, Bahr-Kalus B, Koribalski BS, Lenc E, Ingallinera A, Afonso J, Ahmad A, Ahmed UT, Alexander E, Andernach H, Asorey J, Battisti AJ, Bilicki M, Botteon A, Brown M, Brüggen M, Cowley M, Dage K, Hale CL, Hardcastle M, Kothes R, Lazarević S, Lin Y-T, Luken K, Moss J, Prathap PKJ, ur Rahman SF, Reiprich T, Riseley C, Salvato M, Seymour N, Shabala S, Smith D, Vaccari M, van Loon JT, Wong OII, Zainal Alsaberi R, Asher A, Ball B, Barbosa D, Biava N, Bradley A, Carvajal R, Crawford EJ, Galvin TJ, Huynh M, Leahy D, Matute I, Moss V, Pappalardo C, Smeaton Z, Velović V and Zafar T. (2025) The Evolutionary Map of the Universe: A new radio atlas for the southern hemisphere sky. *Publications of the Astronomical Society of Australia* 42, e071, 1–28. <https://doi.org/10.1017/pasa.2025.10042>

New Zealand, ⁴²Macquarie University Astrophysics and Space Technologies Research Centre, Sydney, NSW, Australia, ⁴³Lahore University of Management Sciences (LUMS), Lahore, Pakistan, ⁴⁴NCBC at NED University of Engineering and Technology, Karachi, Pakistan, ⁴⁵Argelander Institute for Astronomy, University of Bonn, Bonn, Germany, ⁴⁶Astronomisches Institut der Ruhr-Universität Bochum (AIRUB), Universitätsstraße 150, Bochum, Germany, ⁴⁷Dipartimento di Fisica e Astronomia, Università degli Studi di Bologna, Bologna, Italy, ⁴⁸Max-Planck Institute for Extraterrestrial Physics, Garching, Germany, ⁴⁹School of Natural Sciences, University of Tasmania, Hobart, TAS, Australia, ⁵⁰Inter-University Institute for Data Intensive Astronomy, Cape Town, South Africa, ⁵¹Department of Physics and Astronomy, University of the Western Cape, Bellville, Cape Town, South Africa, ⁵²Lennard-Jones Laboratories, Keele University, Keele, UK, ⁵³Faculty of Engineering, Gifu University, Gifu, Japan, ⁵⁴Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, OAL, Tapada da Ajuda, Lisbon, Portugal and ⁵⁵Department of Physics and Astronomy, University of Calgary, Calgary, AB, Canada

Abstract

We present the Evolutionary Map of the Universe (EMU) survey conducted with the Australian Square Kilometre Array Pathfinder (ASKAP). EMU aims to deliver the touchstone radio atlas of the southern hemisphere. We introduce EMU and review its science drivers and key science goals, updated and tailored to the current ASKAP five-year survey plan. The development of the survey strategy and planned sky coverage is presented, along with the operational aspects of the survey and associated data analysis, together with a selection of diagnostics demonstrating the imaging quality and data characteristics. We give a general description of the value-added data pipeline and data products before concluding with a discussion of links to other surveys and projects and an outline of EMU's legacy value.

Keywords: Sky surveys; galaxies; milky way; astronomical techniques; catalogues

(Received 20 December 2024; revised 30 April 2025; accepted 11 May 2025)

1. Introduction

1.1 EMU and radio surveys

The ‘Evolutionary Map of the Universe’ (EMU, Norris *et al.* 2011, 2021b)^a is a landmark project to deliver the touchstone radio atlas of the southern hemisphere sky using the Australian Square Kilometre Array Pathfinder (ASKAP) radio telescope (Hotan *et al.* 2021).

There is a long history of pushing radio telescopes to their limits to maximise sky coverage at the best possible sensitivity and resolution to understand the nature and properties of ever fainter radio source populations (e.g. Willis, Oosterbaan, & de Ruiter 1976; Kron, Koo, & Windhorst 1985; Oort 1987; Windhorst *et al.* 1993; Hopkins *et al.* 1998; Gruppioni *et al.* 1999; Prandoni *et al.* 2000; de Vries *et al.* 2002; Hopkins *et al.* 2003; Schinnerer *et al.* 2004; Huynh *et al.* 2005; Simpson *et al.* 2006; Norris *et al.* 2006; Moss *et al.* 2007; Smolčić *et al.* 2017; Sabater *et al.* 2021; Best *et al.* 2023; Hale *et al.* 2024). These efforts have revealed that the bright (flux densities above a few mJy) extragalactic radio source population is composed primarily of systems powered by supermassive black holes (e.g. White *et al.* 2020a,b), both nearby and extending to the highest redshifts, while the fainter sources are dominated by star-forming galaxies and low luminosity or ‘radio quiet’ (RQ) active galactic nuclei (AGN) systems (e.g. Seymour *et al.* 2008; White *et al.* 2015, 2017; Prandoni *et al.* 2018; Pennock *et al.* 2021; Drake *et al.* 2024). Wide area sky surveys (e.g. Becker, White, & Helfand 1995; Condon *et al.* 1998; Mauch *et al.* 2003; Intema *et al.* 2017; Hurley-Walker *et al.* 2017, 2022b; Shimwell *et al.* 2022) have illustrated the complexity of radio emission associated with radio galaxy jets and lobes (e.g. Gürkan *et al.* 2022; Koribalski *et al.* 2024a), large scale structures in galaxy clusters (e.g. Giovannini, Tordi, & Feretti 1999; Kempner & Sarazin 2001; Duchesne *et al.* 2021b, 2024; Botteon *et al.* 2022), and from supernova remnants (SNRs), neutral atomic hydrogen (H I) emission, and other objects and structures in the Galactic

Plane (e.g. Umana *et al.* 2021; Filipović *et al.* 2023; Lazarević *et al.* 2024b; Smeaton *et al.* 2024b) and Magellanic Clouds (e.g. Pennock *et al.* 2021).

Each new generation of radio telescope technology and improved radio survey scale has enhanced our understanding of the Universe. These developments continued in the late 2000s and early 2010s with extensive preparation worldwide for major projects anticipating the advent of Square Kilometre Array (SKA) precursor facilities (Norris *et al.* 2013), and ultimately the SKA itself. In the northern hemisphere, the Low-frequency Array (LOFAR), has pushed the limits at low frequencies (Sabater *et al.* 2021; Shimwell *et al.* 2022; Sweijen *et al.* 2022; de Gasperin *et al.* 2023; Groeneveld *et al.* 2024; de Jong *et al.* 2024). A key driver for these was the capability of such new facilities to move beyond the practical limitations of then-existing telescopes (e.g. Norris *et al.* 2013; Norris 2017b). In addition to the many planned scientific developments that such major projects could achieve, it has long been established that expanding the available observational parameter space in this fashion leads to new discoveries beyond just those that can be foreseen (Ekers 2009; Norris 2017a). It was in this environment and with this sense of excitement and anticipation that the original concept for the EMU project was formed.

1.2 EMU history

EMU was conceived in 2009, crystallising earlier ideas around the concept of a maximal area highly-sensitive radio continuum science project with the ASKAP radio telescope (Johnston *et al.* 2007, 2008). EMU was one of two concepts equally ranked in that year as the highest priority projects that ASKAP should deliver, the other being the Widefield ASKAP L-band Legacy All-sky Blind survey (WALLABY, Koribalski *et al.* 2020) which is focussed on detecting neutral hydrogen spectral line emission (H I) in the nearby Universe but also delivers deep 1.4 GHz radio continuum data

^aEMU Project page: <https://emu-survey.org/>.

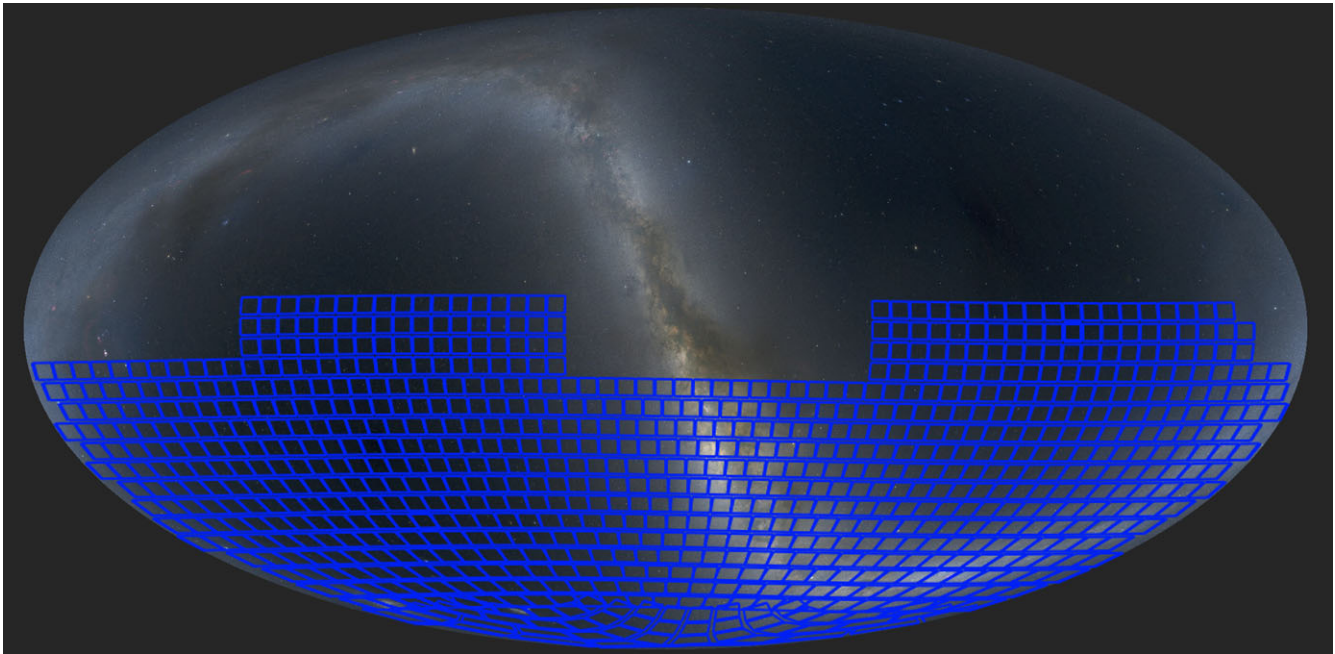


Figure 1. The EMU sky coverage to be delivered in 2028. The background image is the ‘Mellinger coloured’ image (Mellinger 2009) accessed through AladinLite (Boch 2014). Each blue outline represents the footprint of a single ASKAP tile, and there are 853 such footprints comprising the full EMU survey. There is a small overlap between each adjacent tile. North of $\delta = -10^\circ$ each footprint requires two observations, leading to the total of 1 014 tile observations (see Section 3). North of $\delta = -70^\circ$, the tiles follow constant declination strips. Further south, to efficiently cover the pole, the tiles are arrayed in a rectilinear grid centred on the pole.

(Koribalski 2012; Koribalski et al. 2020). Closely linked was the Polarisation Sky Survey of the Universe’s Magnetism (POSSUM) survey (Gaensler et al. 2010; Purcell et al. 2017; Gaensler et al. 2025, in press) to measure the continuum source polarisation properties and Faraday rotation, to develop the best insights into the role of magnetic fields in the Universe. EMU, WALLABY, and POSSUM together were originally conceived as complementary projects that would be carried out commensally through a single observing program with ASKAP. They each capitalise on ASKAP’s unique phased-array feed receiver technology (DeBoer et al. 2009; Chippendale et al. 2015; Hotan et al. 2021) that allows for a very rapid survey speed, a key development necessary for delivering very sensitive all-hemisphere programs. These surveys will provide targets for the SKA, which will not conduct such all-sky surveys itself.

The original EMU concept, detailed in Norris et al. (2011), was for a 3π sr sky survey from the South Celestial Pole up to $\delta = +30^\circ$ reaching to a root-mean-square (rms) noise level of $\sigma = 10 \mu\text{Jy beam}^{-1}$ with a resolution of $\sim 10''$. Such coverage and sensitivity could deliver a survey cataloguing as many as 70 million radio sources at a frequency of 1.4 GHz. As ASKAP commissioning progressed in the late 2010s it became clear that a more realistic performance goal would be a 3π sr sky survey conducted at a frequency around 900 MHz, and limited to a sensitivity of $\sigma = 20 - 30 \mu\text{Jy beam}^{-1}$ with a resolution of $15''$, a consequence arising from a combination of telescope technical performance, the radio frequency interference (RFI) environment at the telescope site (Lourenço et al. 2024), and practical observation scheduling reasons. The original plan for a fully commensal observing program for EMU and WALLABY also became impractical due to the RFI. Mapping H I in the nearby Universe requires WALLABY

to observe close to 1.4 GHz, while EMU, in order to retain the best sensitivity and survey speed, moved to a lower frequency where ASKAP’s continuum sensitivity is optimal. In parallel, the POSSUM project established a clear desire for commensal observing and data sharing with either or (ideally) both EMU and WALLABY.

A review in 2021 of the ASKAP survey science projects, while the telescope was finalising its commissioning activities, reinforced the strong rankings of EMU (jointly with POSSUM) and WALLABY, and recommended that EMU be awarded a total of 8 533 h of ASKAP observing time over the five-year strategic timeline being considered (originally 2022–2027). In line with this allocation, the original 3π sr survey goal was reduced to a coverage of 2π sr, although the original goal remains as a future ambition for EMU following the initial five-year operational period of ASKAP. With these modified survey goals, EMU now anticipates cataloguing about 20 million extragalactic radio sources. This prediction is derived from the Tiered Radio Extragalactic Continuum Simulation (T-RECS; Bonaldi et al. 2019), a simulation of the radio continuum properties of the two main extragalactic radio populations (AGN and star-forming galaxies) over the 150 MHz to 20 GHz range. The final EMU survey five-year coverage is shown in Fig. 1.

1.3 Current EMU status

ASKAP initiated formal full survey operations in May 2023. EMU observations are projected to be complete in 2028. As of the date of writing (April 2025), out of 1014 total tiles (see details in Section 4), there are 307 (30 %) that have been validated as good

and released through the CSIRO ASKAP Science Data Archive (CASDA)^b (see details in Section 4).

This paper details the scientific motivations for conducting EMU (Section 2), the EMU survey design (Section 3), the observations, data processing and validation (Section 4), EMU source statistics (Section 5), and an overview of the value-added data pipeline (Section 6). This is followed by a discussion of EMU in the wider context, including related survey programs (Section 7). We summarise and present the next steps for EMU in Section 8.

Throughout, where relevant, we assume cosmological parameters of $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ and $\Omega_k = 0$.

2. Scientific goals

EMU science spans a vast range of astrophysics and cosmology, and has remained broadly the same as in the original EMU concept (Norris *et al.* 2011), albeit with some evolution as the fields have progressed over the past decade. Here we review the key scientific areas that EMU is primarily aimed at addressing, while also acknowledging there will be a vast wealth of science supported by the survey that extends well beyond these goals. This extends to an expectation of many legacy science outcomes not even anticipated at this early stage (e.g. Norris 2017a).

2.1 Star-forming galaxies and AGN

EMU will detect about 20 million sources to a 5σ limit of $100 \mu\text{Jy beam}^{-1}$. Of these, about 4–5 million will have radio emission dominated by AGN, while 15–16 million will have emission dominated by star formation (SF), based on predictions from the T-RECS simulations (Bonaldi *et al.* 2019). Both populations will span a significant fraction of the age of the Universe, up to reionisation for radio AGN and the most extreme starbursts. EMU will allow large-scale statistical exploration of the evolution of these populations and how it depends on galaxy mass, environment, SF history, interactions and merger history (e.g. Hopkins 2004; Seymour *et al.* 2008; Davies *et al.* 2017; Novak *et al.* 2017). Massive galaxies appear to form their stars early and quickly, progressively becoming less active after redshift $z \sim 2$, while lower-mass galaxies become dominant at lower redshifts (e.g. Mobasher *et al.* 2009). This evolution is mirrored in the AGN accretion rate, suggesting some feedback mechanism couples AGN to galaxy evolution (e.g. Cowley *et al.* 2016; D’Silva *et al.* 2023). EMU will quantify these effects in detail, by providing a deep homogeneously selected sample of both AGN and SF galaxies over the majority of cosmic history, unbiased by dust obscuration (e.g. Afonso *et al.* 2003).

2.1.1 The obscured universe

In the context of dust obscuration, radio observations such as EMU have a critical role to play in uncovering the highly obscured Universe, in a way that complements optical and ultraviolet (UV) surveys. It has been long established that radio-selected samples contain more heavily obscured systems than optically-selected samples (e.g. Afonso *et al.* 2003). But importantly, a recent EMU analysis (Ahmed *et al.* 2024) demonstrates that even in an optically-selected parent sample, radio-detected galaxies exhibit significantly higher levels of dust obscuration, especially for low-mass, low star formation rate (SFR) galaxies. Such results suggest

that a substantial fraction of the cosmic SFR density and black hole accretion history may be hidden in optically obscured systems. Through EMU’s unparalleled sensitivity over its extensive survey area, we can systematically characterise the dust-obscured universe. This will allow us to investigate how the prevalence and properties of such galaxies evolve with redshift, galaxy mass, and local environment.

2.1.2 The cosmic star formation history

Star forming galaxies identified through EMU will provide an unprecedented view of the cosmic star formation history (CSFH; e.g. Hopkins 2004; Hopkins & Beacom 2006; Novak *et al.* 2017; Driver *et al.* 2018; Cochrane *et al.* 2023) and its link to the cosmic stellar mass density history (CSMH; e.g. Wilkins *et al.* 2008a; Wilkins, Trentham, & Hopkins 2008b). This relies on both accurate AGN classification (e.g. Cid Fernandes *et al.* 2010; Heckman & Best 2014), and robust SFR estimation (e.g. Davies *et al.* 2016, 2017; Brown *et al.* 2017).

To develop improved radio-based SFR calibrations, EMU’s data will be combined with multi-wavelength photometric and spectroscopic data from surveys such as the Galaxy and Mass Assembly (GAMA, Driver *et al.* 2011, 2022) and the Wide-Area Vista Extragalactic Survey (WAVES, Driver *et al.* 2019). Drawing on the latest developments in population synthesis tools, (e.g. Robotham & Bellstedt 2024; Bellstedt & Robotham 2024), AGN contributions can be explicitly accounted for, and radio photometry can be used to refine SFR estimates. In turn, samples with extensive multiwavelength data can be used to inform and improve the calibration of radio luminosities to SFRs for those systems without such extensive supporting measurements. Given EMU’s very large sample sizes, different populations can then be separated by astrophysically relevant quantities, such as stellar mass and environment, to construct the CSFH and CSMH for each subset, to explore the mass and environment dependence of the growth of stellar mass in galaxies, along with its link to AGN. The joint constraint of the CSFH and CSMH can also be used to investigate the cosmic evolution of the stellar initial mass function (Wilkins *et al.* 2008b; Hopkins 2018) separated by such populations.

In cases where radio sources lack counterparts, machine learning (ML; e.g. Luken *et al.* 2023) or statistical techniques (e.g. Prathap *et al.* 2025, in press) can be applied to assign redshifts and to infer SFRs and stellar masses probabilistically, enabling population-based analyses even for radio sources without counterparts. This comprehensive effort will involve significant development work on SFR calibrations, AGN/SF diagnostic techniques, ML applications, and redshift and stellar mass assignments, further advancing these fields in addition to our understanding of the cosmic SF and mass history through EMU.

2.1.3 The AGN and star formation link

Construction of obscuration independent samples of AGN hosts is vital in order to robustly establish the AGN duty cycle and its links to SF, the relative timing of AGN and SF activity in galaxies exhibiting both phenomena (Schawinski *et al.* 2007; Wild, Heckman, & Charlot 2010; Shabala *et al.* 2012, 2017), as well as to galaxy transitions through post-starburst or ‘green valley’ stages (e.g. Pennock *et al.* 2022), allowing for a comprehensive overview of galaxy evolution. Complementary multiwavelength photometry and redshifts for all elements of these analyses are critical, not only to supplement EMU detections but also to identify AGN

^b<https://research.csiro.au/casda/>.

hosts that are not dominated by radio emission, such as through X-rays and infrared. Consequently, maximising the EMU survey area is necessary to encompass key complementary surveys in different areas of the sky. It is equally critical to maximise the sample numbers, especially at high- z , due to the inevitable reduction in sample size necessary when measuring evolutionary effects. This arises from (1) limited numbers of counterparts identified in complementary surveys, and (2) construction of luminosity- or mass-limited subsamples split by redshift, mass, environment, galaxy type, and more. This is compounded when the necessary complementary data (different for different types of analyses) only exist over limited regions of sky.

A crucial piece of the galaxy evolution puzzle is the role played by AGN in regulating SF in the host galaxy, commonly referred to as AGN feedback (Croton et al. 2006). While the NRAO VLA sky survey (NVSS, Condon et al. 1998), the Faint Images of the Radio Sky at Twenty centimeters (FIRST, Becker, White, & Helfand 1995) and the Rapid ASKAP Continuum Survey (RACS, McConnell et al. 2020; Hale et al. 2021) are largely dominated by bright radio galaxy populations, EMU is ideal for studying the low luminosity tail of radio-loud (RL) AGN, as well as the radio-quiet (RQ) AGN population, which become significant at $S_{1.4\text{ GHz}} < 100 - 200 \mu\text{Jy}$ (Bonzini et al. 2013). RQ AGN show signatures of nuclear activity at optical, infrared or X-ray bands (e.g. Best et al. 2023; Das et al. 2024; Drake et al. 2024), but are comparatively faint in the radio domain. To identify the RQ AGN population, EMU can rely on spectroscopy from GAMA, Sloan Digital Sky Survey (SDSS, Kollmeier et al. 2019), Dark Energy Spectroscopic Instrument (DESI, Aghamousa et al. 2016), the William Herschel Telescope Enhanced Area Velocity Explorer instrument (WEAVE, Dalton et al. 2012) and soon also 4MOST surveys including WAVES, 4MOST Hemisphere Survey (4HS, Taylor et al. 2023) and the Optical Radio Continuum and H I Deep Spectroscopic Survey (ORCHIDSS, Duncan et al. 2023), as well as WEAVE-LOFAR (Smith et al. 2016) and *Euclid* (Scaramella et al. 2022). Photometry in the optical, infrared, and X-ray will also be obtained from telescopes and surveys such as *Gaia* (Gaia Collaboration et al. 2016, 2023), the Dark Energy Survey (DES, Dark Energy Survey Collaboration et al. 2016), *Herschel* (Pilbratt et al. 2010) and *eROSITA* (Predehl et al. 2021) respectively.

This is illustrated by extensive work drawing on the EMU Early Science observations of the GAMA 23 region, covering an 80 deg^2 area (Gürkan et al. 2022), which links EMU and GAMA data. This resource has already been used extensively to explore aspects of AGN (Prathap et al. 2024), SF (Ahmed et al., submitted), and dust obscuration in galaxies (Ahmed et al. 2024). Gürkan et al. (2022) used the GAMA spectroscopy to identify RQ AGN, marking the first significant work in this domain. Candini et al. (in preparation) are expanding on earlier work (Mullaney et al. 2013) using SDSS and NVSS that show that bright ($L[\text{O III}] > 10^{42} \text{ erg s}^{-1}$) AGN characterised by radio luminosities ($L[1.4 \text{ GHz}] > 10^{23} \text{ W Hz}^{-1}$) tend to have larger [O III] line width. This suggests that [O III] outflows may be linked to the presence of a radio AGN in bright systems, consistent with the increased fraction of high excitation radio galaxies (HERGs) at the highest radio luminosities (Best & Heckman 2012). Despite their relative rarity, these bright radio galaxies dominate the kinetic feedback budget from AGN (Turner & Shabala 2015; Hardcastle et al. 2019).

Radio observations of AGN at low redshift are extensive compared to the scarcity at high redshift. Only a handful are known at $z > 6$, with the highest-redshift radio source at $z = 6.8$ (Bañados et al. 2021) and with the *James Webb Space Telescope* (JWST)

finding a strong candidate at $z \sim 7.7$ (Lambrides et al. 2024). Many more high-redshift AGN are likely to be among the EMU sources (Shobhana et al. 2023), but cannot yet be identified because their redshifts have not been measured. Indeed, RACS (McConnell et al. 2020; Hale et al. 2021) has already found detections of new RL quasi-stellar objects (QSOs) at $z > 6$ (Ighina et al. 2021). AGN near the end of cosmic reionisation contain the (progenitors of the) highest-mass black holes and are the youngest radio sources in the Universe. High- z AGN are critical in understanding the growth of black holes in the early Universe (through rapid accretion), feedback on host galaxies, and their luminosity function and spatial distribution are important for cosmological studies. While extreme dust obscuration and neutral hydrogen (H I) absorption in such systems pose a significant challenge for optical, infrared and X-ray observations, they are transparent to radio emission. Radio observations can detect dust-obscured young AGN in the transition phase of galaxy evolution (from post-merger to quasar), and are a necessary probe of high-redshift galactic environment. A key question is where the radio emission of these high-redshift infrared-luminous sources arises. It may be attributed to a weak jet, quasar winds, disk winds, nuclear starbursts, or something else (e.g. Panessa et al. 2019). Because of the relatively low radio luminosity and surface density on the sky of these high- z sources, highly sensitive and wide-area observations, like those provided by surveys such as EMU and LOFAR, are needed. Low frequency southern hemisphere radio data are also available from the Murchison Wide-field Array (MWA) surveys GaLactic and Extragalactic All-sky MWA survey (GLEAM, Hurley-Walker et al. 2017) and GLEAM-eXtended (GLEAM-X, Hurley-Walker et al. 2022b; Ross et al. 2024), which, despite having much poorer spatial resolution than EMU, will still be important resources for constraining radio spectral indices for the brighter EMU sources.

2.2 Astrophysics of radio galaxies

By ‘radio galaxies’ here we are referring to those systems with radio jets and lobes associated with an AGN. With EMU’s extremely good surface brightness sensitivity (e.g. Brügggen et al. 2021; Norris et al. 2021c), well-resolved extended emission from such radio galaxies will be detected for several 10^5 sources, based on extrapolations from the numbers found in the first EMU Pilot Survey (Norris et al. 2021b). Such large samples of radio galaxies will change our understanding of their overall structure. Detailed studies of their extended structure with EMU (Velović et al. 2022) will influence models of the powering jets as well as the interactions with the surrounding medium (e.g. English, Hardcastle, & Krause 2016; Yates-Jones, Shabala, & Krause 2021; Yates-Jones et al. 2023). In addition, unusual structures which challenge our models are important but rare (e.g. Koribalski et al. 2024a,c), and can only be found by surveying sufficiently large areas of sky with superb surface brightness sensitivity. Spectral index maps for such extended objects can be derived from EMU data, and will be available in unprecedented numbers, enabling the history of relativistic particle gains and losses in radio galaxies to be systematically studied.

EMU’s excellent surface brightness sensitivity ensures efficient detection of giant radio galaxies, including those of relatively low luminosities. Giant radio galaxies are examples of both extreme jet physics and low density environments, making them a valuable probe of astrophysics in extreme conditions. Expanding on recent findings with RACS and LOFAR data (e.g. Andernach, Jiménez-Andrade, & Willis 2021; Oei et al. 2023; Mostert et al.

2024), they will be found by EMU in significant numbers (e.g. Quici *et al.* 2021; Gürkan *et al.* 2022; Simonte *et al.* 2024, Kataria *et al.* in preparation). Having large samples of extreme systems in both hemispheres will be important for exploring any large-scale cosmological implications.

Dying radio galaxies are important for understanding the duty cycle of AGN (Shabala *et al.* 2020) and, when relaxed, are a good probe of surrounding pressures (Murgia *et al.* 2011; Yates, Shabala, & Krause 2018; English *et al.* 2019). Radio remnant morphologies are also excellent probes of jet and environment dynamics, but require excellent surface brightness sensitivity due to the rapid fading of remnant lobes (e.g. Yates-Jones *et al.* 2023; Riseley *et al.* 2025, Stewart *et al.* in preparation). EMU will provide the best constraints on any current jet or hot spot activity in such systems. The bending and distortion of radio galaxy morphologies is an excellent probe of diffuse gas, providing ideal laboratories to explore the physical conditions in the outskirts of clusters, poor groups, and cosmic filaments. The demonstrated surface brightness sensitivity of EMU is critical for these studies, as is a large sky area due to the rarity of key source populations with short fractional lifetimes.

2.3 Galaxy clusters and large scale structure

Galaxy clusters present opportunities to study large-scale structure evolution, turbulence, cosmic rays, shocks, feedback, and more. They evolve and grow through a variety of processes, including passive accretion of gas, consumption of small galaxy groups, and violent merger events. Many clusters host vast and enigmatic diffuse radio continuum sources such as giant and mini radio halos, radio relics from merger shocks, and numerous AGN and remnant radio galaxies interacting with the intracluster environment (van Weeren *et al.* 2019). Structures on the largest scales, superclusters, are also a likely source of radio emission from filaments and from various cluster merger signatures (Veronica *et al.* 2022; Paul *et al.* 2023). EMU has already produced results studying such features, with Hyeong-Han *et al.* (2020) and Duchesne *et al.* (2021a) exploring diffuse radio sources in the massive merging clusters SPT-CL J2023–5535 and SPT-CL J2032–5627, respectively, Di Mascolo *et al.* (2021) investigating the possibility of a halo in SPT-CL J2106–5844 to extend the detection of radio emission in clusters to $z \gtrsim 0.8$, and Loi *et al.* (2023) discovering an unusual extended radio arc in Abell 3718. Böckmann *et al.* (2023) found a correlation between the radio luminosity of cluster central AGN and the X-ray luminosity of the clusters in the first EMU Pilot Field. Early EMU/*eROSITA* results on the Abell 3391/95 galaxy cluster system (Reiprich *et al.* 2021a; Brügggen *et al.* 2021) have helped constrain physical processes in the merger. Detailed analyses of ‘clumps’ in the filaments discovered in the Abell 3391/95 system show the influence of the bright wide angle tailed central galaxy (Veronica *et al.* 2022). The complex interplay of radio sources in the spectacular merger systems Abell 3266 and Abell 3627 has been constrained using EMU data (Riseley *et al.* 2022; Koribalski *et al.* 2024a). Macgregor *et al.* (2024) have mapped the cluster emission in Abell S1136 using EMU Early Science data, finding that the diffuse emission breaks up into filaments when seen with the sensitivity and resolution of ASKAP. Diffuse radio sources in clusters are known to have very steep spectra, with spectral indices^c ranging from $\alpha \sim -1$ to -3 depending on the type of source (halos, relics, radio remnants), and it is anticipated that a large number of

radio halos and relics will be detected by EMU (e.g. Cassano *et al.* 2012; Nuza *et al.* 2017; Nishiwaki & Asano 2022; Duchesne *et al.* 2024)

Filaments of the cosmic web exist on scales larger than clusters, and are now being mapped using weak lensing (Hyeong-Han *et al.* 2024). Galaxy populations inside and outside of filaments have been compared using stacking analyses (e.g. Kleiner *et al.* 2017), but detecting radio emission from the filaments themselves is an emerging field. With the new generation of telescopes, faint bridges of diffuse emission are now becoming detectable in pairs of pre-merging galaxy clusters (e.g. Govoni *et al.* 2019; Botteon *et al.* 2020; de Jong *et al.* 2022; Balboni *et al.* 2023; Pignataro *et al.* 2024). EMU Early Science data (Venturi *et al.* 2022) show that detecting emission extending beyond the central regions of clusters with ASKAP is possible. EMU will enable sensitive studies of the synchrotron cosmic web.

While other instruments, such as the MeerKAT radio telescope (Jonas & MeerKAT Team 2016), will have comparable sensitivity to diffuse cluster phenomena, they will focus only on targeted regions, sometimes including complementary ASKAP data (e.g. Koribalski *et al.* 2024b). The large area proposed for EMU will ensure coverage of a vastly greater number of clusters, and in regions unmatched by such targeted surveys. Only around 150 clusters have so far been found to host diffuse radio sources, such as halos and relics (see, e.g. Botteon *et al.* 2022; Knowles *et al.* 2022; Duchesne *et al.* 2021b, 2024, for recent large collections). EMU allows for an increase in this number to statistically significant samples (thousands) through maximising the sky area covered, along with its excellent spectral index precision and sensitivity to diffuse emission.

2.4 Cosmology and fundamental physics

EMU has the capability to provide important tests of the fundamental physics of the Universe, including questions such as the nature of dark matter (Regis *et al.* 2021), the nature of the mysterious force accelerating the expansion of the Universe (dark energy), and the mechanism that generates the initial conditions, both of which are currently unknown. Dark energy is an established part of the cosmological model, but the evidence at low-redshift comes mainly from standard candles and standard rulers. Radio galaxies from EMU can be used to trace the distribution and the evolution of the gravitational potential at higher redshifts through their clustering statistics. By cross-correlating the EMU galaxy distribution with the cosmic microwave background (CMB) the late-time integrated Sachs-Wolfe (ISW) effect (Crittenden & Turok 1996; Raccanelli *et al.* 2012; Stözlner *et al.* 2018; Bahr-Kalus *et al.* 2022) can be measured (Fig. 2). Cross-correlating with galaxy surveys at lower redshifts, including samples drawn from the *Wide-field infrared survey explorer* (WISE, Wright *et al.* 2010), VISTA hemisphere survey (VHS, McMahon *et al.* 2013), Dark Energy Survey (DES, Dark Energy Survey Collaboration *et al.* 2016), Kilo-Degree Survey (KiDS, de Jong *et al.* 2015), and the Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration *et al.* 2016), further allows cosmic magnification (Scranton *et al.* 2005) to be detected and used. These data will thus enable a highly sensitive test of General Relativity at large scales. Either of these cross-correlation approaches also allows for the characterisation of the redshift distributions and bias of the radio source population (Alonso *et al.* 2021). The cross-correlation of the first EMU Pilot galaxy sample with CMB lensing is an illustration of what the full EMU survey will enable for such characterisation studies, and for

^cWe adopt the convention that spectral index, α , is related to flux density, S , and frequency, ν , through $S \propto \nu^\alpha$.

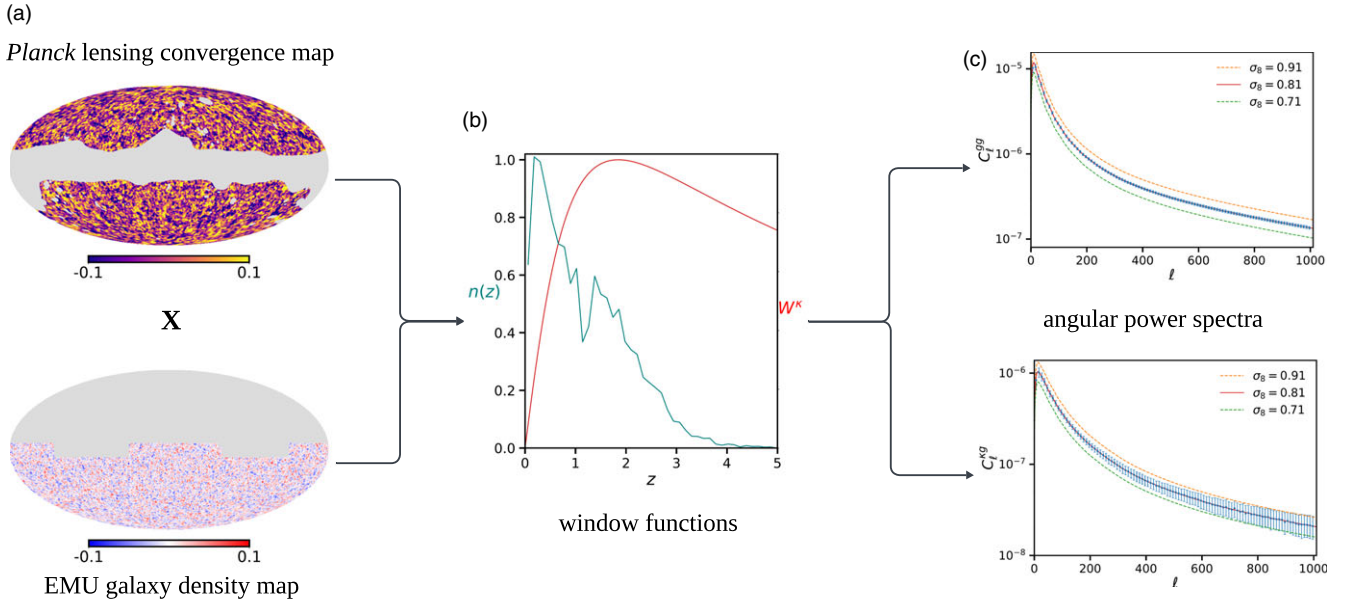


Figure 2. A schematic demonstrating how the EMU galaxy density map (a) can be used to measure the auto angular power spectra, and also cross-correlated with other large-scale structure maps (in this case the Planck CMB lensing convergence map, denoted X). Combining this with information about the redshift distribution $n(z)$ we can compute the window function (b), and compare the measured auto- and cross-power spectra with their theoretical predictions to constrain the cosmological parameters (in this case the amplitude of the density perturbations σ_8).

the measurement of the growth of structures (Tanidis et al. 2024), complementary to weak lensing surveys. EMU data can also be used to expand on an analysis of the cosmic dipole (Oayda et al. 2024) that suggests a tension between the dipoles inferred from the radio source distribution of NVSS and RACS, and the kinematic dipole of the CMB.

The EMU sample will also provide key information to test the inflationary theory by determining the large-scale Gaussianity of the initial distribution of structures (Raccanelli et al. 2017; Bernal et al. 2019). These clustering statistic approaches are complementary to other established cosmological probes (e.g. CMB, type-Ia supernova, and baryon acoustic oscillations and galaxy clusters). However, they require a large contiguous area and high source density for uniform sampling. The limitations of future continuum clustering surveys can be related to the low density of AGNs or intrinsic confusion noise (Asorey & Parkinson 2021). The cosmic radio dipole, caused by our peculiar motion with respect to the rest frame where EMU galaxies are statistically isotropic, also contributes strongly to the large-scale clustering signal and can mimic that of a non-Gaussian initial distribution of structures (Chen & Schwarz 2016). Previous measurements of the cosmic radio dipole direction (Blake & Wall 2002; Singal 2011; Gibelyou & Huterer 2012; Tiwari & Nusser 2016; Siewert, Schmidt-Rubart, & Schwarz 2021) are in agreement with the expectation from the CMB (Planck Collaboration et al. 2020), but their amplitudes are in tension. However, the dipole amplitude from recent MeerKAT absorption line survey data agrees with the CMB after including sub-mJy sources (Wagenveld et al. 2024). The sensitivity of EMU will allow testing of this result using a ~ 5 times larger survey area.

The density of sources in EMU, a factor of 10 or more greater compared to NVSS and RACS, will be significant in addressing other unresolved cosmological questions, with one such being the origin of the CMB cold spot region ($\alpha = 03^h \sim 15^m \sim 05^s$,

$\delta = -19^\circ \sim 35' \sim 02''$; Rudnick, Brown, & Williams 2007; Smith & Huterer 2010; ur Rahman 2020). There is a statistically significant underdensity in radio source numbers and surface brightness in the region of the cold spot, shown by Rudnick et al. (2007) using NVSS data. They argue that this implies it is cosmologically local, resulting from a localised manifestation of the late-time ISW effect. With increased source density, and redshift estimates for many EMU sources, this result can be explored in more detail to identify the extent, and potentially the redshift localisation, of any underdensity, and the likelihood of association with the origin of the cold spot.

2.5 The galaxy and magellanic clouds

EMU will create the most sensitive wide-field atlas of Galactic continuum emission in the southern hemisphere, along with some of the most sensitive maps of the Magellanic Clouds, allowing the study of the formation and evolution of stars in exquisite detail. EMU's observations of the Magellanic Clouds will complement existing MWA (For et al. 2018), Australia Telescope Compact Array (ATCA) higher-frequency (Filipović et al. 2002; Crawford et al. 2011) and similar-sensitivity MeerKAT observations (Cotton et al. 2024; Carli et al. 2024; Sasaki et al. 2025), extending measurements of spectral indices for all detected sources.

EMU will observe the full range of Galactic latitudes, extending well beyond the few degrees on either side of the plane to which many Galactic surveys are limited. EMU's sensitivity, especially to extended emission, is needed to increase samples of supernova remnants (Filipović et al. 2022; Bozzetto et al. 2023; Smeaton et al. 2024b; Zangrandi et al. 2024), planetary nebulae (Asher et al. 2024), and the newly-detected association of low surface brightness H II emission with reflection nebulae such as Lagotis (Bradley et al. 2025), which together support the study of the end stages

of stellar evolution. EMU will reveal all the stages in the evolution of a compact H II region (hypercompact, ultracompact and compact).

Early science and pilot data have shown EMU’s potential for these studies (e.g. Joseph et al. 2019; Pennock et al. 2021; Filipović et al. 2021; Umana et al. 2021; Ball et al. 2023). One highlight is the unexpected detection of asymptotic giant branch (AGB) stars in EMU’s wide-band images through their OH maser emission, despite these objects lacking continuum emission (Ingallinera et al. 2022). The large fractional bandwidth and resolution of EMU, coupled with infrared surveys of the Galactic Plane (VISTA VVV, *Spitzer* GLIMPSE and MIPS GAL, *Herschel* Hi-GAL) allow us to distinguish thermal radio emitters (H II regions, planetary nebulae) from non-thermal (supernova remnants, pulsar wind nebulae, pulsars, active stars). High-energy observations (both gamma rays from High Energy Stereoscopic System (HESS) and the *Fermi* Gamma-ray Space Telescope, and X-rays from *Chandra* X-ray Observatory, *XMM-Newton*, the Neil Gehrels *Swift* Observatory, and the *eROSITA* all-sky survey) also provide strong synergies. ASKAP short baselines recover spatial scales up to ≈ 45 arcmin at ≈ 1 GHz, a unique capability among interferometers, making EMU ideal to study large and complex Galactic structures (e.g. Umana et al. 2021; Filipović et al. 2024).

The Galactic science goals of EMU include: (1) a complete census of the early stages of massive SF in the southern Galactic Plane; (2) detection and characterisation of a significant number of missing supernova remnants up to the edge of the Galactic disk (~ 300 known, up to 2 000 expected, e.g. Ranasinghe & Leahy 2022; Ball et al. 2023; Filipović et al. 2023; Lazarević et al. 2024a; Smeaton et al. 2024a,b; Green 2024; Bufano et al. 2024; Jing et al. 2025); (3) understanding the complex structures of giant H II regions and the inter-relationship of dust, ionised gas and triggered SF (De Horta et al. 2014; Sano et al. 2017, Bradley et al., in preparation); (4) the variety of luminous blue variables (Bordiu et al. 2024); (5) serendipitous discoveries, such as the radio flares from ultra-cool dwarfs found by Berger et al. (2001): these stars are distributed isotropically in the sky, and EMU is the only wide-area southern survey planned at this frequency and sensitivity; (6) a detailed characterisation of planetary nebulae (spectral energy distributions, distances, ionised mass, Bojičić et al. 2021). Detection of other classes of radio stars, adding to the growing numbers recently identified (Driessen et al. 2023, 2024) will also be enhanced through the sensitivity of EMU.

2.6 Pulsars, variables, and transients

While pulsars are primarily detected and observed with high temporal resolution in order to resolve their pulses, the phase-averaged emissions of pulsars can also be detected in radio continuum surveys. Deep all-sky continuum surveys like EMU enable the study of the spectral, polarisation and scintillation properties of a large sample of radio pulsars (Bell et al. 2016; Murphy et al. 2017; Anumalapudi et al. 2023; Sett et al. 2024). EMU will also enable us to carry out targeted searches for radio pulsars over the whole visible sky while avoiding the need for expensive pixel-by-pixel searches with high temporal resolution, strongly complementing other pulsar surveys in the southern hemisphere with MWA (Bhat et al. 2023a,b), MeerKAT (Padmanabh et al. 2023) and the phased array feed on Murriyang/Parkes (Chippendale et al. 2016; Deng et al. 2017). Continuum surveys are equally sensitive to all pulsars, not affected by the dispersion measure smearing,

scattering or orbital modulation of spin periods. This allows for the discovery of extreme pulsars such as sub-millisecond pulsars, pulsar-black hole systems and potentially also pulsars in the Galactic Centre (e.g. Lower et al. 2024). The capability of finding new pulsars with ASKAP has been demonstrated by the discovery of pulsars originally identified as highly polarised radio sources (Kaplan et al. 2019; Wang et al. 2024a) and pulsars associated with SNRs (e.g. Ahmad et al. 2025) and pulsar wind nebulae (Lazarević et al. 2024b). The deeper observations of EMU have the potential to reveal a large population of these sources, especially millisecond pulsars at intermediate and high Galactic latitudes. To distinguish pulsars from other point sources Dai et al. (2016) developed a formalism for computing variance images from standard interferometric radio images and demonstrated its feasibility. Dai, Johnston, & Hobbs (2017) showed that, with the variance imaging technique alone, EMU should discover ~ 40 new millisecond pulsars and ~ 30 new normal pulsars. Variance imaging with EMU will be more sensitive than current pulsar surveys at high Galactic latitudes.

Searches for radio transients and variable sources have previously been limited by small fields of view and poor sensitivity (e.g. Lazio et al. 2010; Obenberger et al. 2014). The ASKAP telescope is conducting the Variables and Slow Transients (VAST, Murphy et al. 2013, 2021) survey specifically to improve this parameter space. EMU data provide a complementary resource to VAST in identifying transients on shorter timescales. With its large sky coverage, EMU enables a large-scale radio transient and variability search on shorter timescales (< 1 h), largely unexplored before (e.g. An et al. 2023; Wang et al. 2024b). Studies on 15 min timescales using the first EMU Pilot Survey have identified 11 such sources and the full EMU survey may identify up to ~ 1 000 variable sources (Wang et al. 2023). More recently, several ultra-long period (ULP) sources, with periods of several minutes, with repeating bursts of coherent radio emission have been reported (e.g. Caleb et al. 2022; Hurley-Walker et al. 2022a, 2023; Dong et al. 2024). ASKAP has demonstrated its capability to discover such sources (Caleb et al. 2024; Dobie et al. 2024) and the full EMU survey is expected to discover many more.

EMU has also been used to search for a radio-continuum counterpart of the recent ultra-high-energy (UHE) neutrino event, KM3-230213A (KM3NeT Collaboration et al. 2025). Among more than 1 000 radio sources within the 68% confidence region of the UHE neutrino event, three distinctive radio sources, a nearby spiral galaxy (UGCA 127), a radio AGN, and a compact variable radio source (blazar), were found to be possible origins for KM3-230213A (Filipović et al. 2025).

2.7 Resolved galaxies in the local universe

The sensitivity and resolution of the EMU survey enable detailed, spatially-resolved studies of several thousand nearby galaxies ($D \lesssim 50$ Mpc). EMU’s resolution of $\sim 15''$, is well matched to the resolution of the *WISE* W4 band ($22 \mu\text{m}$), providing $\gtrsim 5$ synthesised beams (resolution elements) across galaxies with $M_B < -18$ mag ($\log(M_*/M_\odot) \gtrsim 9$) at $D \lesssim 50$ Mpc (Leroy et al. 2019). The EMU data are well-suited for comparison with existing atlases of thousands of nearby galaxies from all-sky imaging surveys, such as GALEX (UV; e.g. Gil de Paz et al. 2007), 2MASS (near-IR; e.g. Jarrett et al. 2003), and *WISE* (mid-IR; e.g. Jarrett et al. 2019; Leroy et al. 2019). A smaller subset of galaxies (~ 50) could also be compared to optical integral field spectroscopic surveys, such

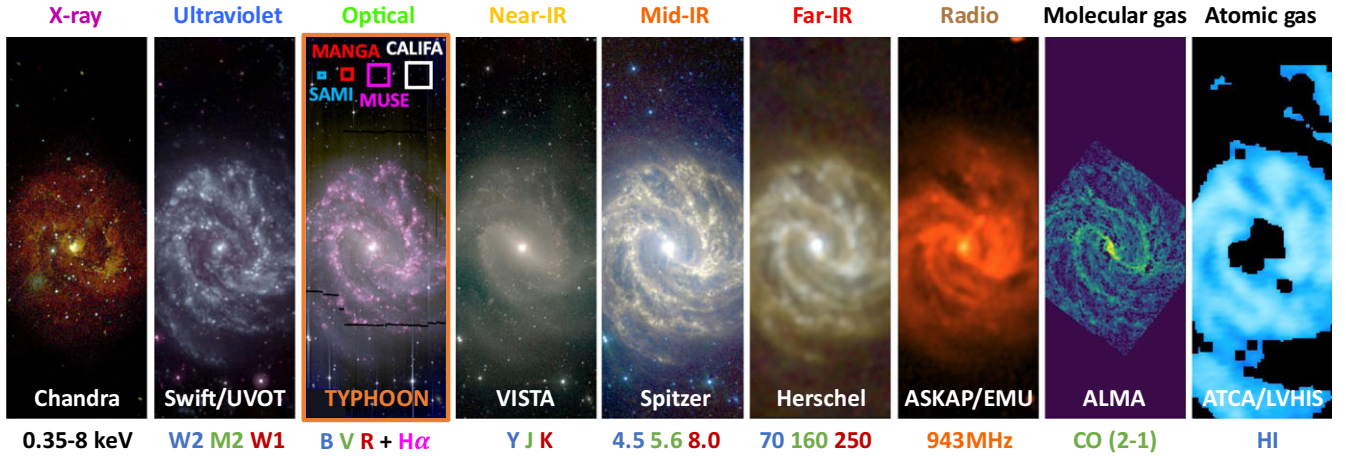


Figure 3. Demonstration of the exquisite ancillary data available for the nearby spiral galaxy M 83. The filters used to create each colour image are indicated on the bottom. The size of the panels were matched to the TYPHOON survey field-of-view (orange box; third panel), which is $\sim 100\text{--}1\,000\times$ larger relative to other optical IFS instruments (smaller boxes). We note a MUSE mosaic of 26 tiles is also available for this galaxy (Della Bruna et al. 2022). The ASKAP/EMU data (seventh panel) are from a single observing block (5 h, half of the total final integration time). Similar nearby galaxies with extensive multiwavelength photometry and spectroscopy offer a unique opportunity for spatially resolved comparisons between all baryonic components within galaxies (stars, ionised gas, molecular gas, atomic gas, and dust). The ancillary data are retrieved from the following: *Chandra* – Harvard/SAO; *Swift/UVOT* – NASA HEASARC service; TYPHOON – priv. comm.; VISTA – ESO science archive; *Spitzer* and *Herschel* – NASA Extragalactic Database; ALMA – PHANGS-ALMA (Leroy et al. 2021); ATCA/LVHIS (Koribalski et al. 2018). Higher resolution H I data from ASKAP/WALLABY for M 83 will be available in the future.

as PHANGS-MUSE (Emsellem et al. 2022), TYPHOON (Seibert et al. in preparation; see Fig. 3), and SDSS-V Local Volume Mapper (LVM, Drory et al. 2024).

In star-forming galaxies, most of the synchrotron emission at ~ 1 GHz comes from cosmic-ray electrons accelerated by SNRs from high mass ($M_\star > 8 M_\odot$) stars, with lifetimes of $\tau \lesssim 50$ Myr. The typical lifetimes of cosmic-ray electrons are $\tau \lesssim 100$ Myr (Condon & Ransom 2016), and as a result, radio continuum emission from normal galaxies traces recent SF on similar timescales. Understanding the exact SF timescale traced by synchrotron emission has been difficult to pin down through global measurements of galaxies (e.g. Cook et al. 2024). By linking resolved EMU data with independent tracers of SF at other wavelengths, such as UV, optical (emission lines), and infrared (Kennicutt & Evans 2012), it is possible to explore this link in fine detail. Closely related to the link with SF, star-forming galaxies also show a well-known correlation between their far-IR and radio emission (e.g. Condon, Anderson, & Helou 1991; Yun, Reddy, & Condon 2001). The physical origin of this correlation is poorly understood and will also be studied in greater detail through such spatially-resolved comparisons of a subset of galaxies with available far-IR data (e.g. Kennicutt et al. 2011).

Additionally, EMU has access to continuum observations with ASKAP Band 2 (at 1.4 GHz), that are commensal with the WALLABY Survey (see Section 7.1.3). Early results from WALLABY’s continuum observations of the Eridanus supergroup demonstrated the potential for the better resolution ($6.1'' \times 7.9''$) continuum observations of Band 2 to be useful for probing the internal physical mechanisms that are occurring in star-forming regions in combination with ancillary observations at shorter wavelengths (Grundy et al. 2023). In the nearby galaxy IC 1952, Grundy et al. (2023) found that the resolved infrared-radio correlation could be used to distinguish between a background AGN (near the plane of the galaxy) and star-forming regions within the galaxy. Disentangling the AGN and SF properties within galaxies

is critical in understanding not only the multiwavelength tracers of SF but also any feedback links between the two.

2.8 Discovering the unexpected

Experience has shown (Norris 2017a) that whenever we observe the sky to a significantly greater sensitivity, or explore a significantly new volume of observational phase space, we make new discoveries. Even the Australia Telescope Large Area Survey (ATLAS) (Norris et al. 2006; Middelberg et al. 2008), which expanded the phase space of sensitive radio surveys by only a factor of a few, discovered two previously unrecognised classes of object (infrared faint radio sources, and radio-loud AGN buried in SF galaxies). With a sensitivity 10–30 times better than other large-scale radio surveys at similar frequencies, EMU is likely to discover new types of object, or new phenomena. While impossible to predict their nature, we might reasonably expect new classes of galaxy or new Galactic populations.

This goal has already been demonstrated in EMU through the successful identification of a new class of radio object, odd radio circles (ORCs, Norris, Crawford, & Macgregor 2021a; Norris et al. 2021c; Koribalski et al. 2021; Gupta et al. 2022). These diffuse circles of radio emission are now known (Norris et al. 2022), at least in some cases, to surround galaxies at $0.2 < z < 0.6$. Other patches of diffuse emission which may be related to ORCs have also been found around galaxies (Koribalski et al. 2024c) or groups of galaxies (Bulbul et al. 2024). The cause of these circles is still debated, but may result from a starburst wind termination shock from the central host galaxy (Norris et al. 2021c), from a spherical shock wave from a cataclysmic event in the central host galaxy (Norris et al. 2021c), from a shock in a galaxy-galaxy merger event (Dolag et al. 2023), or from re-energised electrons in a relic radio lobe (Shabala et al. 2024). The combination of their rarity (about one per 100 deg^2) and low surface brightness explains why they have not been discovered before. We expect to discover about 300 ORCs in the

full EMU survey, making them an important new class of radio objects.

To expedite the discovery of ORCs and other unusual radio morphologies, we have developed ML algorithms that leverage multi-modal foundation models. These models use text and image embeddings to retrieve similar radio sources from the first year of the EMU full survey in under a second. To make these models accessible, we created the EMU Survey Object Search Engine (EMUSE^d, Gupta *et al.*, submitted), which enables users to search for similar objects using text descriptions or image uploads. In another work (Gupta *et al.*, submitted) we present the ORCs and other unusual sources identified using a combination of object detection methods from Gupta *et al.* (2024a) and these foundation models, from data observed during the first year of the full EMU survey.

There are many additional areas in which EMU data will support novel discoveries. These include the detection, measurement, and exploration of binary supermassive black holes; populations of gigahertz peaked spectrum (GPS), compact steep spectrum (CSS), and other young AGN to understand the AGN duty cycle and evolution (e.g. Callingham *et al.* 2017); spatially resolved nearby galaxies to explore SF and extreme stellar evolutionary states; properties of dark matter (Storm *et al.* 2017); the stellar initial mass function and whether, and how, it may evolve (e.g. Hopkins 2018); and more. The discovery of ORCs has demonstrated the value of exploring new regions of observational parameter space. We expect that the full EMU survey will continue to reveal new and rare classes of objects as we maximise the sky areas probed to this unprecedented sensitivity. A few additional examples of objects for which current models are inadequate are presented in Section 6.5.

2.9 Legacy

EMU was always conceived as a project with immense legacy value, in addition to the direct scientific outcomes planned. As now shown through more than three decades of major sky survey projects across the electromagnetic spectrum, each supports a wealth of science beyond that anticipated, and underpins further projects as new complementary surveys arise.

EMU will be the touchstone ~ 1 GHz radio continuum atlas for at least a decade after completion, and possibly beyond. While the SKA's survey plans are not yet finalised, its strengths suggest it may prioritise a wedding-cake-tiered continuum survey program with the widest tier limited to smaller sky areas ($\sim 1\,000$ deg²). Consequently, no planned survey would replace EMU for the foreseeable future. EMU will provide the most uniform deep wide-area radio data at GHz frequencies to complement galaxy and AGN evolution, physical processes and cosmological studies from recent and upcoming surveys and facilities such as *SRG/eROSITA* (Merloni *et al.* 2024; Predehl *et al.* 2021), DES, DESI, GLEAM-X, the LOFAR Two-metre Sky Survey (LoTSS, Shimwell *et al.* 2019), PanSTARRS, Apertif, 4HS, WAVES, the Rubin Observatory's Legacy Survey of Space and Time (LSST, Ivezić *et al.* 2019), *Euclid*, and more. A recent example of this radio, optical, X-ray complementarity using EMU data is the discovery of radio emission from a 15 Mpc filament in the intergalactic medium, a warm gas bridge, infalling matter clumps, and (re-)accelerated plasma in the Abell 3391/95 galaxy cluster system (Reiprich *et al.* 2021b).

EMU is already demonstrating its capacity to serve as a critical and key data resource in ML algorithm development for radio astronomy. The ML applications for EMU range across source finding, host or multiwavelength cross identification, identification of multi-component or giant radio galaxies, image denoising, redshift estimation, and more. Such algorithms will be applicable to many future datasets, especially those delivered by the SKA. In addition to ML, EMU is demonstrating other new software development, and providing an example of unifying survey pipeline, data hosting, value-added data development and delivery, across multiple organisations and technologies. The legacy of software, tools, and data infrastructure and infrastructure linkages developed for the specific needs of EMU will also inform future large scale astronomical developments, in particular for the SKA.

3. Survey design

To determine the survey sky coverage, the EMU and POSSUM teams explored several options that can satisfy all constraints set by the observatory, as well as by both science teams. These constraints are:

- A 2π sr areal coverage, decided by the telescope time awarded to EMU and POSSUM;
- A contiguous sky region, as any missing sky coverage due to having gaps or holes would be problematic for cosmology analyses looking for structures on large angular scales;
- Maximal overlap with equatorial multiwavelength and spectroscopic sky surveys, especially the SDSS Extended Baryon Oscillation Spectroscopic Survey (eBOSS) emission line galaxy (ELG) cosmology programme (Ahumada *et al.* 2020), the GAMA survey (Driver *et al.* 2022), and the WAVES survey (Driver *et al.* 2019);
- Overlap with complementary radio surveys of the northern sky, especially LoTSS and the Very Large Array Sky Survey (VLASS, Lacy *et al.* 2020);
- Coverage of the southern Galactic Plane, as well as the Magellanic System, to enable science at good spatial resolution in our cosmic neighbourhood;
- Coverage of the extreme southern sky ($\delta < -40^\circ$) that cannot be accessed by telescopes in the northern hemisphere, to enable discovery in these less-explored sky areas.

Input from the EMU, POSSUM, and WALLABY teams was solicited, leading ultimately to the current EMU footprint shown in Fig. 1. This option met POSSUM's primary goal of ensuring a large contiguous areal coverage with minimal adjustment for specific targets, as well as EMU's goal of contiguous sky coverage encompassing suitable complementary multiwavelength and spectroscopic datasets (Fig. 4). It also encompasses the full area of the WALLABY survey.

Early in the development of the EMU survey strategy, it was recognised that optimal tiling of the sky would require different choices for the polar cap compared to the bulk of the rest of the survey. For most declinations, bands of tiles at a fixed declination work well in tiling the sky. This approach, however, is not efficient for regions close to the celestial pole. For declinations $\delta < -70^\circ$ a rectilinear grid centred on the pole was chosen (Fig. 5). This choice leads to some increased overlap between adjacent tiles at the

^d<https://askap-emuse.streamlit.app/>.

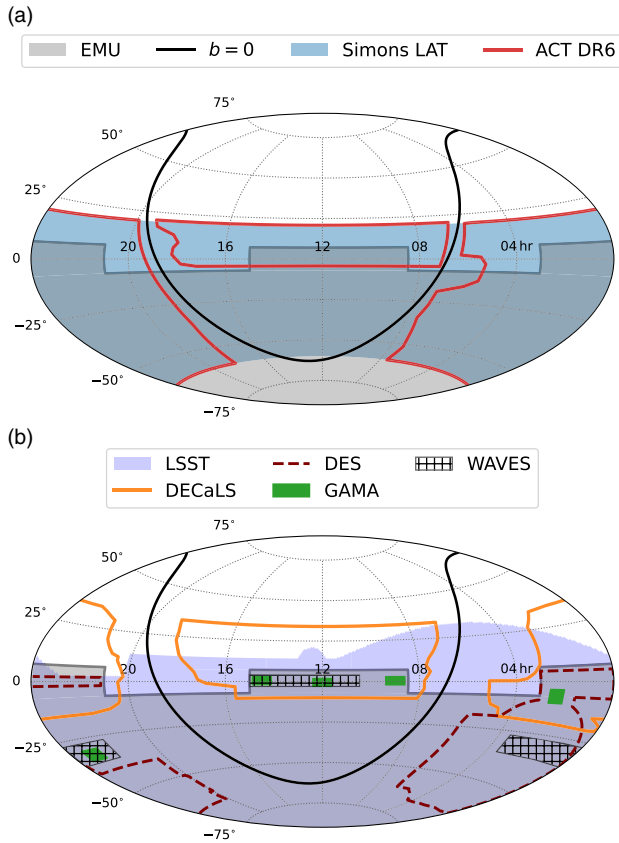


Figure 4. Some examples illustrating the overlap in sky coverage of EMU with (a) mm and (b) optical surveys. Panel (a) shows the footprint of the Atacama Cosmology Telescope 6th Data Release cluster survey (red outline) and the expected coverage of the forthcoming Simons Observatory Large Aperture Telescope survey (blue shading, Ade et al. 2019). Panel (b) shows the survey footprints for DES (dark red dashed outline, Dark Energy Survey Collaboration et al. 2016), DECaLS (orange solid outline, Dey et al. 2019), LSST (blue shaded region, Ivezić et al. 2019), GAMA (green shaded region, Driver et al. 2011; Liske et al. 2015) and WAVES Wide (black hatched region, Driver et al. 2019). In both panels the EMU survey footprint is shown by the grey shaded region and the Galactic equator by the black line.

boundary between the two tiling models, but remains an optimum choice to retain minimal overall duplication, while also facilitating practical telescope operations. The final EMU sky coverage is produced through a tessellation of 853 ASKAP tile footprints. Each requires a 10 h integration to achieve EMU’s sensitivity goals. Of the 853 total regions, 161 lie at declinations far enough north ($\delta > -10^\circ$) that a full 10 h track is not feasible in a single observation. These tiles are split into two separate (5 h) integrations. Each EMU tile is given a name in the format of EMU_hhmm±dd. For those tiles split into two 5 h integrations, the two unique names given are EMU_hhmm±dda and EMU_hhmm±ddb.

In the course of ASKAP Early Science and Pilot science observations, it was established that a central frequency of 943 MHz would be the optimum for EMU. This choice provides the best point source continuum sensitivity while retaining the maximal bandwidth by avoiding the RFI at higher frequencies (Fig. 6).

4. Observations and data processing

EMU observations are optimised with all 36 ASKAP antennas, but can be scheduled with 32/36 (or 34/36 for more northerly tiles), to meet the required uv coverage and sensitivity goals. EMU uses

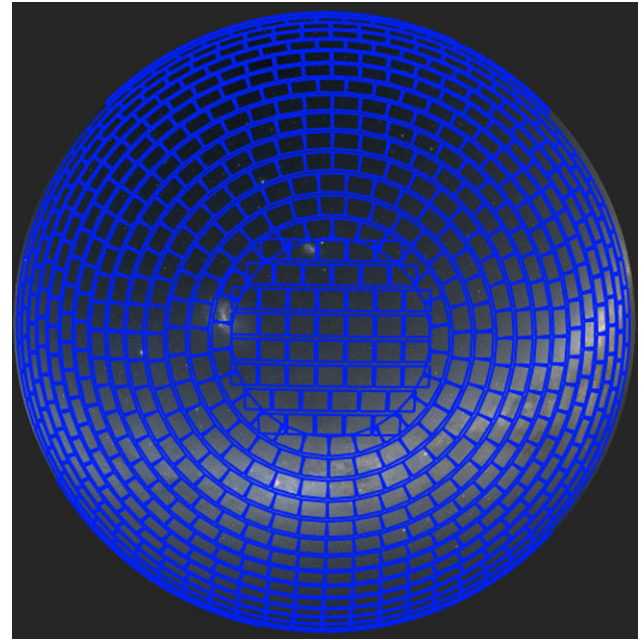


Figure 5. The tessellation selected to cover the celestial pole, showing the change in strategy at $\delta < -70$. The blue outlines correspond to each EMU tile.

a central frequency of 943 MHz with a bandwidth of 288 MHz. Each tile observed by ASKAP is allocated as a ‘scheduling block’ (SB) with an associated unique ID number. Autonomous scheduling of observations is carried out by an observatory-managed tool called SAURON (Moss 2022, Moss et al. in preparation), which takes into account a variety of survey, environmental and system constraints. These include allocation of time to each of the ASKAP surveys in appropriate proportion, consideration of tile availability and observing constraints (such as distance to the Sun), observing program prioritisation for different tiles or tile groups, and more. As a consequence, while still deterministic, the scheduling is dynamic and complex. The motivation for this approach is to maximise the survey efficiency of ASKAP while balancing many competing factors and operating in a remote and extreme environment. In addition to these constraints, EMU’s observations aim to prioritise tiles adjacent to previously completed tiles, in order to build up locally contiguous sky coverage in regions in order to facilitate the triggering of the The EMU Radio and Value-added Catalogue (EMUCAT) pipeline (See Section 6.1 below).

The data are processed using the ASKAPsoft pipeline, including multi-frequency synthesis imaging, multi-scale clean and self-calibration. The process is largely the same as that detailed by Norris et al. (2021b) for the first EMU Pilot survey. In particular, the clean scales and weighting scheme are the same. Some improvements include the fact that the cleaning now goes down to a 5σ level, to take into account varying image quality, rather than using the absolute threshold adopted earlier, as that could potentially result in over- or under-cleaning. In addition, the w -term handling has been improved, to allow almost twice the number of w -planes. Phase self-calibration is now done initially using a sky model derived from RACS (McConnell et al. 2020; Hale et al. 2021) to fix phases over the full integration so that deconvolution has a good starting point. No amplitude self-calibration is performed. The full sequence starts with self-calibration against RACS, an initial deconvolution step, then another self-calibration

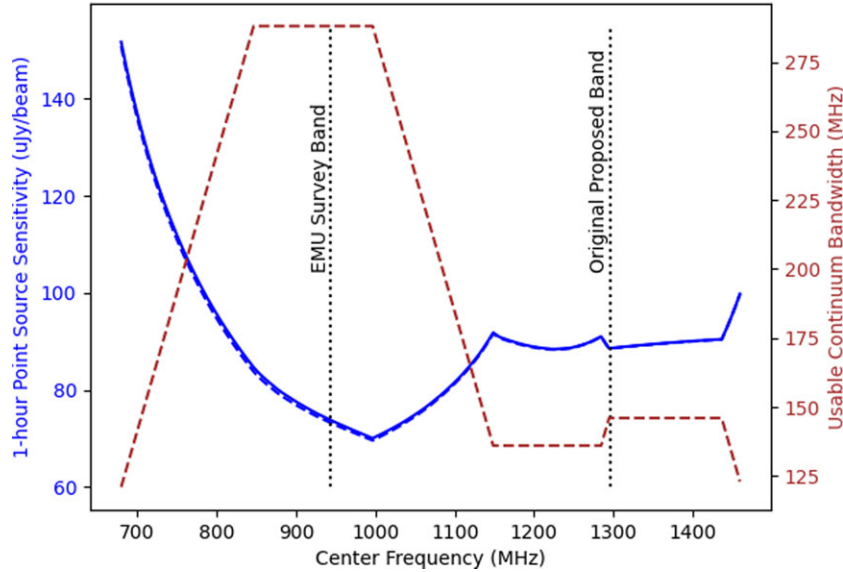


Figure 6. ASKAP sensitivity and usable bandwidth as a function of frequency. The centre frequency refers to the middle of the 288 MHz of correlated bandwidth prior to rejection of unusable channels. The point source sensitivity includes estimates of robust weighting, convolution to a common restoring beam and a mosaicing factor.

against the deconvolution model, followed by a final stage of deconvolution. For details of the array configuration, justification of the choice of baseline weighting, and other technical considerations of ASKAPsoft, see Guzman *et al.* (2019) and Whiting (2020). After validation (see Section 4.1 below) EMU data become available on CASDA. EMU is identified as project AS201, WALLABY as project AS202, and POSSUM as project AS203 on CASDA. The publicly available EMU data include three versions of each validated tile mosaic for each of Stokes I and V. The Stokes Q and U datasets are available as POSSUM data products. The three EMU image versions are produced from the same calibrated dataset, but with different choices of restoring beam for the imaging. The recommended data product includes ‘conv’ as a suffix in the filename, and has been restored with a common 15'' resolution for each beam in the mosaic of a tile. This is consistent for all tiles across the full survey area, and these data are the source for self-consistent flux density measurements across the full survey area. The other two options have the suffixes ‘raw’ and ‘highres’. The ‘raw’ image comes from using the native restoring beam, dependent on non-flagged uv sampling for each beam in the mosaic. This provides a better resolution than the ‘conv’ images, typically around 11 – 13'' but will have varying resolution across the mosaic, and will also differ from tile to tile. Flux densities derived from these images are not necessarily self-consistent between tiles, or even within the same tile mosaic. The ‘highres’ image, made with uniform weighting, is restored with a smaller synthesised beam again, typically 7 – 9''. This is consistent across each beam within a tile mosaic, but will differ from tile to tile, and these images have somewhat poorer sensitivity and imaging fidelity compared to the others. Their value is in highlighting finer-scale structure for (typically brighter) objects of interest, while coming at the expense of losing sensitivity to extended emission. Flux densities in these images, again, will not necessarily be self-consistent.

The reason for choosing the common resolution of 15'' in the ‘conv’ images is a consequence in part of the weighting scheme chosen to maximise sensitivity to faint sources, but is primarily driven by the resolution limitations of tiles close to the equator,

where ASKAP’s uv coverage is somewhat poorer than in tiles further south. This is necessary to ensure a common resolution over the full sky area covered by EMU.

Source catalogues for these images are generated using the *Selavy* source finder (Whiting & Humphreys 2012). *Selavy* estimates local noise statistics within a sliding box to set its detection threshold, accounting for noise variations across an image. It constructs two catalogues for each tile, an ‘island’ and a ‘component’ catalogue. The island catalogue is formed by identifying contiguous regions of emission, or islands, above the local background threshold. The component catalogue is an attempt to resolve blended or complex sources, using a sub-thresholding technique within each island. A series of descending flux thresholds are used to identify distinct peaks, which serve as initial guesses for Gaussian component fitting. The resulting component catalogue consists of the fitted Gaussians within each island, constrained by criteria such as minimum separation, positivity, beam-size limits, and consistency with the island’s total flux. A comparison of the performance of *Selavy* with other point source finders is presented by Boyce *et al.* (2023).

In addition to the EMU images and source catalogues a selection of associated data are available, including the visibilities and a series of ancillary datasets. Ancillary data include weight images used in the mosaicing, clean model images generated during deconvolution, and a series of images generated during source finding, such as background noise images, images containing detected components, and residual images derived from the component image.

4.1 Data validation

The ASKAP operations team delivers fully processed data cubes from the telescope to the science teams, making use of the Pawsey Supercomputing Centre.^e In order to ensure consistent quality control, the science teams are responsible for validating the data.

^e<https://pawsey.org.au>.

Given the scale and logistics involved in operating a major survey facility like ASKAP, and the limited size of the operations team, it is impractical to manually reprocess datasets if they have errors. Instead, it is easier and more streamlined simply to reobserve such a tile. In order to decide whether a tile has or has not met the required quality level, the science team is responsible for a process of data validation, which either accepts the tile as meeting the science requirements, or rejects it. The data from rejected tiles are discarded, and the tiles are placed back in the ASKAP observing queue.

Common problems that may lead to errors in the processed data, and which are typically rectified by this reobservation strategy, include excess solar interference, system issues, or erroneous calibrations (bandpass, holography, or, for some observations early in the process, other key datasets that either had errors or that were not observed close enough in time to the dataset in question). These issues are typically flagged in a processed EMU mosaic through excessively large positional or flux density offsets between measured sources and a reference dataset (such as RACS), unusual source count properties, or excessive imaging artifacts (radial spikes, especially around relatively faint sources, or large scale ripples). To facilitate this validation process, each ASKAP science team has established its own validation workflow, which produces a series of diagnostic figures, statistics, and traffic-light metrics to aid the team in rapidly identifying problematic data cubes; EMU makes use of the ASKAP continuum validation workflow.^f

ASKAP delivers three to four newly observed and processed tiles for EMU per week, on average, which each require validation. The validation team includes more than 30 volunteers from among the EMU collaboration, and uses a roster system to share the load of conducting the manual validation and inspection process. In order that the process is as streamlined as possible, a validation workflow was established, with examples and clear step-by-step instructions. This workflow was explained and demonstrated through several training sessions. The validation workflow includes independent recording of key validation metrics, to facilitate subsequent analysis of the EMU tile properties. To ensure a level of consistency in this quality control, the validation team are instructed to escalate potentially problematic tiles to the validation lead and management team for final review before any tile is rejected in the validation process.

The number of rejected tiles has decreased with time as ASKAP operations have become more streamlined, although a significant spike in these numbers occurred during a recent peak in solar activity and associated solar storm events in mid-2024. Overall, about 20% of EMU observations to date have resulted in rejected tiles through this validation process. Overheads associated with operational improvement over time have been factored into the timing of ASKAP's 5-yr observing plan, which assumes an observational efficiency of 70%.

4.2 Known image processing limitations

The ASKAPsoft pipeline needs to accommodate requirements for each of the ASKAP surveys, and in consequence there are inevitably some limitations in the processing that affect the EMU imaging. We outline the main issues and their impact here.

One key issue is a limited implementation in the handling of the w -term^g. As a result, sources near to or outside the edge of the primary beam's FWHM frequently show radial spike or grating-ring type artifacts. Bright sources outside of the primary beam image result in residual grating-ring artifacts and calibration errors (since such sources are not included in the deconvolution during phase self-calibration). There are a number of approaches under development to improve the way this is handled within ASKAPsoft, although these are not yet implemented for existing EMU data products at the time of writing. This includes the introduction of source peeling, and improvements to w -term handling, supported by better memory utilisation. One recent straightforward approach, shortly to be added to the official ASKAPsoft pipeline, draws on a sky model for the location of known bright sources. Such bright out-of-field sources are imaged, modelled, and subtracted from the uv data before the main imaging steps to reduce the impact of their associated artifacts. This will serve to provide some improvement in image quality for newly observed EMU tiles.

Apart from initial phase calibration as a result of the bandpass observation, no subsequent phase referencing is performed. This is partially compensated for with phase self-calibration, but relies on a good deconvolution model. Any errors in the deconvolution model will result in calibration errors and subsequent imaging artifacts. Significant astrometric errors can also result when the phases are greatly disturbed over the course of the observation (these are not necessarily consistent from beam to beam and so can result in a slight smearing of sources when overlapping beams are mosaicked). This astrometric error is independent of the signal-to-noise (S/N) ratio of the source (further details are presented in Section 5.2 below). Investigations of RACS observations have shown that the bulk of the astrometric error is set during the bandpass observation, as this procedure currently takes over two hours to complete (with an approximately 3 min scan performed on the calibrator source PKS1934–638 in each beam). One option to improve the initial bandpass is to either fix the phases against a reference field (and so align the beams) or to perform a shorter field-based bandpass observation and so determine phase solutions for all beams and antennas simultaneously. A longer-term solution being considered by the observatory is to phase-reference the target field against the RACS sky model. This allows time-dependent phases to be corrected as well.

Solar interference has become problematic with a more active Sun. Solar flares are particularly difficult to deal with as they are exceptionally bright, often highly variable and, owing to their compact nature, affect all baselines. A secondary effect is that Solar activity can shift the true beam positions during the beam-forming process since the Sun is used to determine beam weights. In general though, these are expected to be corrected by the holography which is performed for each beam-forming session.

Telescope pointing errors can result in direction-dependent errors, and increased artifacts at the edge of the beam. These are not corrected for by the third telescope axis (which maintains the orientation of the beam with respect to the sky). They not only affect the target source but also the effectiveness of the holography. Further investigation by the observatory team is underway to determine the magnitude and nature of these pointing errors.

^f<https://github.com/Jordacious/ASKAP-continuum-validation>.

^gThis refers to techniques used to account for the effects of the third spatial coordinate, w , in the uvw -coordinate system, which represents the baseline geometry in Fourier space.

Changes in antenna bandpass can detrimentally affect imaging quality. This can occur where an antenna exhibits a partial failure that is not immediately detected by the system and causes amplitude or phase changes in the bandpass. Such an occurrence is rare and is more easily dealt with during the validation process. A lower level issue that affects amplitude gain has been observed to vary diurnally, possibly being temperature related. This is also currently uncorrected for (in simple arrays this would typically be corrected via phase referencing) and could cause low-level artifacts in imaging.

Overall, while there is clearly still room for improvement in the telescope performance and the delivered image quality, these issues do not significantly compromise the integrity of the EMU data provided. Ongoing developments are expected to provide incremental improvements to the EMU data quality of the remainder of the survey.

The EMU team has a goal by the end of the survey to reprocess early tiles to bring them into line with the final image quality to homogenise the sensitivity and image fidelity, in support of an eventual final data release.

4.3 Confusion limits

A fundamental limit to the sensitivity of radio (and other) images is the ‘confusion’ level, the rms fluctuations in an image due to the contribution of the entire source population that is below the level where individual sources can be detected (Condon 1974). Using the formalism detailed in Condon *et al.* (2012) and further developed by Vernstrom *et al.* (2014) and (2015) we expect the confusion level with EMU’s default 15’’ synthesised beam and 943 MHz frequency to be $\approx 15 \mu\text{Jy beam}^{-1}$. This is comparable to the thermal noise, as measured in the Stokes V images. Further reductions in noise would therefore result in only marginal improvements in sensitivity, so EMU is very close to being as sensitive as possible at its resolution and frequency.

The EMU total intensity images have a typical rms of $\approx 30 \mu\text{Jy beam}^{-1}$, a factor of ≈ 2 brighter than either the thermal rms or confusion level separately, or ≈ 1.5 more than the $\approx 20 \mu\text{Jy beam}^{-1}$ expected from their combination (a similar excess is found for the 151 MHz, 20’’ data products of LoTSS). To assess the contributions to the EMU image rms from various factors, we made difference images between repeated observations of the same fields. The rms fluctuations in the difference images did not increase above that of the input images, as would be expected if it arose purely from thermal noise, which is independent for each observation. Thus, the measured rms fluctuations contain contributions both from confusion and instrumental effects such as the cumulative sidelobes from very faint sources, as well as the thermal noise. Further processing developments may allow us to push closer to the pure confusion limit.

To provide a more concrete illustration of EMU’s sensitivity to low surface brightnesses, we convert it to a brightness temperature detection limit, which is given by

$$T = \frac{\lambda^2 S_{\text{lim}}}{2k\Omega_b}, \quad (1)$$

where $\lambda = 0.318 \text{ m}$ at EMU’s 943 MHz, S_{lim} is the limiting flux density of detected sources, $k = 1.38 \times 10^{-23} \text{ J K}^{-1}$ is the Boltzmann constant, and Ω_b is the restoring beam solid angle. Adopting a flux limit of $S_{\text{lim}} = 150 \mu\text{Jy}$ and EMU’s 15’’ resolution, giving $\Omega_b = 5.99 \times 10^{-9} \text{ sr}$, we have EMU’s limiting brightness

temperature $T = 0.92 \text{ K}$. The brightness temperature of nearby face on spiral galaxies is about $T \approx 1 \text{ K}$ at 1.4 GHz (Condon *et al.* 1998), and so about $T \approx 3 \text{ K}$ at EMU’s 943 MHz observing frequency. The brightness temperature falls with redshift (for synchrotron sources with $\alpha = -0.7$) as $T \propto (1+z)^{-3+\alpha}$, implying that EMU will be able to detect such typical face on spiral or star forming galaxies to $z \approx 0.3 - 0.4$.

Given EMU’s excellent uv coverage, it is also sensitive over a wide range of angular scales. This enables the survey to probe much lower than the nominal confusion level at scales larger than 15’’, by first subtracting sources at the 15’’ resolution, and then convolving to larger scales. An automated process for doing this is described in Section 6.2, and results in rms fluctuations reaching $< 9 \mu\text{Jy (15'' beam)}^{-1}$, equivalent to $54 \mu\text{Jy (45'' beam)}^{-1}$ for structures on that scale and above. Thus, EMU will provide extremely sensitive detections of faint extended sources, such as cluster halos and dying radio lobes and supernova remnants, that would not be otherwise possible.

5. EMU source statistics

Providing a complete and homogeneous source catalogue for EMU will require combination of independently observed tiles to exploit the full depth of the survey. This will account for the overlap regions between tiles as well as the northern tiles that are only observed in 5 h scheduling blocks, as well as de-duplication of source detections where fields overlap (see Section 6.1). For a wide-area multi-year survey such as EMU, this process is most efficiently achieved once large contiguous regions of the survey footprint have been observed, and are implemented as the core stage of EMUCAT (Section 6.1). More immediately, when an EMU tile is observed the source finder *Selavy* (Whiting & Humphreys 2012) is automatically run on the image as part of the ASKAP data pipeline. The resultant source catalogues for individual EMU tiles provide (a) broad accessibility to EMU data for a range of scientific objectives prior to the availability of EMUCAT data products, and (b) readily interpretable measurements with which to characterise the performance of the survey.

Using the second of these benefits, we can analyse the source statistics to quantify the performance of EMU to date. In this analysis, we draw on the data available as of July 2024. At that time 220 EMU fields ($\sim 20\%$ of the full EMU survey) had been observed, passed validation, and were released on CASDA (Fig. 7). Of these, 143 are full-depth 10 h observations and 77 are 5 h observations in the northern part of the EMU footprint, 36 of which are re-observations of the same field. We note here that the 5 h observing blocks (each of which are observed twice) are combined in the image plane as part of the EMUCAT pipeline (see Section 6.1), providing equivalent depth to the 10 h observations, in order to deliver a uniform sensitivity sky catalogue. Reimaging the fields by combining the uv data for large areas of a hemispheric survey is a substantial undertaking, and is currently deferred to a later phase of the project. Throughout the remainder of this section we provide statistics based on the currently available automatically generated catalogues, including separate catalogues for each of the repeated 5 h tiles.

The 10 h observations have a median rms noise of $30 \mu\text{Jy beam}^{-1}$, typically detecting ≈ 750 sources deg^{-2} . A few of the 10 h observations are subject to increased noise, usually as a result of a bright radio source in the field, e.g. the image from SB 46986 has an rms of $258 \mu\text{Jy beam}^{-1}$ (the highest noise

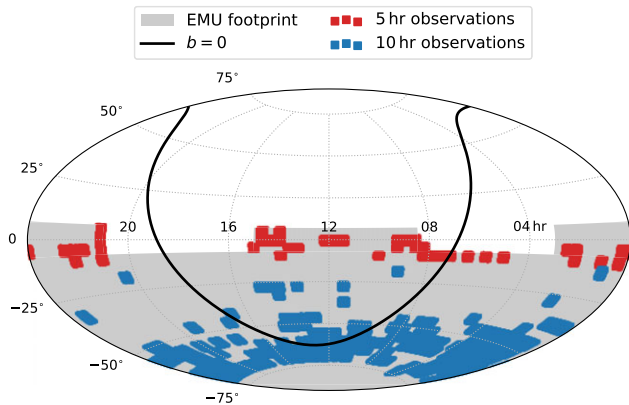


Figure 7. Sky map (Aitoff projection) showing the EMU tiles released as of July 2024. The grey shaded region shows the full EMU footprint, the blue regions show released 10 h observations, and the red regions show released 5 h observations.

in any EMU tile observed prior to July 2024) a consequence of containing PKS 0131-36 (a 10 Jy source at 843 MHz, Jones & McAdam 1992), but more than 95 % of these observing blocks have an rms noise $< 45 \mu\text{Jy beam}^{-1}$. The 5 h observations have a median rms of $46 \mu\text{Jy beam}^{-1}$ (ranging between $40 \mu\text{Jy beam}^{-1}$ and $60 \mu\text{Jy beam}^{-1}$), detecting ≈ 460 sources deg^{-2} on average. Given the overlap between adjacent EMU tiles that have been observed, simply concatenating the catalogues for individual tiles will result in some duplicate sources. For the statistics presented here, a first-order removal of duplicates is achieved by finding sources where the distance to its nearest neighbour is less than $10''$ and where that nearest neighbour is not in the same EMU field. For the 5 h observations, the catalogue from the second observation of the tile will be dominated by the same sources that were detected in the first observation. To avoid duplicates from these re-observations, we only include sources detected in the first of the two scheduling blocks dedicated to these tiles.

In Fig. 8 we show the peak flux distribution of all sources detected in EMU so far. Requiring a minimum S/N of 5, the 10 h observations (blue line) detect sources down to $\approx 0.17 \text{ mJy beam}^{-1}$ and the 5 h observations (red line) detect sources down to $\approx 0.31 \text{ mJy beam}^{-1}$. For reference, we also show the peak flux distribution of sources identified by RACS in their low-frequency observations (RACS-Low, McConnell et al. 2020). RACS-Low covers $\approx 30\,000 \text{ deg}^2$ at $\nu \sim 888 \text{ MHz}$ down to an rms noise level of $\sim 300 \mu\text{Jy beam}^{-1}$, and is in effect a shallower analogue of EMU. The black dotted line in Fig. 8 is obtained by randomly sampling the RACS-Low catalogue (Hale et al. 2021) to normalise for the smaller observed footprint of EMU, and scaled assuming a typical spectral index of $\alpha = -0.7$ to account for the small difference in the observing frequencies of the two surveys. Qualitatively, the bright end of the EMU flux distribution behaves as expected for a deeper version of RACS-Low with the two surveys having near identical distributions above $S_{943 \text{ MHz}} \approx 2 \text{ mJy beam}^{-1}$. The notable increase in radio source number counts at $S_{943} \lesssim 1 \text{ mJy}$ can be attributed to the increasing proportion of star-forming galaxies at these faint levels (e.g. Windhorst et al. 1985; Hopkins et al. 1998; Seymour et al. 2008), highlighting the ability of EMU to probe radio source populations that are missed by shallower surveys such as RACS.

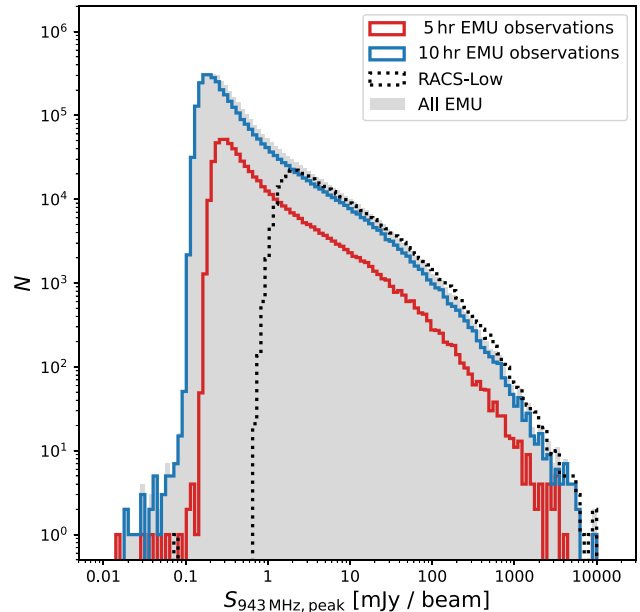


Figure 8. Distribution of the peak flux density of EMU sources. The solid grey histogram shows the EMU sources analysed here. The blue line shows sources from 10 h EMU observations and the red line shows those sources from 5 h EMU observations. For reference, the black dotted line shows the RACS low peak flux distribution normalised to account for differences in survey footprint.

5.1 Fraction of resolved sources

The majority of radio sources detected by EMU are expected to be unresolved by the EMU beam (Norris et al. 2006, 2021b). In such cases, if the EMU point spread function is well modelled by a Gaussian, then the total flux of a point source in EMU should be equal to the amplitude of the Gaussian fitted to that source by *Selavy*, at least within measurement uncertainties. For resolved sources however, one would expect the total flux density, S_{total} , to be greater than the peak flux density, S_{peak} . In order to identify unresolved sources in EMU, we adopt a commonly-used approach of defining an envelope in the parameter space of $S_{\text{total}}/S_{\text{peak}}$ versus S/N (Fig. 9), where the S/N is given by $S_{\text{peak}}/\text{rms}$ (e.g. Smolčić et al. 2017; Shimwell et al. 2019; Hale et al. 2021). Such an envelope is expected to be of the form:

$$\frac{S_{\text{total}}}{S_{\text{peak}}} = \text{med} \left(\frac{S_{\text{total}}}{S_{\text{peak}}} \right) \pm A (S/N)^{-B}. \quad (2)$$

To define our envelope, we first select a sample of likely compact and isolated sources. In practice, we obtain these by requiring that the source is the only Gaussian fitted to that source island, and that the estimated major axis of the source is less than $20''$. For the most robustly measured of such ‘quasi compact’ sources ($S/N > 100$) we determine the median $S_{\text{total}}/S_{\text{peak}}$ to be 1.05. To quantify the scatter in $S_{\text{total}}/S_{\text{peak}}$ as a function of S/N, we measure the 2.5th percentile of $S_{\text{total}}/S_{\text{peak}}$ in 10 bins of S/N with equal logarithmic width. Our lower envelope is then determined using a non-linear least-squares fit assuming the form of Equation (2), finding $A = 1.09$ and $B = 0.54$ (dashed red line in Fig. 9). Taking the mirror of this line around the median of $S_{\text{total}}/S_{\text{peak}}$ as the upper envelope (solid red line in Fig. 9), we expect 95% of the compact sources to

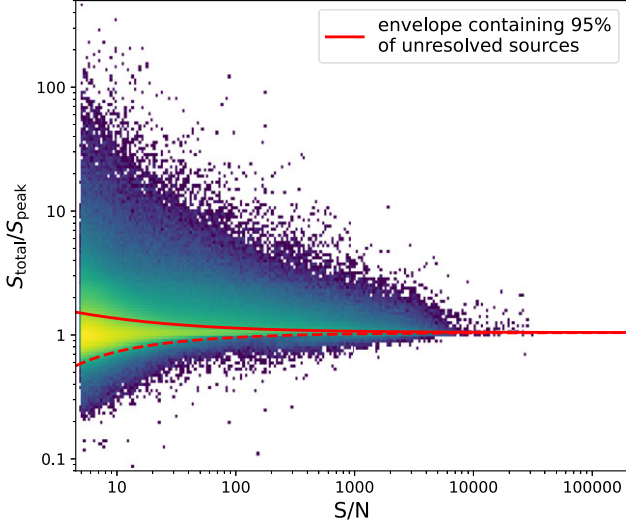


Figure 9. Distribution of total to peak flux density of EMU sources as a function of S/N. The red solid line denotes the cut off for identifying resolved sources and is a mirror of the red dashed line based on the 95% scatter in compact sources.

lie in the region bounded by:

$$\frac{S_{\text{total}}}{S_{\text{peak}}} < 1.05 + 1.09 (S/N)^{-0.54}. \quad (3)$$

We then find that 17% of all the EMU sources (not just those defined as compact) have $S_{\text{total}}/S_{\text{peak}} > 1.05 \pm 1.09 (S/N)^{-0.54}$ and can be considered resolved, with the remaining 83% likely being point sources at EMU’s resolution.

5.2 Astrometry

In order to quantify the typical astrometric precision of EMU, we cross-match the point sources, i.e. those lying below the solid red line in Fig. 9, to the closest source in two other surveys, using a $10''$ search radius. First we match EMU sources with the *Gaia* third data release, a catalogue with sub-milliarcsecond astrometric precision. *Gaia* is an optical telescope, and so the vast majority of sources observed by *Gaia* will not have radio counterparts. To minimise the contamination of our cross-match by such objects, we only cross-match with *Gaia* sources considered likely to be quasars ($p_{\text{QSO}} > 0.8$), finding $\approx 35\,000$ matches. In Fig. 10a we show the distribution of offsets in right ascension, α , and declination, δ , for these matches where $\Delta\alpha = (\alpha_{\text{EMU}} - \alpha_{\text{Gaia}}) \cos(\delta_{\text{EMU}})$ and $\Delta\delta = \delta_{\text{EMU}} - \delta_{\text{Gaia}}$. We find the means of the positional offsets to be $\Delta\alpha = -0.319 \pm 0.011''$ and $\Delta\delta = -0.109 \pm 0.011''$ (uncertainties are standard errors of the means), with a standard deviation in both $\Delta\alpha$ and $\Delta\delta$ of $2.14''$.

Second, we match EMU sources to the first epoch ‘Quick Look’ catalogue from VLASS (Lacy et al. 2020; Gordon et al. 2021). With a typical angular resolution of $\approx 3''$ and sub-arcsecond astrometric precision, VLASS can provide some of the most robust radio positions of any wide area sky survey. VLASS is substantially less deep than EMU, being sensitive to point sources brighter than ~ 1 mJy, and consequently one would expect the majority of VLASS sources in the survey overlap regions to be detected by EMU, significantly reducing the contamination by spurious matches that may impact our cross-match with *Gaia*. To account for the differences in resolution between the two surveys we only cross-match with likely

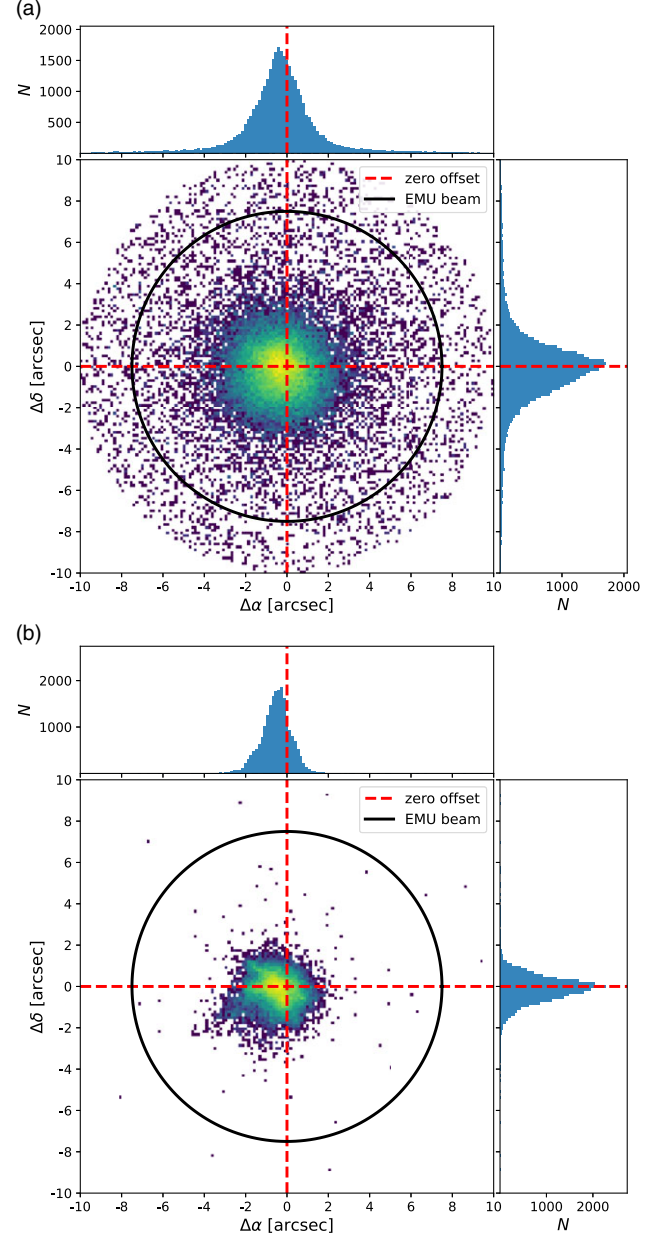


Figure 10. Positional offsets between EMU sources and *Gaia* QSOs (top) and VLASS point sources (bottom). The red dashed lines show zero offset, and the black circle shows the FWHM of the typical EMU beam ($15''$).

compact sources in VLASS (deconvolved major axis of less than $1''$), finding $\approx 21\,000$ matches and showing the distribution of their positional offsets in Fig. 10b. The typical positional offsets between EMU and VLASS sources are $\Delta\alpha = -0.538 \pm 0.005''$ and $\Delta\delta = -0.137 \pm 0.005''$, with standard deviations in $\Delta\alpha$ and $\Delta\delta$ of $0.760''$ and $0.678''$, respectively.

Both our cross-match with *Gaia* and our cross-match with VLASS are consistent with EMU sources being offset from the reference frame by $\sim 0.5''$ to the west and $\sim 0.1''$ to the north. For the first EMU Pilot Survey Norris et al. (2021b) found a similar astrometric offset in declination, but report $\Delta\alpha = +0.11''$ when comparing to CatWISE. Notably, while our early EMU main survey data cover the entire southern sky, the EMU Pilot Survey spanned

the more limited declination range of $-63^\circ < \delta < -48^\circ$. When analysing the RA bias in our cross-match with *Gaia* split by the median declination of our sample ($\delta = -54^\circ$), those EMU/*Gaia* matches at $\delta < -54^\circ$ have a smaller RA bias ($\Delta\alpha = -0.18 \pm 0.02''$) than those at $\delta > -54^\circ$ ($\Delta\alpha = -0.44 \pm 0.02''$), suggesting a potential declination dependence in the EMU astrometric bias. The origin of this astrometric bias is currently unknown and being investigated. Furthermore, we note that the scatter in positions we find is greater than that found in the EMU Pilot Survey (see Figure 14 of Norris et al. 2021b) although it is unclear whether this is due to increasing solar activity, the greater use of daytime observations, or some other cause. The sub-arcsecond scale of the effect is substantially smaller than the typical resolution of the survey ($15''$) and the pixel size of the EMU images ($2''$), and is unlikely to be problematic for most science. However, until the cause is well understood, care should be taken in accounting for this known astrometric bias when matching EMU data to very high resolution multiwavelength observations (e.g. those using adaptive optics or space-based telescopes).

5.3 Accuracy of flux density measurements

To quantify the accuracy of the flux density measurements made by EMU, we compare these measurements for point sources in EMU that are also detected in RACS-Low. We pick RACS-Low as a comparison survey for a number of reasons. First, like EMU, RACS-Low was observed using ASKAP and consequently any systematic biases resulting from the telescope design are expected to be similar for both surveys. Second, the entire EMU footprint is covered by RACS-Low. Third, RACS-Low is observed at median frequency $\nu \sim 888$ MHz, using a 288 MHz bandwidth, resulting in an overlap in the frequency coverage of the two surveys. Consequently, spectral curvature should have a limited impact on the flux density measurements of non-variable sources observed by both EMU and RACS-Low. Fourth, the accuracy of measurements from RACS-Low has already been characterised by comparison with other radio surveys, allowing the possibility of this single direct comparison with EMU. Indeed, Hale et al. (2021) found a high level of accuracy in the RACS-Low flux density measurements, showing for sources observed in both RACS-Low and the Sydney University Molonglo Sky Survey (SUMSS, Bock, Large, & Sadler 1999) the ratio of flux measurements from these two surveys to be $S_{\text{RACS-Low}}/S_{\text{SUMSS}} = 1.00 \pm 0.16$, when accounting for spectral index.

While the native resolution of RACS-Low is similar to that of EMU ($\sim 15''$), the RACS-Low catalogue was produced from images convolved to a common resolution of $25''$ (Hale et al. 2021). So as to not encounter issues arising from comparing flux densities of sources observed at different resolutions, we require a sample of sources that are unresolved in both EMU and RACS-Low, and isolated in EMU (the better resolution data). For EMU we select sources satisfying Equation (3) and being the only source fit to their flux island by *Selavy*. Unresolved sources in RACS-Low are identified by $S_{\text{total}}/S_{\text{peak}} < 1.025 + 0.69 (S/N)^{-0.62}$ (Hale et al. 2021). We match our isolated and unresolved EMU sources to unresolved RACS-Low sources within $3''$, finding ≈ 22 000 sources. This search radius is large enough that it accounts for the scatter in the EMU astrometric precision (see Section 5.2). A larger search radius however would not find many more matches, with $> 90\%$ of those RACS-Low sources that are within $10''$ of an EMU source being separated by $< 3''$.

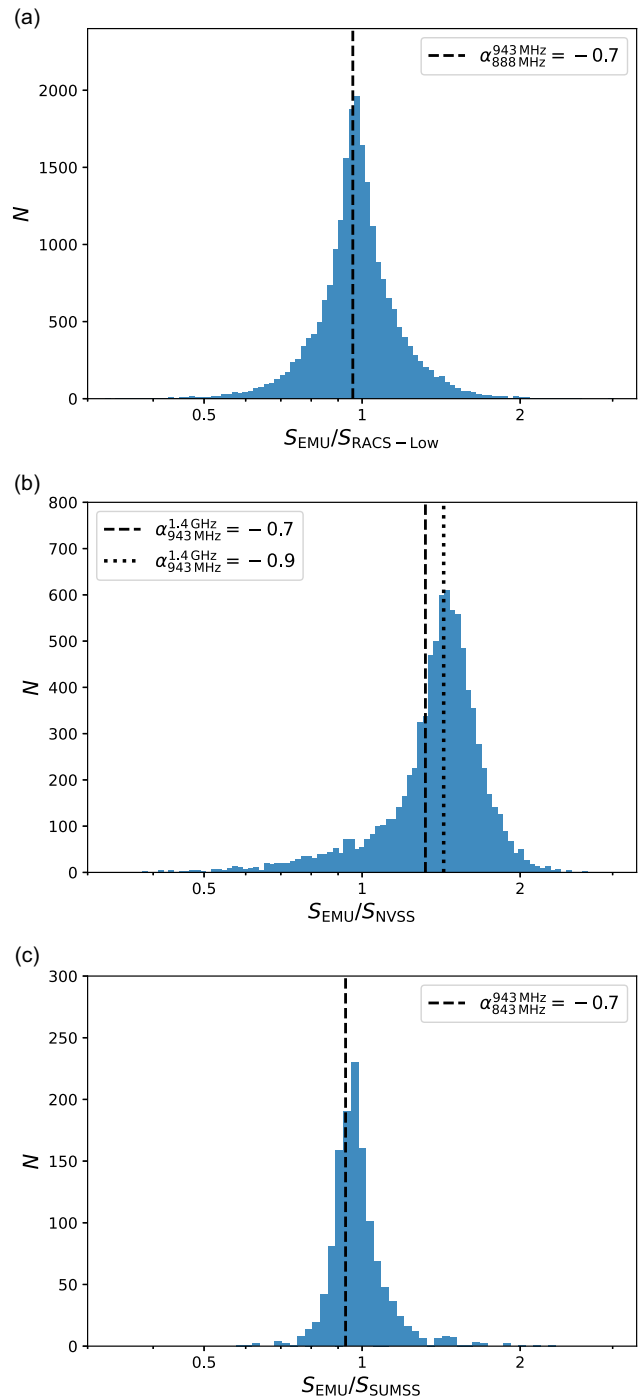


Figure 11. Distributions of the ratio of flux density in EMU to the flux density measurement in (a) RACS-Low, (b) NVSS, and (c) SUMSS, for compact and isolated sources. The black dashed vertical lines in all panels show the expected ratio for a source with $\alpha = -0.7$, and in panel (b) the black dotted vertical line shows the expected ratio assuming $\alpha = -0.9$.

Figure 11 presents comparisons of the EMU flux densities against those from other surveys. In Fig. 11a, we show the distribution of EMU to RACS-Low flux density ratio, $S_{\text{EMU}}/S_{\text{RACS-Low}}$. Assuming a typical two-point spectral index between the RACS-Low and EMU central observing frequencies of $\alpha_{888 \text{ MHz}}^{943 \text{ MHz}} = -0.7$ and neglecting any

potential variability (the majority of EMU and RACS sources are not expected to show measurable variability), the flux density in EMU should be 96% of the flux density in RACS-Low. We find the median value of $S_{\text{EMU}}/S_{\text{RACS-Low}}$ to be $0.98^{+0.16}_{-0.13}$, where the uncertainties are defined by 16th and 84th percentiles.

In addition to comparing to the RACS-Low measurements, we also compare the EMU flux density measurements to sources detected by NVSS and SUMSS, two radio continuum surveys not conducted using ASKAP. Both surveys have a poorer resolution than EMU, $45''$ for NVSS and $45'' \times 45'' \text{cosec}|\delta|$ for SUMSS. In order to ensure we are not comparing the measurements from multiple EMU sources blended into a single source in the poorer resolution surveys, we only consider EMU sources where the angular separation to their nearest neighbour in EMU is larger than the beam size of the comparison survey. Point sources in EMU that are sufficiently isolated, by this criterion, are then matched to the catalogues of NVSS and SUMSS using a $3''$ matching radius, and the flux measurements for sources that are considered unresolved in those surveys are compared to the EMU flux density. The NVSS catalogue does not provide both a peak and integrated flux density measurement, so we adopt those sources where the major axis of the fitted source size is smaller than or equal to or the NVSS beam size of $45''$ as being unresolved. Similarly, SUMSS provide source sizes for resolved sources in their catalogue and we only use sources in the SUMSS catalogue they identify as unresolved (see Section 4.2 of Mauch et al. 2003).

In Figs. 11b and 11c we show the distributions of the flux measurement ratios $S_{\text{EMU}}/S_{\text{NVSS}}$ and $S_{\text{EMU}}/S_{\text{SUMSS}}$, respectively. In both these panels, the black dashed vertical line shows the expected ratio for a source having a typical power-law spectrum with a spectral index of $\alpha = -0.7$. We would expect a such a source to be 32% brighter in EMU than in NVSS, which at 1.4 GHz is observed at a substantially higher frequency than EMU. Instead, we find the median value of $S_{\text{EMU}}/S_{\text{NVSS}}$ to be $1.43^{+0.21}_{-0.28}$, approximately 8% larger than expected. There are two likely explanations for this result. First, there could be a systematic bias in the ASKAP data that causes the flux density measurements to be too high. Second, this result could be explained by a steeper spectral index of $\alpha = -0.9$ between 943 MHz and 1.4 GHz being more typical of the population sampled in our matches, as shown by the black dotted vertical line in Fig. 11b, and previously suggested by Hale et al. (2021).

The distribution of $S_{\text{EMU}}/S_{\text{SUMSS}}$ shown in Fig. 11c can help distinguish between these two scenarios. Like NVSS, SUMSS also used a different telescope to EMU, in this case the Molonglo Observatory Synthesis Telescope (Mills 1981), and consequently if there were a systematic bias in the ASKAP flux density measurements then we would expect the median value of $S_{\text{EMU}}/S_{\text{SUMSS}}$ to be overestimated, similarly to what we see in the $S_{\text{EMU}}/S_{\text{NVSS}}$ distribution. The 843 MHz observations of SUMSS are much closer to the EMU observing frequency than the 1.4 GHz observations of NVSS, and consequently $S_{\text{EMU}}/S_{\text{SUMSS}}$ is less sensitive to spectral gradient (and curvature) than $S_{\text{EMU}}/S_{\text{NVSS}}$. If the spectral index is driving the apparent overestimation of $S_{\text{EMU}}/S_{\text{NVSS}}$, then we would expect to see a smaller effect in the $S_{\text{EMU}}/S_{\text{SUMSS}}$ distribution. Given the observing frequencies, we would expect a source with $\alpha = -0.7$ to be 93% as bright in EMU as in SUMSS. We find the median value of $S_{\text{EMU}}/S_{\text{SUMSS}}$ to be $0.96^{+0.11}_{-0.06}$, approximately 3% larger than might be expected. While this doesn't entirely eliminate the possibility of measurement biases in the EMU data, our results here are consistent with being driven by the spectral shape

of the radio sources used, and suggest that any systematic biases in the actual ASKAP flux density measurements are typically only on the order of a few percent.

The repeated observation strategy used in the northern fields provides an opportunity to quantify the internal scatter in the EMU flux density measurements. For any given source, the ratio of the flux density measurements taken in the second (B) observation to the first (A) observation, $S_{\text{EMU B}}/S_{\text{EMU A}}$, is a consequence of both the repeatability of the measurement, and any intrinsic source variability. The majority of radio sources are not expected to be variable. Moreover, the mean time interval between the two observations of the 5 h fields observed up until July 2024 is low at 18 days, and Hajela et al. (2019) found that $< 0.06\%$ of radio sources vary significantly on similar timescales. The spread in the distribution of $S_{\text{EMU B}}/S_{\text{EMU A}}$ is therefore dominated by the intrinsic measurement scatter rather than variability. We match unresolved sources from the *Selavy* catalogues for the A and B images of the 5 h fields using a $3''$ radius, finding $\approx 320\,000$ matches. In comparing the flux measurements for these sources, we find the mean and standard deviation in $\log_{10}(S_{\text{EMU B}}/S_{\text{EMU A}})$ to be 0 and 0.06, respectively, consistent with a $\sim 15\%$ scatter in the EMU flux density measurements.

6. EMU value-added data

6.1 EMUCAT

EMUCAT (Marvil et al., in preparation) is a sophisticated resource being developed by EMU team members in coordination with the Australian SKA Regional Centre (AusSRC). The primary aims of EMUCAT are to combine all individual ASKAP EMU observations into a single cohesive all-sky data product and then add value through the inclusion of other multi-wavelength survey data and associated derived data products. EMUCAT will substantially improve the usability and scientific value of the ASKAP data and is expected to be the primary resource used by EMU science programs requiring large statistical samples.

EMUCAT improves upon the sensitivity and uniformity of the individual EMU tile observations and provides a solution to their consolidation and de-duplication. An individual EMU observation, which consists of a 36-beam mosaic, has a direction-dependent sensitivity variation due to antenna primary beam and receiver sensitivity effects. EMU's survey strategy (Section 3) includes enough overlap of adjacent tiles so as to allow for the recovery of a more uniform sensitivity when these tiles are combined. The EMUCAT workflow obtains validated tiles from CASDA and creates a larger linear mosaic of these mosaics, i.e. a 'super-mosaic.' The workflow then runs the ASKAP source finder *Selavy* on the super-mosaic and ingests the results into a database.

EMUCAT is constructed as a relational database and assembled via automatically triggered workflows. EMUCAT uses a set of non-overlapping region definitions to divide the full EMU survey area into 281 patches of roughly 70 deg^2 each, and the workflow is triggered once per EMUCAT region. After ingesting the *Selavy* results, the workflow performs a series of post-processing and cross-identification tasks and computes various derived astrophysical properties, updating the database at each stage. The value-added data produced include identification of single and multi-component radio sources through cross-matching to the *WISE* catalogues (AllWISE, CatWISE, and unWISE). About 70% of *Selavy* components have counterparts within $5''$ in unWISE

and CatWISE (dropping to about 50% in AllWISE). This is supplemented through delivery of multiwavelength photometry and redshifts from a subsequent series of cross-identification stages, and deriving or measuring a range of properties (colours, sizes, morphologies, AGN/SF classification, etc.). The catalogues for cross-matching currently include VHS (DR 5), Legacy Survey (DR 10), 2MASS, 6dFGS, 2dFGRS, WiggleZ, GAMA, SDSS (DR 16), DES (DR 2), and this list is being actively expanded. Full details of the catalogue matching process, associated data structures and derived products are continuing to evolve. These will be frozen in advance of EMU public data releases, and full details will be presented in the EMUCAT description paper (Marvil et al., in preparation).

EMUCAT products are currently served through a VO-compliant TAP server hosted at the AusSRC. They are accessible through TAP-aware tools such as TOPCAT, through python notebooks at the AusSRC co-located with the data, and via a user portal for basic database queries. The super-mosaics produced by the EMUCAT workflow are made available on CASDA as EMU ‘derived’ data products. Periodic EMUCAT data releases will be issued over the course of the EMU survey, for which we anticipate the first public data release around mid-2025. We expect a total of three public EMU data releases over the life of the survey, with the second release consisting of between one half and two-thirds of the final dataset, and the final release to be the full dataset.

Additionally, we have developed a comprehensive detection pipeline, (RG-CAT, Gupta et al. 2024b), aimed at constructing radio galaxy catalogues through the use of advanced computer vision algorithms. This pipeline follows a two-step approach. First it detects compact and extended radio galaxies along with their potential host galaxies jointly in radio (EMU) and infrared (AllWISE, Cutri et al. 2021) images. This is achieved using computer vision networks built on transformers (based on convolutional neural networks), supplemented by a novel keypoint detection approach to identify host galaxies. Then it uses these predictions to build a consolidated catalogue.

The computer vision model identifies sources within an image and places bounding boxes around multiple components of extended radio sources, grouping them as individual radio galaxies. The catalogue construction pipeline (see Figure 5 of Gupta et al. 2024b) relies on predictions from the Gal-DINO model, first introduced in RadioGalaxyNET (Gupta et al. 2024a), using cutouts generated from the *Selavy*-based components catalogue. For each cutout, bounding boxes, categories, and infrared host predictions for the central source are used to assemble the final catalogue.

These automated approaches require a training/validation set of galaxies. For this we have used the catalogue of about 3 600 double-lobed radio sources obtained from the first EMU Pilot Survey by Norris et al. (PASA, submitted). Sources were manually identified, thus avoiding the biases introduced by automated source-finders. Each source is also manually given a morphological classification and a cross-identification to infrared and optical catalogues.

Catalogue assembly depends on the prediction confidence scores from the Gal-DINO network. For each cutout, the central source with the highest confidence score is prioritised, added to the catalogue, and removed from the *Selavy* catalogue. In cases of extended radio galaxies, multiple components are grouped as a single entry, while for compact sources, a single component is included. The final catalogue integrates both compact and

extended radio galaxies. A cross-matching process is then used to link radio galaxies with their multi-wavelength counterparts from infrared (CatWISE) and optical catalogues, including DES, DESI Legacy Surveys, and SuperCosmos. Gupta et al. (in preparation) present the value added catalogues from the RG-CAT pipeline for the first year’s observations of the full EMU survey.

6.2 Diffuse image pipeline

The EMU data products contain not only a high density of compact and somewhat extended sources, but are also highly sensitive to more diffuse emission (see also Section 4.3). This sensitivity to many angular scales can hinder detection of extremely low brightness diffuse emission, and so it becomes useful to create images containing *only* the diffuse component. A common method for removing compact emission is to image data with a (u,v) cut corresponding to the angular scales to be removed. The resulting CLEAN component model is then subtracted from the visibilities leaving only the extended emission. This process can be computationally expensive, requiring multiple additional imaging rounds on each visibility dataset. For the full EMU survey, we opt to employ an image-based approach to separate out diffuse emission from the embedded compact emission. This process makes use of the filtering algorithm described by Rudnick (2002). The EMU implementation of this pipeline performs two rounds of filtering that scale with the size of the restoring beam (θ_{beam}): a lower filter removing emission $< 3 \theta_{\text{beam}}$ ($45''$) and an upper filter removing emission $> 27 \theta_{\text{beam}}$ ($405''$). A schematic outline of the pipeline is shown in Fig. 12.

The small-scale filter removes point sources and the large-scale filter is intended to remove some of the larger features of the Galactic Plane and large-scale background ripples found in the data. Between application of the two filters, we also convolve the resulting image to $9 \theta_{\text{beam}}$ ($135''$) to help reduce artifacts, although that image is not included as a final product. Finally, we reset the zero level of the images by subtracting the σ -clipped mean value over the image. We note that the zero level is not necessarily accurate at all locations on the image and the filters do not cut off sharply at the filter size. Some emission remains up to approximately three times the size of a particular filter. Note that this process is qualitatively very different from a simple smoothing. If the $15''$ images had merely been convolved to $45''$ resolution, the diffuse emission would have been blended with the compact emission, and their fluxes and extents would be very uncertain.

The pipeline is implemented in a python package^h which also has generic implementations of the filtering algorithm described by Rudnick (2002). It provides a number of options for the user. Caution is advised for quantitative use of these images, and we suggest that those interested in large-scale emission in the Galactic Plane should run the filtering code without the large-scale filter applied.

A comparison of the filtering algorithm with other filtering scales and (u,v) -based methods is provided by Duchesne et al. (2024). Two example cases are shown in Fig. 13; Fig. 13a highlights a typical extragalactic use-case where the filter highlights diffuse emission in the galaxy cluster ACT-CL J0046.4–3911 (Knowles et al. 2021), and Fig. 13b shows an isolated unclassified region of diffuse emission at $(\text{RA}, \text{Dec}) = (19:10:36.2 -71:05)$, with no obvious host. In the bottom example, the average S/N of the diffuse

^hDiffuseFilter: <https://gitlab.com/Sunmish/diffusefilter>.

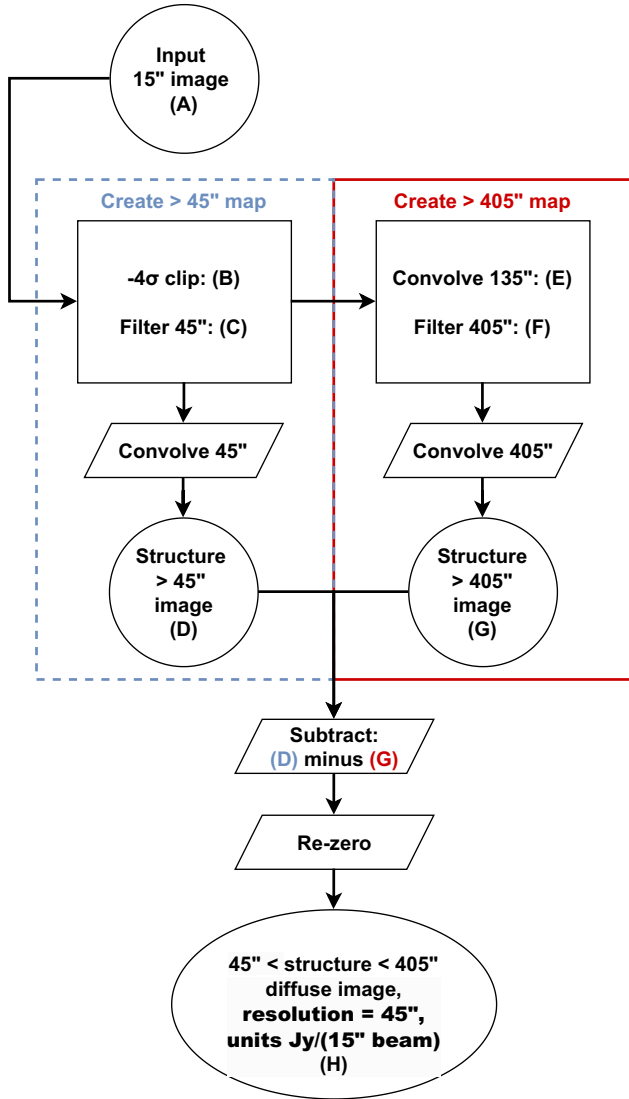


Figure 12. EMU diffuse imaging pipeline, which produces images with structures on scale sizes of $\sim 45''$ to $\sim 405''$. The intermediate steps are used to eliminate artifacts and improve the filtering. The script implementing this, with options for obtaining intermediate step outputs, is described in the text.

emission is ~ 1.6 in the $15''$ image, and rises to ~ 5.5 in the diffuse image.

6.3 Source identification and classification

While $\sim 80\%$ of EMU sources are unresolved (Section 5.1), the remaining $\sim 20\%$ have extended or diffuse morphologies whose study is expected to yield some of the major astrophysical insights from EMU. Traditionally, the identification and classification of sources has been carried out by manual inspection of images, but that is impractical for the large data volumes of EMU. Nevertheless, the manual identification and classification of ~ 3 600 galaxies with complex morphologies in the first EMU Pilot Survey (Norris et al. 2021b) has been conducted by Norris et al. (PASA, submitted), largely to provide a training set for more scalable techniques. Such techniques fall into two broad categories: ML and citizen science.

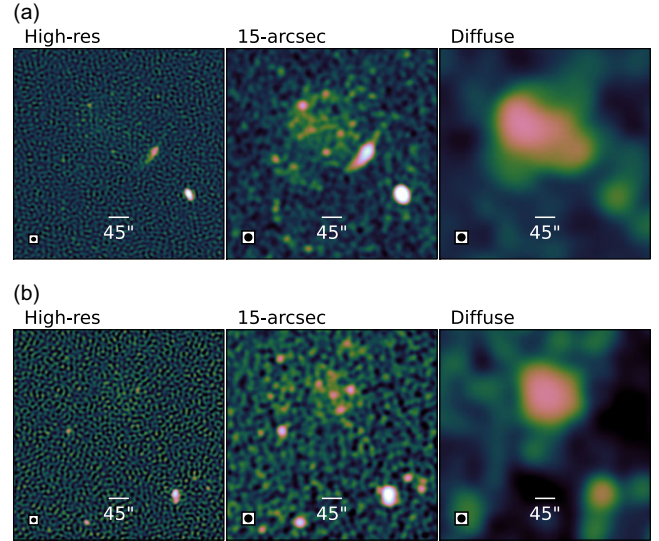


Figure 13. Comparisons of the standard Stokes I output images for (a): the galaxy cluster ACT-CL J0046.4–3911 and (b): a source of unknown origin. For each we show the high resolution image (left), the main image at $15''$ resolution (centre), and the diffuse image (right). The colour scales are linear in the range $[-2\sigma_{rms}, 3\sigma_{rms}]$, and logarithmic in the range $(3\sigma_{rms}, 50\sigma_{rms})$, where σ_{rms} is the rms noise of the particular image.

One innovative approach to classification, with the specific goal of finding complex sources with unusual morphologies, was pioneered by Segal et al. (2019, 2023) who introduced a complexity measure to find unusual sources. An alternative approach to finding peculiar objects is to look for outliers in sources classified using autoencoders (Ralph et al. 2019) or self-organising maps (Galvin et al. 2019, 2020; Mostert et al. 2021; Gupta et al. 2022).

Lochner & Bassett (2021) introduced ASTRONOMALY, a general purpose anomaly detection framework which uses active learning to refine the algorithm. Recently, an extension of this process, ASTRONOMALY: PROTEGE (Lochner & Rudnick 2024), introduced a human trainer to teach the algorithm how to find a broad range of ‘interesting’ sources.

Gupta et al. (2023b) presented a weakly-supervised deep learning algorithm for detecting and classifying radio galaxies. Building on this, as detailed in Section 6.1 above, Gupta et al. (2023a, 2024a,b) developed object detection methods and compared convolutional neural networks and transformer based backbones to classify radio galaxies while simultaneously detecting radio bounding boxes and their infrared hosts. In a recent advance, we developed a multi-modal foundation model to classify EMU radio sources using text and image prompts, deployed and available to the community in EMUSE. These developments largely focus on extragalactic sources, but ML approaches have also been successfully applied to both compact and extended Galactic sources by Riggi et al. (2016, 2024).

An alternative approach to using ML techniques is to enlist the help of thousands of citizen scientists who can classify radio morphologies and multiwavelength identifications by eye. The first such project at radio wavelengths was the EMU-driven Radio Galaxy Zoo project (Banfield et al. (2015), Wong et al., submitted) which resulted in over 100 000 radio source classifications of radio galaxies from the FIRST and ATLAS surveys. A successor project, Radio Galaxy Zoo (LOFAR) was used at scale to

generate source associations and optical identifications for LoTSS (Hardcastle et al. 2023). This work has now been built upon by the Radio Galaxy Zoo EMU project (Bowles et al. 2023), which will be highly complementary to the ML approaches. The project is now also available in Greek, Chinese, and Urdu, making it more accessible to non-English-speaking citizen scientists.¹

An important feature of all these classification schemes is that it is now recognised that classifying a galaxy as ‘Type A’ or ‘Type B’ is restrictive, as in general, a galaxy may have several overlapping classifications. EMU will therefore adopt the approach advocated by Rudnick (2021) in which sources are tagged with multiple labels, rather than being assigned to mutually exclusive classifications.

6.4 Redshifts for EMU

While spectroscopy remains the gold standard of redshift measurement, even the most powerful multi-object spectrograph cannot measure redshifts for more than a small fraction of the tens of millions of sources that will be catalogued by EMU. This motivates a drive to exploit photometric redshifts and other redshift estimators, and to carefully understand their accuracy and any limitations associated with spectral template availability or training sets.

The advent of large optical and infrared surveys has triggered the development of photometric redshift techniques capable of estimating the redshifts of millions of sources. ML plays a significant role in the estimation of these photometric redshifts. Work by Norris et al. (2019) demonstrates that the data used to train the ML model significantly changes the reliability of the final estimations. Where the training sample is not representative, the number of catastrophic failures the ML models produce increases significantly. Given that optically-selected samples have quite different properties (e.g. redshift distribution) from radio-selected samples, the choice of training sample is particularly important in the radio regime, and different studies have chosen different ways of dealing with the selection of appropriate training data.

For example, Duncan (2022) uses as wide a range of sources as possible in their training set, but takes the additional step before training their estimation model to segment the data using a Gaussian mixture model. Individual sparse Gaussian Process regression models are then trained on each segment, allowing for the segment that most closely matches the source needing a redshift to be used as the training sample for the estimation model. As part of the effort to create a single, all-purpose redshift catalogue, Duncan (2022) examined their results for specific rare and under-represented populations, including sources detected in the radio continuum. They find that, when compared with the results by Duncan et al. (2019), the estimated photometric redshifts perform $\sim 10\%$ worse for optically luminous QSOs. Duncan (2022) notes that the worse performance may be due in part to the use of a training set specifically made up of representative sources by Duncan et al. (2019), rather than a general training set segmented into best-matching sources. Creating specific training sets based on the sources being estimated is an alternative method of constructing training sets used, for example, when estimating redshifts for QSO samples (Curran 2022; Curran, Moss, & Perrott 2022; Carvajal et al. 2023), and for radio continuum sources more generally (Luken et al. 2022, 2023). Another recent ML approach combines

information from the imaging together with the catalogue-level spectral energy distributions (Roster et al. 2024) to improve the performance of photo- z estimation for AGNs.

Specifically, Luken et al. (2023) has constructed training sets matching (as closely as possible) the expected properties of the EMU survey, based on data from both the northern hemisphere (with radio sources detected by the NVSS, FIRST, and LoTSS surveys with SDSS optical photometry) and the southern hemisphere (with data from ATLAS, and data taken from ASKAP observations of the Stripe 82 equatorial field, matched with DES optical photometry). Given the differences in optical photometry used, Luken et al. (2023) used the data from the Stripe 82 equatorial field to find corrections between the SDSS and DES photometry. They further tested commonly-used redshift estimation algorithms, finding that simple ML algorithms can out-perform more complex algorithms, albeit without simple methods of obtaining the uncertainties and probability distribution functions (PDFs) required by the community. Future work will include incorporating better uncertainties and PDFs for the simple algorithms, investigating methods of imputing missing data, building on work started by Luken, Padhy, & Wang (2021). The resulting work has now been applied to the first EMU Pilot Survey (Luken et al, in preparation) and will be further developed to be used routinely as part of EMUCAT.

A particular challenge arises from the lack of high-redshift radio sources that have accurately measured redshifts. This results in existing algorithms that perform poorly for high-redshift sources, even if they deliver estimates with small uncertainties at low redshift. While for some sources redshift estimates can be derived from radio continuum data alone (Turner et al. 2020), such approaches are only applicable to extended radio galaxies with broadband spectral data. It is hoped that future spectroscopic surveys, especially those specifically targeting radio continuum selected targets such as ORCHIDSS (Duncan et al. 2023) and WEAVE-LOFAR (Smith et al. 2016), will improve on this, though these will have limited overlap with EMU.

Other approaches can also help, such as the use of drop-outs to estimate the redshift of high-redshift objects (Shobhana et al. 2023). While this technique can only provide a redshift range, it could add significantly to the number of sources identified as high-redshift in EMU, particularly if they can be used as priors in further refining ML algorithms.

Statistical approaches, as well, can be important in constraining the population properties and evolution of the radio sources, even in the absence of individual redshift estimates for each source. The use of clustering properties to estimate redshift distributions and luminosity functions has been demonstrated for both radio (Ménard et al. 2013) and optical (Karademir et al. 2022) samples. More recently, (Prathap et al. 2025, in press) have demonstrated that radio luminosity functions can be robustly estimated from redshift distributions without the use of directly measured redshifts for individual sources.

With other facilities and programs such as WAVES, 4HS, the LSST with the Vera Rubin Observatory, the *Euclid* satellite, and others, there will be extensive redshifts, photometric and otherwise, available to complement these resources as well.

6.5 Examples of the unexpected

As discussed above (Section 2.5), we expected to find sources that were rare enough, or faint enough, that they would not have been found in previous surveys. Many extended radio sources can

¹<https://www.zooniverse.org/projects/hongming-tang/radio-galaxy-zoo-emu>.

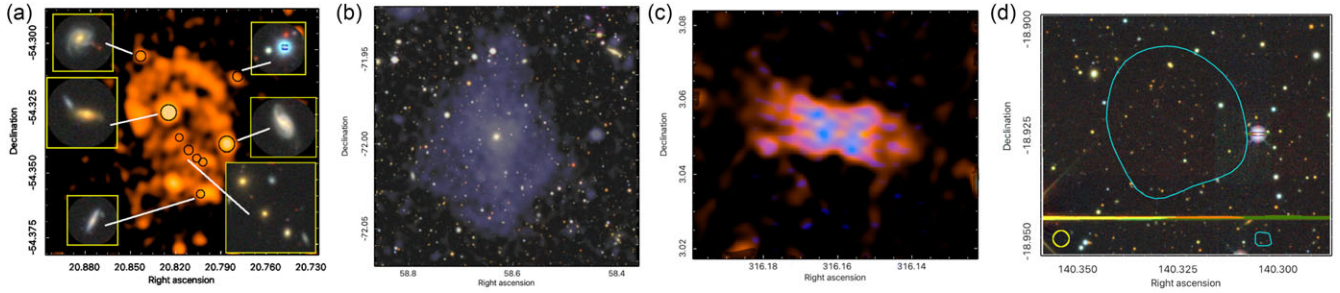


Figure 14. Examples of unusual sources found in EMU. Each represents examples of physical processes that differ from current models. (a) Potentially interacting spiral galaxies; insets drawn from the Legacy Survey. (b) Diffuse radio emission (shown in purple) overlaid on a Legacy Survey image. (c) EMU $\sim 15''$ radio image in red, overlaid with $8''$ image in blue. (d) A single contour from the diffuse image at $10 \mu\text{Jy}/15''$ beam, outlining the location of very low brightness diffuse radio emission, overlaid on the Legacy Survey image.

be explained by ‘weather,’ i.e. variations in the jets or winds or inhomogeneities in the surrounding medium. There are a small subset, though, which go beyond this. As with the ORCs discussed above, this subset contains sources that appear inconsistent with our current physical models, and have the potential to allow genuinely new insight into the associated astrophysical processes. We present four such examples here, shown in Fig. 14.

Our first example, Fig. 14a, highlights a surprising new phenomenon associated with interacting spiral galaxies. While it is well known that spiral galaxies are often radio sources, we have found larger structures surrounding some interacting spirals, resembling intersecting rings. Only one of the galaxies, J01230881–542015 (2MASX J01230901–5420130), the middle inset on the right in Fig. 14a has a redshift. Adopting its value of $z = 0.0729$ for the whole of the extended emission, the structure spans more than 320 kpc, with a luminosity of $\sim 10^{23.3} \text{ W Hz}^{-1}$, similar to the characteristics of radio AGN. While some of the radio emission may arise from background ellipticals in superposition, this is unlikely to be the origin of the structures seen here. A very similar system is seen in the MGCLS survey, and at least two other examples are currently found in EMU. The origin of these extended structures associated with spiral galaxies is completely unknown, and represents new physical processes accelerating relativistic particles in their interactions.

Fig. 14b shows a diffuse structure around an elliptical galaxy that is not what is expected from sources produced by pairs of jets. The host galaxy, with a 1.8 mJy radio core, is catalogued as WISEA J035435.09–715943.9 (2MASX J03543507–7159439), with a $J = 14.38$ at $z = 0.0474$. The quasi-elliptically shaped diffuse radio structure is $420 \times 315 \text{ kpc}$, with a monochromatic luminosity of $10^{23.6} \text{ W Hz}^{-1}$, similar to radio AGN. Quasi-spherical outflows from elliptical galaxies now appear to be responsible for some ORC-like objects (Koribalski *et al.* 2024c), and this diffuse structure could be a related phenomenon. Whether jetted structures could evolve into quasi-elliptical shapes such as seen here would need to be investigated through simulations.

Physically, jets from AGN fall into two broad classes. First are those where the jets maintain their collimation for long distances, dumping their energy at ‘hot-spots’ and generating backflows that appear as ‘lobes’. Other jets expand and slowly fade in brightness as they travel away from the nucleus. The Fanaroff-Riley classification (Fanaroff & Riley 1974) is one attempt to capture this dichotomy. In Fig. 14c, however, we show an example of a radio galaxy that does not fit easily into either scheme. It lacks

terminal hot-spots, but the extended regions near the nucleus are edge-brightened, unlike the expectations for expanding jets. There are some poorly resolved interior structures as well, with hints of a central jet. There is no obvious centrally located host, although a DESI DR10 blue disk-like galaxy with a faint W1 counterpart (WISEA J210436.49+030203.6) appears coincident with one of the radio ‘patches.’ Another example of such a source, with a clear central jet, bright radio core and faint, red host galaxy, is shown in Knowles *et al.* (2022). They argue that none of the proposed mechanisms for creating this lateral brightening appear plausible. In the case shown here, there is no obvious host galaxy. If the source is at a redshift of $z = 0.5$ ($z = 1$) it would have a size of ~ 1 (1.3) Mpc, and a luminosity of $10^{24.3}$ (10^{25}) W Hz^{-1} , making it a very large and extremely luminous radio galaxy. MHD simulations are needed to show how these sources can form, along with observed spectral structure to validate such models.

Using the diffuse images described in Section 6.2 in combination with the $15''$ images, we can find regions of diffuse emission with no compact counterparts. Many of these show double structures with a compact radio source or bright optical galaxy between them, likely the last fading stages of a radio galaxy. There are, however, small isolated regions of diffuse emission with no likely counterparts, occurring with a surface number density approximately one per $50\text{--}100 \text{ deg}^2$. One such region is shown in Fig. 14d. The contour outlines a region of diffuse emission with a brightness of $\sim 71 \pm 7 \mu\text{Jy}$ ($15'' \text{ beam}$) $^{-1}$. It is also faintly visible on the original $15''$ image with a brightness of $56 \pm 33 \mu\text{Jy}$ ($15'' \text{ beam}$) $^{-1}$ and an area of ~ 20 beams. The origins of these isolated diffuse patches are unclear. Among the possibilities are the one remaining lobe of a dying radio galaxy, emission from a poor cluster or group undetected in X-rays or optically, or some completely new phenomenon. Statistical studies of their optical environment and spectral indices are two areas where future studies could provide insights.

7. EMU collaborations

7.1 Coordination with related radio surveys

There are three major radio survey programs that have a close relationship to the EMU survey, which provide complementary data. We describe these here.

7.1.1 POSSUM

POSSUM¹ is a groundbreaking radio polarisation survey designed to observe commensally with EMU, sharing the same observational setup and strategy but with distinct polarimetric calibration and imaging requirements.

POSSUM aims to achieve three primary objectives: (1) generate a dense Faraday rotation measure (RM) grid of up to one million extragalactic sources across EMU/POSSUM's survey area; (2) map the intrinsic linearly-polarised emission and RM properties of a wide range of Galactic and extragalactic objects; and (3) study the diffuse Galactic interstellar medium (ISM) by contributing interferometric data with excellent surface brightness sensitivity, complementing single-dish data. POSSUM will achieve an RM grid density of 30–50 RMs deg⁻² with a median uncertainty of ~ 1 rad m⁻² and an angular resolution of 20'' (since the observing resolution must be smoothed to a common resolution set by the lowest frequency channel of the polarimetric data cubes). POSSUM will also be supplemented by observations covering 1 296–1 440 MHz over 38% of the sky, operating commensally with the WALLABY survey, further enhancing its capabilities.

The POSSUM science case (Gaensler et al. 2025, in press) is highly complementary to that of EMU. POSSUM is focused on probing environments where magnetic fields play crucial astrophysical roles, such as AGN and radio galaxies, galaxy clusters and groups, the cosmic web and intergalactic medium (IGM), the Magellanic System and nearby galaxies, the circumgalactic medium (CGM) and galaxy halos, and the Milky Way's ISM. POSSUM will address two more fundamental questions: How were the first magnetic fields generated? And what processes have sustained, organised, and strengthened these fields through to the present day?

To leverage the complementarity between EMU and POSSUM, the primary POSSUM polarisation catalogue is designed to have a one-to-one correspondence with the EMU catalogue, ensuring that each entry in the latter contains comprehensive polarisation information from the former, whether or not the source components are significantly polarised. Additionally, POSSUM will benefit from EMU's better image fidelity and resolution, particularly through the use of EMU's multifrequency synthesis images smoothed to POSSUM's resolution to create fractional polarisation maps. For resolved sources, the structural details visible in EMU's high resolution images will aid in identifying targets and interpreting POSSUM polarisation maps, and vice versa. The EMUCAT data products from cross-matching sources with other waveband surveys and determining distance estimates will also be invaluable for interpreting POSSUM data. Furthermore, while only a small percentage of EMU sources will yield reliable individual RMs, POSSUM can study the polarisation properties of fainter sources through stacking techniques, aiding population-level studies, and making the most of the complementary datasets.

Finally, a joint project proposal and approval process has been established between EMU and POSSUM to streamline collaborative research efforts, ensuring that the scientific potential of both surveys is fully realised. The policies accommodate cases where data products from either EMU or POSSUM are used in projects led by the other survey, but where the contributions are minor

¹POSSUM is an open collaboration that welcomes applications from qualified astronomers and students. See <https://possum-survey.org/>.

enough that they do not warrant the initiation of a formal joint project.

7.1.2 PEGASUS

The sensitivity of ASKAP to extended emission declines at scales above a few arcminutes and drops to zero at $\approx 43'$ and $\approx 60'$ at the two ends of the EMU band, set by the shortest baseline. About 17% of EMU sources have extended emission (§ 5.1) and $\approx 0.1 - 1\%$ of extragalactic sources (based on results from the first EMU Pilot survey), as many as 20 000–200 000 (or about 20–200 in each tile), have extended emission on scales large enough that they could be mischaracterised. Large objects are poorly imaged and sensitivity to large spatial scales varies with frequency, impacting the reliability of spectral indices of these large objects.

Large Galactic and extragalactic objects, such as SNRs, H II regions, planetary nebulae, Galactic ISM diffuse emission, nearby galaxy cluster intracluster medium (ICM), the Magellanic Clouds, Centaurus A, and giant radio galaxies, lose a significant fraction of their flux if measured with interferometers alone, compromising the estimate of total energy content, spectral index, and the inferred age of their relativistic electrons (Mao et al. 2012; Ingallinera et al. 2014; Govoni et al. 2017; Murgia et al. 2024).

To address this issue with the largest objects, the single-dish survey POSSUM EMU GMIMS All-Stokes UWL Survey (PEGASUS; Parkes project P1123; Carretti et al., in preparation) is being conducted with the CSIRO 64 m Murrumbidgee radio telescope at Parkes. This complements EMU by providing missing large-scale information. PEGASUS uses the Ultra Wide-bandwidth Low (UWL, Hobbs et al. 2020) receiver to map the Stokes parameters I , Q , and U of the entire southern sky up to $\delta = +20^\circ$, in the frequency range 704–1 440 MHz with a resolution of 0.5 MHz. With a diameter of 64 m and an angular resolution of $\approx 22'$ at the EMU centre frequency, PEGASUS data are well suited to be combined with EMU (Kothes et al., in preparation). The observations, begun in 2023, are about 70% complete, and are expected to be finalised later in 2025.

Besides EMU, PEGASUS complements two more projects. It provides large scale information to POSSUM, described above (Section 7.1.1). It also completes the frequency coverage of the Global Magneto-Ionic Medium Survey (GMIMS) project in the south. GMIMS (Wolleben et al. 2019, 2021; Raycheva et al. 2024, Sun et al., submitted) is a spectropolarimetric survey to map the entire sky at 300–1 800 MHz with single-dish radio telescopes and study the properties of the Galactic magnetic fields and of the magneto-ionic medium. PEGASUS is the Mid-Band (704–1 440 MHz) South survey of GMIMS. The Low-Band (300–480 MHz) (Wolleben et al. 2019) and High-Band (1 300–1 800 MHz) South surveys (Raycheva et al. 2024, Sun et al., submitted) have already been completed.

7.1.3 WALLABY

Another ASKAP sky survey that delivers deep radio continuum data is WALLABY (Koribalski 2012; Koribalski et al. 2020). While its main purpose is to map neutral atomic hydrogen (H I) in and between galaxies, it also delivers 1.4 GHz radio continuum images. WALLABY was allocated a total of 8 832 h over five years (Westmeier et al. 2022). With an integration time of 16 h (typically 2×8 h) per field, the WALLABY rms is ~ 1.6 mJy beam⁻¹

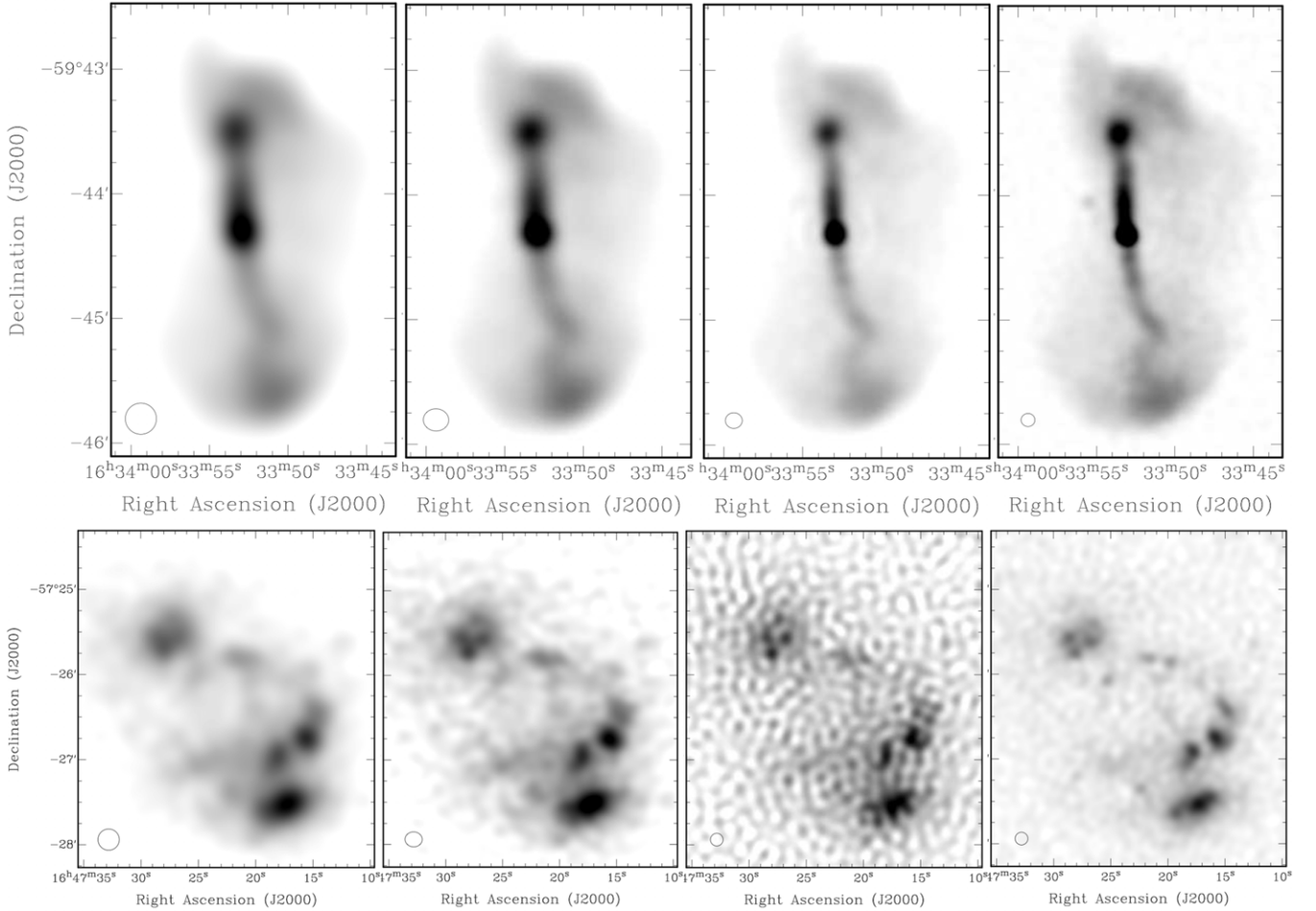


Figure 15. ASKAP radio continuum images of (top) a double-lobe radio galaxy and (bottom) ESO 179-IG013, also known as Kathryn’s Wheel (Parker *et al.* 2015; Paliya & Saikia 2024). The first three images left to right are from EMU, showing the ‘conv’, ‘raw’, and ‘highres’ imaging at 943 MHz respectively. The fourth (rightmost) image is from WALLABY at 1.4 GHz. The strengths and limitations of the EMU ‘highres’ data, and the value of WALLABY continuum imaging as a complement, can be seen in particular in the Kathryn’s Wheel example.

per 4 km s^{-1} channel for the H I 21 cm spectral line ($\sim 30''$ resolution) and $\sim 30 \mu\text{Jy beam}^{-1}$ for the radio continuum ($\sim 8''$ resolution). WALLABY was initially proposed in 2008 as a 21 cm sky survey including H I spectral line, continuum and polarisation data. Due to RFI, only half the available ASKAP bandwidth (i.e. 144 MHz) is currently being processed, providing images over the 1 300–1 440 MHz frequency range. As of April 2025, only $\sim 6\%$ (62) of the 1104 WALLABY fields have been observed and validated as good, due to a delayed survey start and solar avoidance constraints.

WALLABY data products include H I images, spectra and redshifts of galaxies in the nearby Universe (resolution $30''$, 4 km s^{-1}) as well as 1.4 GHz radio continuum images (resolution $\sim 8''$, see Fig. 15). It is worth noting in this comparison that the sensitivity cannot be compared between the EMU ‘high-res’ image and the WALLABY image, although both have similar resolution. The EMU high-res image is not optimised for sensitivity, given it is constructed with uniform weighting, which optimises resolution at the expense of sensitivity. The EMU ‘conv’ image has a comparable sensitivity, from a shorter integration time, but poorer resolution due to the lower observing frequency. High resolution

($12''$) H I cut-outs are also created for pre-selected gas-rich galaxies (Koribalski *et al.* 2020; Murugesan *et al.* 2024), selected from the H I Parkes All-Sky Survey (HIPASS, Koribalski *et al.* 2004; Meyer *et al.* 2004). An example multi-wavelength image is shown in Norris *et al.* (2011, their Figure 8) highlighting the inner, star-forming disk of the spiral galaxy M 83.

Radio continuum images from EMU and WALLABY together deliver extended frequency coverage, useful to determine spectral indices (e.g. Koribalski *et al.* 2024a), and being used jointly in deriving POSSUM data products (Gaensler *et al.*, 2025, in press).

8. Summary and next steps

The Evolutionary Map of the Universe will be the touchstone radio atlas of the southern hemisphere. It will deliver imaging with a median sensitivity of $\sigma = 30 \mu\text{Jy beam}^{-1}$ and a resolution of $15''$ with a 288 MHz bandwidth centred at a frequency of 943 MHz. Source catalogues and value-added data products will be produced for an estimated 20 million extragalactic sources, along with high fidelity imaging of the Galactic Plane and the Magellanic Clouds. This resource is already proving invaluable in supporting scientific

goals spanning galaxy evolution, galaxy environment and large-scale structure including clusters, the astrophysics of supermassive black holes, star formation and stellar evolution, the ISM, IGM, and ICM, cosmology, and much more.

The EMU collaboration consists of more than 400 researchers internationally, from 24 different countries. The collaboration is open to adding new members who are willing to contribute to team activities, and to abide by team policies. The EMU publication policy^k has been developed to encourage collaboration within and between teams, and aims to be flexible in order not to stifle projects and collaborative work. Those wishing to join should contact the EMU management team^l outlining their interest in joining the collaboration, what they expect to contribute, and that they have reviewed and agree to abide by the EMU publication policy. Publicly available ASKAP data, including EMU data, are available through CASDA. EMU is identified as project AS201, and a search on this project number will return all publicly available EMU data. EMU publications are being tracked through an ADS library.^m

The legacy value of EMU is in its large and well-characterised dataset. Well-characterised survey projects such as EMU provide an enormously valuable resource for future work that extends into areas and domains beyond those originally envisaged. As the most sensitive GHz survey to span the whole of the southern hemisphere for the foreseeable future, EMU's legacy value is expected to be vast, in line with the high impact of earlier generations of large sky surveys (e.g. Becker et al. 1995; Condon et al. 1998; Bock et al. 1999; Abazajian et al. 2003).

EMU tiles and *Selavy* catalogue products are publicly available on CASDA, listed under 'Project AS201'. The value-added EMUCAT data are proprietary to the EMU team until distributed through the regularly planned public data releases. These products include the super-mosaics and diffuse images, full-sensitivity de-duplicated catalogues and other catalogue cross-matched and derived data products. EMU observations with ASKAP will continue into 2028 at which stage this set of observations is expected to be completed. We anticipate that a future phase of the project will extend the sky coverage to $\delta = +30^\circ$, in order to deliver on the full potential of such a radio atlas with ASKAP, and to complement other wide-area surveys including VLASS, and LoTSS, along with current and future multiwavelength resources. Such an extension will add substantial value to the cosmology science cases and others, as well as enhancing the already significant legacy value of the project.

Acknowledgements. This scientific work uses data obtained from Inyarrimanha Ilgari Bundara, the CSIRO Murchison Radio-astronomy Observatory. We acknowledge the Wajarri Yamaji People as the Traditional Owners and native title holders of the Observatory site. CSIRO's ASKAP radio telescope is part of the Australia Telescope National Facility (<https://ror.org/05qajvd42>). Operation of ASKAP is funded by the Australian Government with support from the National Collaborative Research Infrastructure Strategy. ASKAP uses the resources of the Pawsey Supercomputing Research Centre. Establishment of ASKAP, Inyarrimanha Ilgari Bundara, the CSIRO Murchison Radio-astronomy Observatory and the Pawsey Supercomputing Research Centre are initiatives of the Australian Government, with support from the Government of Western Australia and the Science and Industry Endowment Fund.

This paper includes archived data obtained through the CSIRO ASKAP Science Data Archive, CASDA (<http://data.csiro.au>).

M.V. acknowledges financial support from the Inter-University Institute for Data Intensive Astronomy (IDIA), a partnership of the University of Cape Town, the University of Pretoria and the University of the Western Cape, and from the South African Department of Science and Innovation's National Research Foundation under the ISARP RADIOMAP Joint Research Scheme (DSI-NRF Grant Number 150551) and the CPRR HIPPO Project (DSI-NRF Grant Number SRUG22031677). R.C. acknowledges support from FCT through Fellowship PD/BD/150455/2019 (PhD: SPACE Doctoral Network PD/00040/2012) and POCH/FSE (EC). D.B. acknowledges support from FCT through Fellowship UI/BD/152315/2021. J.A., R.C. and D.B. acknowledge support from FCT through research grants UIDB/04434/2020 and UIDP/04434/2020 (DOI: 10.54499/UIDB/04434/2020 and DOI: 10.54499/UIDP/04434/2020). M.B. acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG) under Germany's Excellence Strategy – EXC 2121 'Quantum Universe' – 390833306 and the DFG Research Group 'Relativistic Jets'. C.L.H. acknowledges support from the Oxford Hintze Centre for Astrophysical Surveys which is funded through generous support from the Hintze Family Charitable Foundation. C.S.A. acknowledges funding from the Australian Research Council in the form of FT240100498. C.J.R. acknowledges financial support from the German Science Foundation DFG, via the Collaborative Research Center SFB1491 'Cosmic Interacting Matters – From Source to Signal'. M.D.F. and S.L. acknowledge the Australian Research Council funding through grant DP200100784.

Data Availability. See descriptions throughout the text for how to access EMU data and data products.

References

- Abazajian, K., et al. 2003, *AJ*, **126**, 2081
 Ade, P., et al. 2019, *J. Cosmology Astropart. Phys.*, **2019**, 056
 Afonso, J., Hopkins, A., Mobasher, B., & Almeida, C. 2003, *ApJ*, **597**, 269
 Aghamousa, A., et al. 2016, arXiv preprint [arXiv:1611.00036](https://arxiv.org/abs/1611.00036)
 Ahmad, A., et al. 2025, *MNRAS*, **537**, 2868
 Ahmed, U. T., et al. 2024, *PASA*, **41**, e021
 Ahumada, R., et al. 2020, *ApJS*, **249**, 3
 Alonso, D., Bellini, E., Hale, C., Jarvis, M. J., & Schwarz, D. J. 2021, *MNRAS*, **502**, 876
 An, T., et al. 2023, *MNRAS*, **526**, 1809
 Andernach, H., Jiménez-Andrade, E. F., & Willis, A. G. 2021, *Galaxies*, **9**, 99
 Anumarlapudi, A., et al. 2023, *ApJ*, **956**, 28
 Asher, A. D., et al. 2024, *Ap&SS*, **369**, 85
 Asorey, J., & Parkinson, D. 2021, *MNRAS*, **506**, 4121
 Bahr-Kalus, B., et al. 2022, *MNRAS*, **517**, 3785
 Balboni, M., et al. 2023, *A&A*, **679**, A107
 Ball, B. D., et al. 2023, *MNRAS*, **524**, 1396
 Bañados, E., et al. 2021, *ApJ*, **909**, 80
 Banfield, J. K., et al. 2015, *MNRAS*, **453**, 2326
 Becker, R. H., White, R. L., & Helfand, D. J. 1995, *ApJ* **450**, 559
 Bell, M. E., et al. 2016, *MNRAS*, **461**, 908
 Bellstedt, S., & Robotham, A. S. G. 2024, arXiv e-prints, [arXiv:2410.17698](https://arxiv.org/abs/2410.17698)
 Berger, E., et al. 2001, *Nature*, **410**, 338
 Bernal, J. L., et al. 2019, *J. Cosmology Astropart. Phys.*, **2019**, 030
 Best, P. N., & Heckman, T. M. 2012, *MNRAS*, **421**, 1569
 Best, P. N., et al. 2023, *MNRAS*, **523**, 1729
 Bhat, N. D. R., et al. 2023 a, *PASA*, **40**, e021
 Bhat, N. D. R. et al. 2023 b, *PASA*, **40**, e020
 Blake, C., & Wall, J. 2002, *Natur*, **416**, 150
 Boch, T. 2014, Aladin Lite: Lightweight sky atlas for browsers, *Astrophysics Source Code Library*, record ascl:1402.005
 Bock, D. C. J., Large, M. I., & Sadler, E. M. 1999, *AJ*, **117**, 1578
 Böckmann, K., et al. 2023, *A&A*, **677**, A188
 Bojčić, I. S., Filipović, M. D., Urošević, D., Parker, Q. A., & Galvin, T. J. 2021, *MNRAS*, **503**, 2887

^k[http://askap.pbworks.com/w/page/140981337/EMU Publication Policy](http://askap.pbworks.com/w/page/140981337/EMU%20Publication%20Policy).

^lemu_mt@mq.edu.au.

^mhttps://ui.adsabs.harvard.edu/user/libraries/H_YrZLcBTmmw3vGcki-eog.

- Bonaldi, A., et al. 2019, *MNRAS*, **482**, 2
- Bonzini, M., et al. 2013, *MNRAS*, **436**, 3759
- Bordiu, C., et al. 2024, *A&A*, **690**, A53
- Botteon, A., et al. 2020, *MNRAS*, **499**, L11
- Botteon, A., et al. 2022, *A&A*, **660**, A78
- Bowles, M., et al. 2023, *MNRAS*, **522**, 2584
- Boyce, M. M., et al. 2023, *PASA*, **40**, e027
- Bozzetto, L. M., et al. 2023, *MNRAS*, **518**, 2574
- Bradley, A. C., et al. 2025, *PASA*, e032
- Brown, M. J. L., et al. 2017, *ApJ*, **847**, 136
- Brüggen, M., et al. 2021, *A&A*, **647**, A3
- Bufano, F., et al. 2024, *A&A*, **690**, A203
- Bulbul, E., et al. 2024, *A&A*, **685**, L2
- Caleb, M., et al. 2022, *NatAs*, **6**, 828
- Caleb, M., et al. 2024, *NatAs*, **8**, 1159
- Callingham, J. R., et al. 2017, *ApJ*, **836**, 174
- Carli, E., et al. 2024, *MNRAS*, **531**, 2835
- Carvajal, R., et al. 2023, *A&A*, **679**, A101
- Cassano, R., et al. 2012, *A&A*, **548**, A100
- Chen, S., & Schwarz, D. J. 2016, *A&A*, **591**, A135
- Chippendale, A. P., et al. 2015, in 2015 International Conference on Electromagnetics in Advanced Applications (ICEAA), 541
- Chippendale, A. P., et al. 2016, in 2016 International Conference on Electromagnetics in Advanced Applications (ICEAA), 909
- Cid Fernandes, R., et al. 2010, *MNRAS*, **403**, 1036
- Cochrane, R. K., et al. 2023, *MNRAS*, **523**, 6082
- Condon, J. J. 1974, *ApJ*, **188**, 279
- Condon, J. J., Anderson, M. L., & Helou, G. 1991, *ApJ*, **376**, 95
- Condon, J. J., et al. 1998, *AJ* **115**, 1693
- Condon, J. J., et al. 1998, *AJ* **115**, 1693
- Condon, J. J., et al. 2012, *ApJ*, **758**, 23
- Condon, J. J., & Ransom, S. M. 2016, *Essential Radio Astronomy* (Princeton University Press)
- Cook, R. H. W., et al. 2024, *MNRAS*, **531**, 708
- Cotton, W. D., et al. 2024, *MNRAS*, **529**, 2443
- Cowley, M. J., et al. 2016, *MNRAS*, **457**, 629
- Crawford, E. J., et al. 2011, *SerAJ*, **183**, 95
- Crittenden, R. G., & Turok, N. 1996, *Phys. Rev. Lett.*, **76**, 575
- Croton, D. J., et al. 2006, *MNRAS*, **365**, 11
- Curran, S. J. 2022, *MNRAS*, **512**, 2099
- Curran, S. J., Moss, J. P., & Perrott, Y. C. 2022, *MNRAS*, **514**, 1
- Cutri, R. M., et al. 2021, *VizieR Online Data Catalog: ALLWISE Data Release (Cutri+ 2013)*, *VizieR On-line Data Catalog: II/328*. Originally published in: *IPAC/Caltech* (2013)
- Dai, S., et al. 2016, *MNRAS*, **462**, 3115
- Dai, S., Johnston, S., & Hobbs, G. 2017, *MNRAS*, **472**, 1458
- Dalton, G., et al. 2012, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 8446, *Ground-based and Airborne Instrumentation for Astronomy IV*, ed. I. S. McLean, S. K. Ramsay, & H. Takami, 84460P
- Dark Energy Survey Collaboration, et al. 2016, *MNRAS*, **460**, 1270
- Das, S., et al. 2024, *MNRAS*, **531**, 977
- Davies, L. J. M., et al. 2016, *MNRAS*, **461**, 458
- Davies, L. J. M., et al. 2017, *MNRAS*, **466**, 2312
- de Gasperin, F., et al. 2023, *A&A*, **673**, A165
- De Horta, A. Y., et al. 2014, *AJ*, **147**, 162
- de Jong, J., et al. 2022, *A&A*, **668**, A107
- de Jong, J. M. G. H. J., et al. 2024, *A&A*, **689**, A80
- de Jong, J. T. A., et al. 2015, *A&A*, **582**, A62
- de Vries, W. H., et al. 2002, *AJ*, **123**, 1784
- DeBoer, D. R., et al. 2009, *IEEE Proceedings*, **97**, 1507
- Della Bruna, L., et al. 2022, *A&A*, **660**, A77
- Deng, X., et al. 2017, *PASA*, **34**, e026
- DESI Collaboration, et al. 2016, *arXiv e-prints*, [arXiv:1611.00036](https://arxiv.org/abs/1611.00036)
- Dey, A., et al. 2019, *AJ*, **157**, 168
- Di Mascolo, L., et al. 2021, *A&A*, **650**, A153
- Dobie, D., et al. 2024, *MNRAS*, **535**, 909
- Dolag, K., Böss, L. M., Koribalski, B. S., Steinwandel, U. P., & Valentini, M. 2023, *ApJ*, **945**, 74
- Dong, F. A., et al. 2024, *arXiv e-prints*, [arXiv:2407.07480](https://arxiv.org/abs/2407.07480)
- Drake, A. B., et al. 2024, *MNRAS*, **534**, 1107
- Driessen, L. N., et al. 2023, *PASA*, **40**, e036
- Driessen, L. N., et al. 2024, *PASA*, **41**, e084
- Driver, S. P., et al. 2011, *MNRAS*, **413**, 971
- Driver, S. P., et al. 2018, *MNRAS*, **475**, 2891
- Driver, S. P., et al. 2019, *The Msngr*, **175**, 46
- Driver, S. P., et al. 2022, *MNRAS*, **513**, 439
- Drory, N., et al. 2024, *AJ*, **168**, 198
- D’Silva, J. C. J., et al. 2023, *ApJ*, **959**, L18
- Duchesne, S. W., et al. 2021 b, *PASA*, **38**, e010
- Duchesne, S. W., et al. 2024, *PASA*, **41**, e026
- Duchesne, S. W., Johnston-Hollitt, M., Bartalucci, I., Hodgson, T., & Pratt, G. W. 2021 a, *PASA*, **38**, e005
- Duncan, K., et al. 2023, *Msngr*, **190**, 25
- Duncan, K. J., et al. 2019, *A&A*, **622**, A3
- Duncan, K. J. 2022, *MNRAS*, **512**, 3662
- Ekers, R. D. 2009, in *Proceedings of the special session “Accelerating the Rate of Astronomical Discovery” of the 27th IAU General Assembly*. August 11-14 2009. Rio de Janeiro, 7
- Emsellem, E., et al. 2022, *A&A*, **659**, A191
- English, W., Hardcastle, M. J., & Krause, M. G. H. 2016, *MNRAS*, **461**, 2025
- English, W., Hardcastle, M. J., & Krause, M. G. H. 2019, *MNRAS*, **490**, 5807
- Fanaroff, B. L., & Riley, J. M. 1974, *MNRAS*, **167**, 31P
- Filipović, M. D., et al. 2002, *MNRAS*, **335**, 1085
- Filipović, M. D., et al. 2021, *MNRAS*, **507**, 2885
- Filipović, M. D., et al. 2022, *MNRAS*, **512**, 265
- Filipović, M. D., et al. 2023, *AJ*, **166**, 149
- Filipović, M. D., et al. 2024, *PASA*, **41**, e112
- Filipović, M. D., et al. 2025, *arXiv e-prints*, [arXiv:2503.09108](https://arxiv.org/abs/2503.09108)
- For, B.-Q., et al. 2018, *MNRAS*, **480**, 2743
- Gaensler, B. M., et al. 2025, *arXiv e-prints*, [arXiv:2505.08272](https://arxiv.org/abs/2505.08272)
- Gaensler, B. M., Landecker, T. L., Taylor, A. R., & POSSUM COLLABORATION. 2010, in *American Astronomical Society Meeting Abstracts*, Vol. 215, *American Astronomical Society Meeting Abstracts #215*, 470.13
- GAIA Collaboration, et al. 2016, *A&A*, **595**, A1
- GAIA Collaboration, et al. 2023, *A&A*, **674**, A1
- Galvin, T. J., et al. 2019, *PASP*, **131**, 108009
- Galvin, T. J., et al. 2020, *MNRAS*, **497**, 2730
- Gibelyou, C., & Huterer, D. 2012, *MNRAS*, **427**, 1994
- Gil de Paz, A., et al. 2007, *ApJS*, **173**, 185
- Giovannini, G., Tordi, M., & Feretti, L. 1999, *NewA*, **4**, 141
- Gordon, Y. A., et al. 2021, *ApJS*, **255**, 30
- Govoni, F., et al. 2017, *A&A*, **603**, A122
- Govoni, F., et al. 2019, *Sci*, **364**, 981
- Green, D. A. 2024, *arXiv e-prints*, [arXiv:2411.03367](https://arxiv.org/abs/2411.03367)
- Groeneveld, C., et al. 2024, *NatAs*, **8**, 786
- Grundy, J. A., et al. 2023, *PASA*, **40**, e012
- Gruppioni, C., et al. 1999, *MNRAS*, **305**, 297
- Gupta, N., et al. 2022, *PASA*, **39**, e051
- Gupta, N., et al. 2023 b, *PASA*, **40**, e044
- Gupta, N., et al. 2024 b, *PASA*, **41**, e027
- Gupta, N., Hayder, Z., Norris, R. P., Huynh, M., & Petersson, L. 2024 a, *PASA*, **41**, e001
- Gupta, N., Hayder, Z., Norris, R. P., Huynh, M., & Petersson, L. 2023 a, *arXiv e-prints*, [arXiv:2312.06728](https://arxiv.org/abs/2312.06728)
- Gürkan, G., et al. 2022, *MNRAS*, **512**, 6104
- Guzman, J., et al. 2019, *ASKAPsoft: ASKAP science data processor software*, *Astrophysics Source Code Library*, record ascl:1912.003, ascl:1912.003
- Hajela, A., Mooley, K. P., Intema, H. T., & Frail, D. A. 2019, *MNRAS*, **490**, 4898
- Hale, C. L., et al. 2021, *PASA* **38**, e058
- Hale, C. L., et al. 2024, *MNRAS*, [arXiv:2411.04958](https://arxiv.org/abs/2411.04958)
- Hardcastle, M. J., et al. 2019, *A&A*, **622**, A12
- Hardcastle, M. J., et al. 2023, *A&A*, **678**, A151

- Heckman, T. M., & Best, P. N. 2014, *ARA&A*, **52**, 589
- Hobbs, G., et al. 2020, *PASA*, **37**, e012
- Hopkins, A. M., et al. 2003, *AJ*, **125**, 465
- Hopkins, A. M. 2004, *ApJ*, **615**, 209
- Hopkins, A. M. 2018, *PASA*, **35**, e039
- Hopkins, A. M., & Beacom, J. F. 2006, *ApJ*, **651**, 142
- Hopkins, A. M., Mobasher, B., Cram, L., & Rowan-Robinson, M. 1998, *MNRAS*, **296**, 839
- Hotan, A. W., et al. 2021, *PASA*, **38**, e009
- Hurley-Walker, N., et al. 2017, *MNRAS*, **464**, 1146
- Hurley-Walker, N., et al. 2022 a, *Natur*, **601**, 526
- Hurley-Walker, N., et al. 2022 b, *PASA*, **39**, e035
- Hurley-Walker, N., et al. 2023, *Natur*, **619**, 487
- Huynh, M. T., Jackson, C. A., Norris, R. P., & Prandoni, I. 2005, *AJ*, **130**, 1373
- HyeongHan, K., Jee, M. J., Cha, S., & Cho, H. 2024, in *EAS2024, European Astronomical Society Annual Meeting*, 1223
- HyeongHan, K., et al. 2020, *ApJ*, **900**, 127
- Ighina, L., et al. 2021, *A&A*, **647**, L11
- Ingallinera, A., et al. 2014, *MNRAS*, **445**, 4507
- Ingallinera, A., et al. 2022, *MNRAS*, **512**, L21
- Intema, H. T., Jagannathan, P., Mooley, K. P., & Frail, D. A. 2017, *A&A*, **598**, A78
- Ivezić, Ž., et al. 2019, *ApJ*, **873**, 111
- Jarrett, T. H., et al. 2019, *ApJS*, **245**, 25
- Jarrett, T. H., Chester, T., Cutri, R., Schneider, S. E., & Huchra, J. P. 2003, *AJ*, **125**, 525
- Jing, W., et al. 2025, *ApJ*, **980**, 162
- Johnston, S., et al. 2007, *PASA*, **24**, 174
- Johnston, S., et al. 2008, *ExA*, **22**, 151
- Jonas, J., & MeerKAT Team. 2016, in *MeerKAT Science: On the Pathway to the SKA*, 1
- Jones, P. A., & McAdam, W. B. 1992, *ApJS*, **80**, 137
- Joseph, T. D., et al. 2019, *MNRAS*, **490**, 1202
- Kaplan, D. L., et al. 2019, *ApJ*, **884**, 96
- Karademir, G. S., et al. 2022, *MNRAS*, **509**, 5467
- Kempner, J. C., & Sarazin, C. L. 2001, *ApJ*, **548**, 639
- Kennicutt, R. C., et al. 2011, *PASP*, **123**, 1347
- Kennicutt, R. C., & Evans, N. J. 2012, *ARA&A*, **50**, 531
- Kleiner, D., Pimblett, K. A., Jones, D. H., Koribalski, B. S., & Serra, P. 2017, *MNRAS*, **466**, 4692
- KM3NeT Collaboration, et al. 2025, *Nature*, **638**, 376
- Knowles, K., et al. 2021, *MNRAS*, **504**, 1749
- Knowles, K., et al. 2022, *A&A*, **657**, A56
- Kollmeier, J., et al. 2019, *BAAS*, **51**, 274
- Koribalski, B. S., et al. 2004, *AJ*, **128**, 16
- Koribalski, B. S. 2012, *PASA*, **29**, 359
- Koribalski, B. S., et al. 2018, *MNRAS*, **478**, 1611
- Koribalski, B. S., et al. 2020, *Ap&SS*, **365**, 118
- Koribalski, B. S., et al. 2021, *MNRAS*, **505**, L11
- Koribalski, B. S., et al. 2024 a, *MNRAS*, **533**, 608
- Koribalski, B. S., et al. 2024 b, *MNRAS*, **531**, 3357
- Koribalski, B. S., et al. 2024 c, *MNRAS*, **532**, 3682
- Kron, R. G., Koo, D. C., & Windhorst, R. A. 1985, *A&A*, **146**, 38
- Lacy, M., et al. 2020, *PASP*, **132**, 035001
- Lambrides, E., et al. 2024, *ApJ*, **961**, L25
- Lazarević, S., et al. 2024 a, *RNAAS*, **8**, 107
- Lazarević, S., et al. 2024 b, *PASA*, **41**, e032
- Lazio, T. J. W., et al. 2010, *AJ*, **140**, 1995
- Leroy, A. K., et al. 2019, *ApJS*, **244**, 24
- Leroy, A. K., et al. 2021, *ApJS*, **257**, 43
- Liske, J., et al. 2015, *MNRAS*, **452**, 2087
- Lochner, M., & Bassett, B. A. 2021, *A&C*, **36**, 100481
- Lochner, M., & Rudnick, L. 2024, arXiv e-prints, arXiv:2411.04188
- Loi, F., et al. 2023, *A&A*, **672**, A28
- Lourenço, L., et al. 2024, *PASA*, **41**, e012
- Lower, M. E., Dai, S., Johnston, S., & Barr, E. D. 2024, *ApJ*, **967**, L16
- Luken, K. J., et al. 2023, *PASA*, **40**, e039
- Luken, K. J., Norris, R. P., Park, L. A. F., Wang, X. R., & Filipović, M. D. 2022, *A&C*, **39**, 100557
- Luken, K. J., Padhy, R., & Wang, X. R. 2021, in *Machine Learning for Physical Sciences workshop at NeurIPS 2021*, 1
- Macgregor, P. J., et al. 2024, *PASA*, **41**, e050
- Mao, S. A., et al. 2012, *ApJ*, **759**, 25
- Mauch, T., et al. 2003, *MNRAS*, **342**, 1117
- McConnell, D., et al. 2020, *PASA*, **37**, e048
- McMahon, R. G., et al. 2013, *Msngr*, **154**, 35
- Mellinger, A. 2009, *PASP*, **121**, 1180
- Ménard, B., et al. 2013, arXiv e-prints, arXiv:1303.4722
- Merloni, A., et al. 2024, *A&A*, **682**, A34
- Meyer, M. J., et al. 2004, *MNRAS*, **350**, 1195
- Middelberg, E., et al. 2008, *AJ*, **135**, 1276
- Mills, B. Y. 1981, *PASA*, **4**, 156
- Mobasher, B., et al. 2009, *ApJ*, **690**, 1074
- Moss, D., et al. 2007, *MNRAS*, **378**, 995
- Moss, V. A. 2022, in *SciOps 2022: Artificial Intelligence for Science and Operations in Astronomy (SCIOPS)*, 20
- Mostert, R. I. J., et al. 2021, *A&A*, **645**, A89
- Mostert, R. I. J., et al. 2024, *A&A*, **691**, A185
- Mullaney, J. R., et al. 2013, *MNRAS*, **433**, 622
- Murgia, M., et al. 2011, *A&A*, **526**, A148
- Murgia, M., et al. 2024, *MNRAS*, **528**, 6470
- Murphy, T., et al. 2013, *PASA*, **30**, e006
- Murphy, T., et al. 2017, *PASA*, **34**, e020
- Murphy, T., et al. 2021, *PASA*, **38**, e054
- Murugesan, C., et al. 2024, arXiv e-prints, arXiv:2409.13130
- Nishiwaki, K., & Asano, K. 2022, *ApJ*, **934**, 182
- Norris, R. P. 2017 a, *PASA*, **34**, e007
- Norris, R. P. 2017 b, *NatAs*, **1**, 671
- Norris, R. P., et al. 2006, *AJ*, **132**, 2409
- Norris, R. P., et al. 2011, *PASA*, **28**, 215
- Norris, R. P., et al. 2013, *PASA*, **30**, e020
- Norris, R. P., et al. 2019, *PASP*, **131**, 108004
- Norris, R. P., et al. 2021 b, *PASA*, **38**, e046
- Norris, R. P., et al. 2021 c, *PASA*, **38**, e003
- Norris, R. P., et al. 2022, *MNRAS*, **513**, 1300
- Norris, R. P., Crawford, E., & Macgregor, P. 2021a, *Galaxies*, **9**, 83
- Novak, M., et al. 2017, *A&A*, **602**, A5
- Nuza, S. E., Gelsinnis, J., Hoefl, M., & Yepes, G. 2017, *MNRAS*, **470**, 240
- Oayda, O. T., Mittal, V., Lewis, G. F., & Murphy, T. 2024, *MNRAS*, **531**, 4545
- Obenberger, K. S., et al. 2014, *ApJ*, **785**, 27
- Oei, M. S. S. L., et al. 2023, *A&A*, **672**, A163
- Oort, M. J. A. 1987, *A&AS*, **71**, 221
- Padmanabh, P. V., et al. 2023, *MNRAS*, **524**, 1291
- Paliya, V. S., & Saikia, D. J. 2024, *ApJ*, **967**, L26
- Panessa, F., et al. 2019, *NatAs*, **3**, 387
- Parker, Q. A., et al. 2015, *MNRAS*, **452**, 3759
- Paul, S., et al. 2023, *JAA*, **44**, 38
- Pennock, C. M., et al. 2021, *MNRAS*, **506**, 3540
- Pennock, C. M., et al. 2022, *MNRAS*, **515**, 6046
- Pignataro, G. V., et al. 2024, *A&A*, **682**, A105
- Pilbratt, G., et al. 2010, *A&A*, **518**, L1
- PLANCK Collaboration, et al. 2020, *A&A*, **641**, A1
- Prandoni, I., et al. 2000, *A&AS*, **146**, 31
- Prandoni, I., et al. 2018, *MNRAS*, **481**, 4548
- Prathap, J., et al. 2024, *PASA*, **41**, e016
- Prathap, J., et al. 2025, arXiv e-prints, arXiv:2505.11453
- Predehl, P., et al. 2021, *A&A*, **647**, A1
- Predehl, P., et al. 2021, *A&A*, **647**, A1
- Purcell, C., Sun, X., Kaczmarek, J., & POSSUM Team. 2017, in *Astronomical Society of the Pacific Conference Series*, Vol. 512, *Astronomical Data Analysis Software and Systems XXV*, ed. N. P. F. Lorente, K. Shorridge, & R. Wayth, 359
- Quici, B., et al. 2021, *PASA*, **38**, e008
- Raccanelli, A., et al. 2012, *MNRAS*, **424**, 801

- Raccanelli, A., et al. 2017, *PhDU*, **15**, 35
- Ralph, N. O., et al. 2019, *PASP*, **131**, 108011
- Ranasinghe, S., & Leahy, D. 2022, *ApJ*, **940**, 63
- Raycheva, N., et al. 2024, arXiv e-prints, [arXiv:2406.06166](https://arxiv.org/abs/2406.06166)
- Regis, M., et al. 2021, *J. Cosmology Astropart. Phys.*, **2021**, 046
- Reiprich, T. H., et al. 2021 a, *A&A*, **647**, A2
- Reiprich, T. H., et al. 2021 b, *A&A*, **647**, A2
- Riggi, S., et al. 2016, *MNRAS*, **460**, 1486
- Riggi, S., et al. 2024, *PASA*, **41**, e029
- Riseley, C. J., et al. 2022, *MNRAS*, **515**, 1871
- Riseley, C. J., et al. 2025, arXiv e-prints, [arXiv:2503.08840](https://arxiv.org/abs/2503.08840)
- Robotham, A. S. G., & Bested, S. 2024, arXiv e-prints [arXiv:2410.17697](https://arxiv.org/abs/2410.17697)
- Ross, K., et al. 2024, *PASA*, **41**, e054
- Roster, W., et al. 2024, arXiv e-prints [arXiv:2411.07305](https://arxiv.org/abs/2411.07305)
- Rudnick, L. 2002, *PASP*, **114**, 427
- Rudnick, L. 2021, *Galaxies*, **9**, 85
- Rudnick, L., Brown, S., & Williams, L. R. 2007, *ApJ*, **671**, 40
- Sabater, J., et al. 2021, *A&A*, **648**, A2
- Sano, H., et al. 2017, *ApJ*, **843**, 61
- Sasaki, M., et al. 2025, *A&A*, **693**, L15
- Scaramella, R., et al. 2022, *A&A* **662**, A112
- Schawinski, K., et al. 2007, *MNRAS*, **382**, 1415
- Schinnerer, E., et al. 2004, *AJ*, **128**, 1974
- Scranton, R., et al. 2005, *ApJ*, **633**, 589
- Segal, G., et al. 2023, *MNRAS*, **521**, 1429
- Segal, G., Parkinson, D., Norris, R. P., & Swan, J. 2019, *PASP*, **131**, 108007
- Sett, S., Sokolowski, M., Lenc, E., & Bhat, N. D. R. 2024, arXiv e-prints [arXiv:2407.13282](https://arxiv.org/abs/2407.13282),
- Seymour, N., et al. 2008, *MNRAS*, **386**, 1695
- Shabala, S. S., et al. 2012, *MNRAS*, **423**, 59
- Shabala, S. S., et al. 2017, *MNRAS*, **464**, 4706
- Shabala, S. S., et al. 2020, *MNRAS*, **496**, 1706
- Shabala, S. S., et al. 2024, *PASA*, **41**, e024
- Shimwell, T. W., et al. 2019, *A&A*, **622**, A1
- Shimwell, T. W., et al. 2022, *A&A*, **659**, A1
- Shobhana, D., et al. 2023, *MNRAS*, **519**, 4902
- Siewert, T. M., Schmidt-Rubart, M., & Schwarz, D. J. 2021, *A&A*, **653**, A9
- Simonte, M., Andernach, H., Brüggem, M., Miley, G. K., & Barthel, P. 2024, *A&A*, **686**, A21
- Simpson, C., et al. 2006, *MNRAS*, **372**, 741
- Singal, A. K. 2011, *ApJ*, **742**, L23
- Smeaton, Z. J., et al. 2024a, *MNRAS*, **534**, 2918
- Smeaton, Z. J., et al. 2024b, *RNAAS*, **8**, 158
- Smith, D. J. B., et al. 2016, in SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, ed. C. Reylé, J. Richard, L. Cambrésy, M. Deleuil, E. Pécontal, L. Tresse, & I. Vauglin, 271
- Smith, K. M., & Huterer, D. 2010, *MNRAS*, **403**, 2
- Smolčić, V., et al. 2017, *A&A* **602**, A1
- Stözlner, B., Cuoco, A., Lesgourgues, J., & Bilicki, M. 2018, *Phys. Rev. D*, **97**, 063506
- Storm, E., Jeltama, T. E., Spletstoesser, M., & Profumo, S. 2017, *ApJ*, **839**, 33
- Sweijen, F., et al. 2022, *NatAs*, **6**, 350
- Tanidis, K., et al. 2024, arXiv e-prints, [arXiv:2411.05913](https://arxiv.org/abs/2411.05913)
- Taylor, E. N., et al. 2023, *Msngr*, **190**, 46
- Tiwari, P., & Nusser, A. 2016, *J. Cosmology Astropart. Phys.*, **2016**, 062
- Turner, R. J., Drouart, G., Seymour, N., & Shabala, S. S. 2020, *MNRAS*, **499**, 3660
- Turner, R. J., & Shabala, S. S. 2015, *ApJ*, **806**, 59
- Umana, G., et al. 2021, *MNRAS*, **506**, 2232
- ur Rahman, S. F. 2020, *PhW*, **33**, 36
- van Weeren, R. J., et al. 2019, *Space Sci. Rev.*, **215**, 16
- Velović, V., et al. 2022, *MNRAS*, **516**, 1865
- Venturi, T., et al. 2022, *A&A*, **660**, A81
- Vernstrom, T., et al. 2014, *MNRAS*, **440**, 2791
- Vernstrom, T., Norris, R. P., Scott, D., & Wall, J. V. 2015, *MNRAS*, **447**, 2243
- Veronica, A., et al. 2022, *A&A*, **661**, A46
- Wagenveld, J. D., et al. 2024, arXiv e-prints [arXiv:2408.16619](https://arxiv.org/abs/2408.16619)
- Wang, Y., et al. 2023, *MNRAS*, **523**, 5661
- Wang, Z., et al. 2024 b, arXiv e-prints, [arXiv:2409.10316](https://arxiv.org/abs/2409.10316)
- Wang, Z., et al. 2024a, *ApJ*, **961**, 175
- Westmeier, T., et al. 2022, *PASA*, **39**, e058
- White, S. V., et al. 2017, *MNRAS*, **468**, 217
- White, S. V., et al. 2020a, *PASA*, **37**, e018
- White, S. V., et al. 2020b, *PASA*, **37**, e017
- White, S. V., Jarvis, M. J., Häußler, B., & Maddox, N. 2015, *MNRAS*, **448**, 2665
- Whiting, M., & Humphreys, B. 2012, *PASA*, **29**, 371
- Whiting, M. T. 2020, in *Astronomical Society of the Pacific Conference Series*, Vol. 522, *Astronomical Data Analysis Software and Systems XXVII*, ed. P. Ballester, J. Ibsen, M. Solar, & K. Shorridge, 469
- Wild, V., Heckman, T., & Charlot, S. 2010, *MNRAS*, **405**, 933
- Wilkins, S. M., Hopkins, A. M., Trentham, N., & Tojeiro, R. 2008a, *MNRAS*, **391**, 363
- Wilkins, S. M., Trentham, N., & Hopkins, A. M. 2008b, *MNRAS*, **385**, 687
- Willis, A. G., Oosterbaan, C. E., & de Ruiter, H. R. 1976, *A&AS* **25**, 453
- Windhorst, R. A., Fomalont, E. B., Partridge, R. B., & Lowenthal, J. D. 1993, *ApJ*, **405**, 498
- Windhorst, R. A., Miley, G. K., Owen, F. N., Kron, R. G., & Koo, D. C. 1985, *ApJ*, **289**, 494
- Wolleben, M., et al. 2019, *AJ*, **158**, 44
- Wolleben, M., et al. 2021, *AJ*, **162**, 35
- Wright, E. L., et al. 2010, *AJ*, **140**, 1868
- Yates, P. M., Shabala, S. S., & Krause, M. G. H. 2018, *MNRAS*, **480**, 5286
- Yates-Jones, P. M., et al. 2023, *PASA*, **40**, e014
- Yates-Jones, P. M., Shabala, S. S., & Krause, M. G. H. 2021, *MNRAS*, **508**, 5239
- Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, *ApJ*, **554**, 803
- Zangrandi, F., et al. 2024, *A&A*, **692**, A237