Publications of the Astronomical Society of Australia (2026), **43**, e042, 17 pages
doi:10.1017/pasa.2026.10178

CAMBRIDGE UNIVERSITY PRESS

Research Article

Extended radio galaxies in EMU: A comparative look at source-finding techniques

Lachlan J. Barnes[1], Andrew M. Hopkins[1], Yjan Gordon[2], Nikhel Gupta[3], Gary Segal[4], Heinz Andernach[5], Michael J.I. Brown[6], Duncan Farrah[7], Stanislav S. Shabala[8], Sarah V. White[9] and O. Ivy Wong[10]

[1]Macquarie University, Australia, [2]University of Wisconsin-Madison, USA, [3]Space & Astronomy, CSIRO Astronomy and Space Science, Australia, [4]School of Mathematics and Physics, The University of Queensland, Australia, [5]Depto de Astronomia, DCNE, Universidad de Guanajuato, Cjonde Jalisco, Mexico, [6]School of Physics, Monash University, Australia, [7]Department of Physics and Astronomy, University of Hawai'i at Manoa, USA, [8]School of Natural Sciences, University of Tasmania, Australia, [9]International Centre for Radio Astronomy Research, Curtin University, Australia and [10]Space & Astronomy, CSIRO, Australia

Abstract

Extended radio sources present unique challenges for automated detection and classification in wide-field radio surveys. With current surveys such as the Evolutionary Map of the Universe (EMU), robust and scalable methods are essential to identify and catalogue these complex sources. We apply three automatic approaches to detect complex radio emission in EMU observations of the Galaxy And Mass Assembly (GAMA) 09 field (EMU-G09) in order to evaluate their relative strengths and limitations in preparation for large-scale application across future EMU data releases. These include DRAGNHUNTER, designed to detect likely DRAGNs (Double Radio sources associated with Active Galactic Nuclei) from a component catalogue; coarse-grained complexity, a metric designed to highlight regions of complex emission; and RG-CAT, a machine learning pipeline trained on radio sources identified in the EMU pilot survey. We find that together, the three methods recover nearly all extended sources in EMU-G09 but identify largely distinct, partially overlapping subsets, with only 375 sources identified by all finders. This demonstrates that a combination of complementary techniques will be required to achieve a complete census of extended radio sources in future large-scale surveys.

Keywords: Radio continuum: galaxies; galaxies: jets; galaxies: structure

(Received 5 December 2025; revised 26 February 2026; accepted 11 March 2026)

1. Introduction

The study of extended radio sources has expanded rapidly over recent decades. Historically, radio telescopes did not have the sensitivity to detect faint sources, nor the resolution to detect extended sources, limiting the total number of extended radio sources in survey catalogues (see Norris 2017, for a review). Large-scale surveys such as FIRST (Becker, White, & Helfand 1995; White et al. 1997), NVSS (Condon et al. 1998) and LoTSS (Shimwell et al. 2017, 2019, 2022) have increased the known number of radio sources dramatically (cataloguing ∼900 000, ∼1.8 million and ∼4.4 million sources, respectively), many of which show extended emission (∼80%, ∼16%, ∼80%, respectively). Current radio telescope facilities, such as the Square Kilometre Array pathfinders (see Norris 2011; Norris et al. 2013, for more details), and their respective large-scale surveys will further increase the number of known radio sources into the tens of millions.

Double Radio sources associated with Active Galactic Nuclei (DRAGNs; Leahy 1993) dominate the population of extended radio sources at high flux densities (Norris et al. 2025). These systems often show symmetric double-lobed structures with jets, hotspots, lobes, and compact cores coincident with a host galaxy (e.g. Laing & Bridle 2014; Hardcastle & Croston 2020; Ndung'u et al. 2024). Deeper and more sensitive observations reveal an increasingly broader diversity of AGN-powered structures that depart from traditional classifications, such as the traditional Fanaroff–Riley morphology types (Fanaroff & Riley 1974). These include X-shaped or winged radio galaxies (e.g. Gopal-Krishna, Biermann, & Wiita 2003), remnant or dying radio sources with diffuse fading lobes (e.g. Brienza et al. 2017), and giant radio galaxies (GRGs, Dabhade et al. 2017; Dabhade et al. 2020) extending over megaparsec scales. Extended radio emission is also not limited to AGN activity, and arises in a wide range of astrophysical environments. In nearby spiral galaxies, HII regions and supernova remnants contribute compact and filamentary radio features (Condon 1992). On larger scales, galaxy groups and clusters host low surface brightness sources such as radio halos, and relics in the intracluster medium (e.g. Norris et al. 2021; Hopkins et al. 2025). Extended radio morphologies are diverse, often complex, and are shaped by various physical mechanisms beyond classical AGN jets.

As the diversity of extended radio sources becomes more apparent, the process of identifying, cataloguing, and classifying them has become increasingly challenging. Identifying the host galaxies of extended systems also adds further complexity, but is required to derive any physical parameters of the sources (such as size or luminosity). Manual cross-matching is known to be labour-intensive and nonviable for large datasets (e.g. Fan

Corresponding author: Lachlan J. Barnes; Email: lachlan.barnes3@hdr.mq.edu.au

Cite this article: Barnes LJ, Hopkins AM, Gordon Y, Gupta N, Segal G, Andernach H, Brown MJI, Farrah D, Shabala SS, White SV and Wong OI. (2026) Extended radio galaxies in EMU: A comparative look at source-finding techniques. *Publications of the Astronomical Society of Australia* **43**, e042, 1–17. https://doi.org/10.1017/pasa.2026.10178

© The Author(s), 2026. Published by Cambridge University Press on behalf of Astronomical Society of Australia. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

et al. 2015; White et al. 2020a,b; Fan et al. 2020). With modern surveys detecting millions of sources, automated approaches have become essential for scaling source identification and characterisation beyond what is feasible through human inspection alone.

Automatic source detection software such as *Selavy* (Whiting & Humphreys 2012), *PyBDSF* (Python Blob Detector and Source Finder, Mohan & Rafferty 2015), and *CAESAR* (Riggi et al. 2016, 2019) have been incorporated into radio telescope pipelines to accommodate the large influx of radio data. They have been shown to accurately identify and characterise compact sources (e.g. Hopkins et al. 2015; Riggi et al. 2021; Boyce et al. 2023a,b), however there is mixed performance on extended sources. New tools have since been designed to identify and characterise extended radio emission, such as DRAGNHUNTER (Gordon et al. 2023), and coarse-grained complexity (Segal et al. 2019). Various machine learning approaches have also been developed (e.g. Ma et al. 2019; Wu et al. 2019; Galvin et al. 2020; Gupta et al. 2022; Gupta et al. 2024b; Riggi et al. 2023; Alam, Pimbblet, & Gordon 2025), which have been shown to be capable at handling the vast amounts of radio data being obtained and detecting extended sources. Complementary to these approaches, large-scale citizen-science projects such as Radio Galaxy Zoo (RGZ, Banfield et al. 2015; Wong et al. 2025) have shown that crowd-sourced human classifications remain highly effective for identifying complex extended systems, and can provide training sets to further advance machine learning approaches.

Identifying and characterising extended radio sources remains a central challenge in the era of large-scale radio continuum surveys. In this work, we perform a comparative analysis of three independent extended source-finding algorithms, applied to the same region. Rather than comparing to a predefined ground truth, we focus on understanding the relative behaviour of each source finder: the types of extended systems they successfully identify, the cases they miss, and the degree of overlap between their detections. This comparison provides practical insight into the strengths and limitations of current methods, and informs future development and application of extended source-finding techniques for current and upcoming wide-field surveys. This paper is structured as follows: Section 2 describes the multi-wavelength data used in this work. Section 3 introduces and outlines the three extended source-finding algorithms. Section 4 presents a comparative analysis of their performance and discusses implications for future surveys. Where relevant, we adopt a flat ΛCDM cosmology with $H_0 = 70$ km s^{-1}, Mpc^{-1}, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, and $\Omega_k = 0$.

2. Data

2.1. GAMA

Galaxy and Mass Assembly[a] (GAMA, Driver et al. 2011, 2022; Bellstedt et al. 2020) is a multi-wavelength imaging and spectroscopic survey using the AAOmega multi-object spectrograph (Smith et al. 2004) on the Anglo-Australian Telescope (AAT). The GAMA survey covers five separate fields (G02, G09, G12, G15, and G23), spanning a collective area of 286 square degrees. Each region has a survey limit of $r < 19.8$ mag, except for G23, which has a survey limit of $i < 19.2$ mag. Optical spectroscopy has been obtained for ∼300 000 galaxies, with a completeness of ∼98% at the main survey limit (Driver et al. 2011; Liske et al. 2015). The spectra cover 3 750–8 850 Å at a resolving power of $R \sim 1\,300$, yielding typical redshift uncertainties of 50–100 km s^{-1} (Hopkins et al. 2013; Liske et al. 2015). In addition to spectroscopy, GAMA incorporates multi-wavelength photometry, including ultraviolet (GALEX FUV, NUV), optical (*ugriz* from SDSS/KiDS), near-infrared (UKIDSS/VISTA YJHK$_s$), mid-infrared (WISE), and far-infrared/sub-mm (Herschel PACS/SPIRE) imaging (Driver et al. 2016; Bellstedt et al. 2020). The G09 field covers ∼60 square degrees (RA: 129° to 141°, Dec: −2° to +3°), with spectroscopic coverage matching the general GAMA limits and completeness described above.

We merge several public data products available through the GAMA Data Management Units[b] (DMUs; see e.g. Bellstedt et al. 2020; Driver et al. 2022). Specifically, we use the gkvScienceCatv02, which provides target selection, spectroscopic redshifts, and base photometry, ProSpectAGNv02 for AGN-star formation decomposition from SED fitting, StellarMassesGKVv24 for stellar mass and population parameters, and gkvFarIRv03 for far-infrared fluxes. The merged dataset provides a unified set of spectroscopic, photometric, and derived astrophysical parameters for each galaxy. We then restricted this catalogue to the G09 region. The code used to construct this merged GAMA catalogue, hereafter referred to as our GAMA catalogue, is available from the authors upon request.

2.2. EMU

The Australian Square Kilometre Array Pathfinder (ASKAP) is a radio telescope facility located in Western Australia. It is comprised of 36 dishes, each 12 m in diameter, located at the Murchison Radio-astronomy Observatory in Western Australia (Johnston et al. 2007). ASKAP has a phased-array feed (PAF) located at the focus of each antenna (Hotan et al. 2021) and an instantaneous field of view (FOV) of up to 30 square degrees at 800 MHz (Norris et al. 2021), meaning that ASKAP is able to survey the sky approximately thirty times faster than traditional interferometers (Norris 2011). The Evolutionary Map of the Universe[c] (EMU, Norris et al. 2011, 2021; Hopkins et al. 2025) is an ongoing wide-field radio continuum survey that is producing a highly sensitive atlas of the Southern Hemisphere using the ASKAP facility. The EMU survey is conducted at a frequency ∼900 MHz, with a sensitivity of 20–30 μJy beam^{-1} at a resolution of ∼15″. It is anticipated that EMU will detect ∼20 million radio sources, with an estimate of hundreds of thousands showing well-resolved extended emission (Hopkins et al. 2025). As there are no dedicated EMU observations on the G09 region, we combine three adjacent EMU tiles that overlap G09 (Figure 1). These observations form part of the EMU main survey (Hopkins et al. 2025). For all subsequent analysis, we restrict this combined field to the G09 footprint. The resulting EMU field, referred to hereafter as EMU-G09, has observational characteristics similar to those of the EMU Pilot Survey (EMU-PS; Norris et al. 2021), with a central frequency of 944 MHz, a native resolution of 11″ × 13″, and a sensitivity of 1-$\sigma \approx 25$ μJy beam^{-1}.

[a] https://www.gama-survey.org/.

[b] https://www.gama-survey.org/dr4/schema/.

[c] https://emu-survey.org/.

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

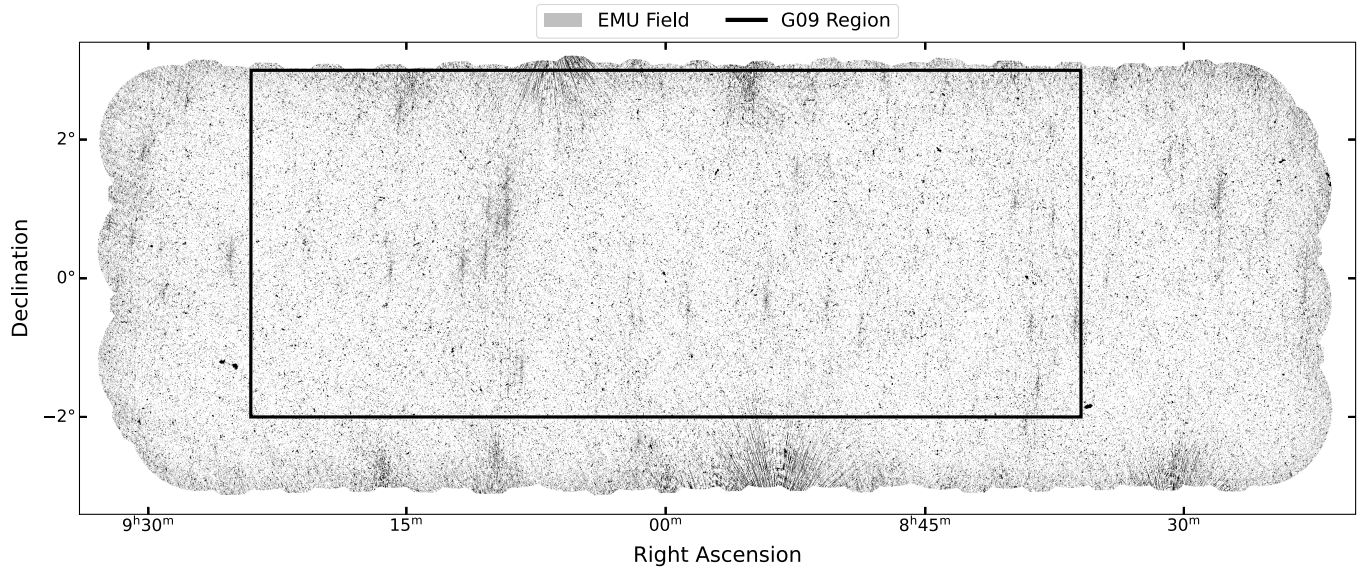


Figure 1. EMU mosaic constructed from the three EMU tiles that overlap the GAMA09 region with the black box indicating the GAMA09 footprint. The EMU observations are centred on a frequency of 944 MHz, a native resolution of 11″ × 13″, and an rms sensitivity of ∼ 25 μJy beam^{-1}. These data constitute the EMU-G09 region used throughout this paper.

3. Source-finding approaches

3.1. *Selavy*

Selavy (Whiting & Humphreys 2012), derived from *Duchamp* (Whiting 2012), is the source-finding algorithm built into the ASKAP Science Data Processing software (ASKAPsoft,[d] Whiting et al. 2017; Norris et al. 2021). *Selavy* identifies groups of pixels with flux densities above a threshold (typically $> 3\sigma$ of the local rms of a radio image, see Section 3.2 of Whiting & Humphreys 2012, for more detail), which are referred to as 'islands'. The *Selavy* island catalogue contains the sizes, positions and flux densities of each island. Gaussian components are then fit to the emission peaks of each island, producing a separate component catalogue (see Norris et al. 2021). *Selavy* identified a total of 34 685 islands in EMU-G09, which were fit with 37 838 components, with approximately 91.1% of islands fit by a single component. While *Selavy* is well-suited to detecting compact radio sources, it is not designed to reliably catalogue extended sources, which can appear as multiple distinct regions on the sky (e.g. two lobes and a core). The island and component catalogues can, however, provide a useful foundation for higher-level algorithms designed to detect more complex emission. Here, we apply three such extended source-finders to explore the range of complex extended radio sources present in EMU-G09. The source finders we focus on in this work are DRAGNHUNTER (Gordon et al. 2023), coarse-grained complexity (Segal et al. 2019, 2023), and RG-CAT (Gupta et al. 2024b). We describe these source finders, and their implementation, in more detail in subsequent sections.

3.2. *DRAGN*HUNTER

DRAGNHUNTER[e] (*DH* hereafter, Gordon et al. 2023) is a source-finding script designed to identify likely DRAGNs, based on an input catalogue of detected radio components. This script was originally designed to be applied to data from the Karl G. Jansky Very Large Array (VLA) Sky Survey (VLASS; Lacy et al. 2020). The *DH* script identifies candidate radio galaxies by searching for nearest-neighbour pairs of extended radio components from an input component catalogue. We make use of the *Selavy* component catalogue of EMU-G09 for this purpose.

The *DH* script requires input parameters defining the minimum component flux density and a minimum angular size. We adopt the 10th percentile (P10) of peak flux density as a conservative threshold to exclude the faintest components, which are more likely to be affected by noise and have poorly constrained fluxes. The P10 value corresponds to a minimum peak flux density of 0.20 mJy, roughly 8σ of the EMU sensitivity (25 μJy beam^{-1}).

We find that ∼7% of the *Selavy* components are unresolved with a deconvolved major axis, Ψ, indistinguishable from zero and an additional ∼9% with $\Psi < 2''$. To exclude the smallest and potentially unreliable components while retaining the majority of sources, we again adopt the 10th percentile (P10) threshold of the range of Ψ values in our sample. The P10 size threshold of corresponds to 3.56″, and was determined after excluding unresolved components, ensuring that only genuinely resolved components contribute to the selection candidate pairs. Using these values, we identify a total 19 320 component pairs from the EMU-G09 *Selavy* component catalogue. A given component may appear in more than one pair if it is the nearest neighbour of multiple candidates. In such cases, the pair with the smallest angular separation is flagged as the preferred pair. This results in a sample of 9 515 preferred pairs.

These preferred component pairs are further refined into likely DRAGNs by considering both the angular separation between components and the mean misalignment between components. Gordon et al. (2023) define the mean misalignment as:

$$\Delta\theta_{\rm mean} = \frac{\Delta\theta_1 + \Delta\theta_2}{2}, \quad (1)$$

where $\Delta\theta_n$ is the relative misalignment (between 0° and 90°) of the position angle of component n, θ_n, relative to the position angle of

[d] https://www.atnf.csiro.au/data-software/software/.
[e] https://github.com/ygordon/dragn_hunter/tree/master/dragnhunter.

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

the pair, $\theta_{\rm pair}$ (see Figure 4 in Gordon et al. 2023), given by:

$$\Delta\theta_n = |\theta_n - \theta_{\rm pair}|. \quad (2)$$

We bin the identified pairs in mean misalignment and measure the local minima of angular separation in each bin (Figure 2), to define where genuine multi-component sources decline and chance pairings begin to dominate. At small angular separations, pairs of components can generally be accepted as belonging to the same source since the probability of a chance association at small separations is low, even if the position angles of the pairs are misaligned. Conversely, for widely separated components, reliable associations preferentially occur when the components are well aligned, as larger separations increase the likelihood of random pairings. Conversely, for widely separated components, reliable associations are found when the components are well aligned, as larger separations increase the likelihood of random pairings.

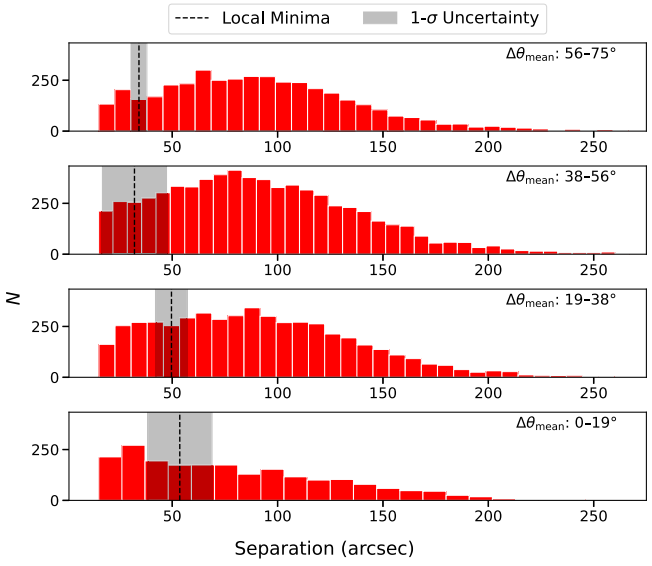


Figure 2. Distributions of angular separation of identified pairs in EMU-G09 (red bars) in bins of mean misalignment. Black lines mark the local minima, with the shaded region showing the 1σ uncertainty of their positions.

To refine the identified pairs into likely DRAGNs, we derive an upper limit in mean misalignment as a function of angular separation:

$$\Delta\theta_{\rm mean,\,EMU} < -213.45\log_{10}(d) + 389.84, \quad (3)$$

where d is the angular separation in arcseconds. Only pairs with $d > 15''$ are considered, corresponding roughly to separations of at least one EMU beam. Pairs below this threshold may belong to a single *Selavy* island or a marginally resolved source (see Appendix A).

Figure 3 shows the identified pairs and the derived EMU upper limit, with the VLASS limit for comparison (Gordon et al. 2023). The cluster of pairs at $\sim$100″ is likely dominated by chance associations, while pairs at smaller separations are more likely to be genuine multi-component sources. Comparing the EMU limit to the VLASS relation, $\Delta\theta_{\rm mean,\,VLASS} < -96.01\log_{10}(d) + 225.32$, we find that the EMU relation has a steeper slope and higher intercept, likely due to differences in resolution and observing frequency ($\sim$15″ vs $\sim$ 2.5″, and 944 MHz vs $\sim$3 GHz, respectively). Using

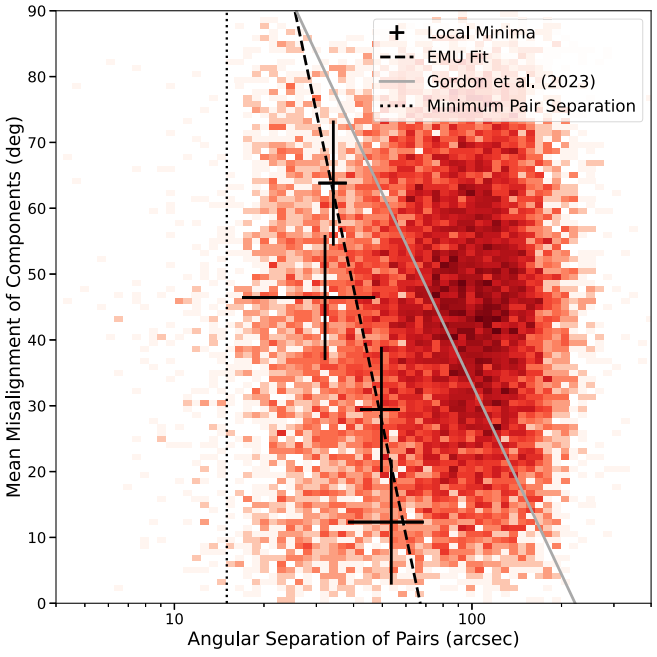


Figure 3. Mean misalignment and angular separation for candidate lobe pairs in EMU-G09. Black crosses show the local minima of pair separation in bins of mean misalignment, with vertical bars indicating bin size and horizontal bars the 1σ uncertainty. The black dashed and grey solid lines show the derived upper limits for likely DRAGNs in EMU and VLASS, respectively, reflecting their different observational parameters. The black dotted line represents the minimum pair separation (15″) used to select DRAGNs in this work.

Equation (3), we identify 2 687 likely DRAGNs from the 9 515 preferred candidate pairs in EMU-G09.

The *DH* script estimates the total flux density, S, defined by Gordon et al. (2023) as the sum of the constituent components of the DRAGN (i.e. both lobes and the core). Note that compact components ($S > 0.20$ mJy and $\Psi < 2''$), which may represent radio cores, are flagged and excluded from the initial lobe–lobe pair-finding stage, such that only extended components are considered when forming candidate pairs. To attempt to recover the missed cores, a search within 30″ or half the pair separation (whichever is smaller) of the central position of the likely DRAGNs is conducted. For most sources however, no core is identified, and the total flux density corresponds to just the sum of both lobes. *DH* also estimates the largest angular size (LAS) of the DRAGN. Here LAS is defined as the distance between the catalogue positions of the pair components, plus the semi-major axis lengths of each lobe.

The distributions of S and LAS are shown in Figure 4. Gordon et al. (2023) found that larger DRAGNs were generally brighter, however we do not explicitly see that trend in our EMU sample. The sources identified by *DH* in EMU-G09 show little to no correlation between size and brightness. The VLASS DRAGNs also occupy a different region of this parameter space, though this is partly a consequence of the differing minimum component–separation thresholds applied to each survey. In EMU we exclude systems with separations below 15″, whereas VLASS can identify structures down to 6″. This means the EMU DRAGNs appear larger in Figure 4 even though this is mostly a selection effect. As shown in the LAS histogram (top panel), the VLASS sample extends to larger angular sizes than our EMU sample. In addition to resolution effects, differences in survey sensitivity also contribute to the separation of the two samples in this parameter

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

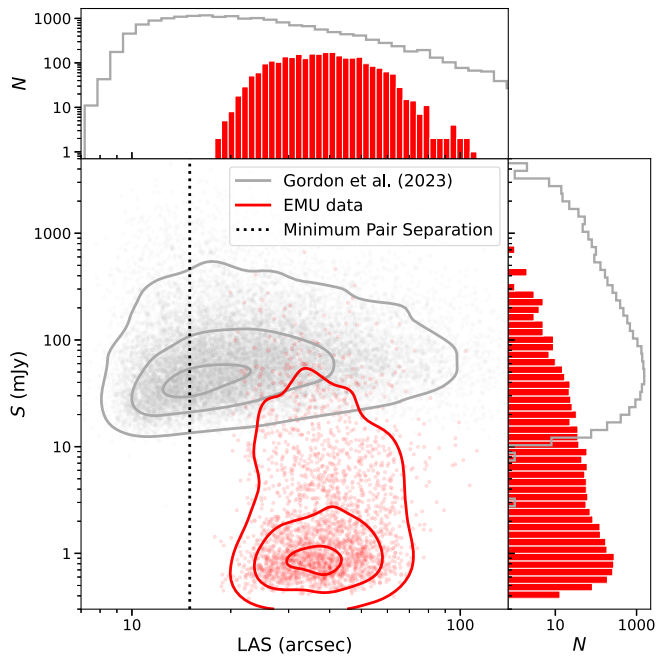


Figure 4. Distributions of LAS and integrated flux density (*S*) for likely DRAGNs in the EMU G09 field. Density contours of the distributions of LAS and *S* for EMU DRAGNs (red) for the VLASS DRAGNs (grey). Contours contain 90%, 50%, and 10% of the data points. The black dotted line represents the minimum pair separation (15″) used to select DRAGNs in this work. The difference between these distributions likely arises from the observational differences between EMU and VLASS.

space. The significantly better sensitivity of EMU relative to VLASS means that the EMU-G09 sources are at systematically lower flux densities than VLASS sources.

3.3. Coarse-grained complexity

The coarse-grained complexity measure (CG-Complexity, Segal et al. 2019, 2023) is a tool that can be applied to large datasets to identify morphologically complex sources in an efficient and automatic way. Coarse-grained complexity is an approximation of the Kolmogorov complexity (see Aaronson, Carroll, & Ouellette 2014, for more details). The CG-Complexity tool operates by first applying a smoothing kernel to an image (e.g. Gaussian or median filter), then compresses the smoothed image using gzip (Levine 2012), treating the byte length as an upper limit of the Kolmogorov complexity (Figure 5).

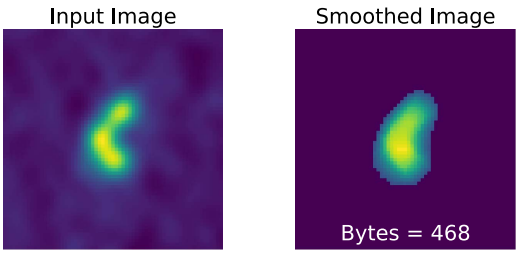


Figure 5. Illustrative example of the coarse-grained complexity process. An input image (left) is smoothed by a Gaussian filter to produce the smoothed image (right) which is then compressed using gzip. The byte length is then used as a proxy for apparent complexity of the radio source. The images here are 64 × 64 pixels, corresponding to an angular size of 128″ × 128″.

Segal et al. (2019) and (2023) use cutouts of 64 × 64 and 256 × 256 pixels, respectively, and employ a blind search to identify complex regions of emission. We do not use a blind approach in this work. We instead create cutouts centred on the right ascension (RA) and declination (Dec) of each *Selavy* island in EMU-G09. This distinction means we identify complex islands rather than complex regions, and therefore retain a dependence on the *Selavy* source detections. When testing the larger cutout size, we found that multiple nearby islands were often included in a single cutout, resulting in repeated or correlated complexity estimates. We therefore adopt the smaller 64 × 64 cutouts, to minimise overlap between adjacent islands. While this means we may not capture the full extent of sources larger than ∼128″ (since one pixel is 2″ × 2″), we expect the obtained complexity values to more accurately represent the apparent complexity of each island. The smaller cutout size also significantly decreases the computation time when applying the coarse-grained complexity calculation to cutouts of the *Selavy* islands.

A manual inspection indicated that many cutouts were contaminated by bright nearby sources towards the edges of the cutouts. Such contamination can artificially increase the measured apparent complexity. To identify and remove these cases systematically, we applied a central crop of each cutout (32 × 32 pixels) and compared the peak pixel intensity within the crop to that of the full 64 × 64 cutout. Instances where the peak intensity in the full cutout was higher than the central crop were flagged as contaminated. This ensures that, in regions with multiple islands (where multiple cutouts may have very similar complexity), cutouts centred on fainter islands are discarded while the brightest island is retained. We identified 10 220 (29.47%) contaminated cutouts, which were excluded from further analysis, leaving 24 465 cutouts of the *Selavy* islands in our sample.

The *Selavy* island catalogue also reports the number of Gaussian components fit to each island. We examine how this relates to the CG-Complexity measure in Figure 6. To quantify the relationship, we compute a Spearman rank correlation (Siegel & Castellan 1988; Conover 1999) using only sources with two or more components, since islands with zero or one component clearly deviate from any linear trend. We obtain a correlation coefficient of 0.281 with a significance of $p = 8.2 \times 10^{-44}$. Although statistically significant, this weak correlation indicates that the number of fitted components is a much poorer proxy for apparent source complexity than the CG-Complexity metric. We discuss the zero-component islands in more detail in Appendix B.

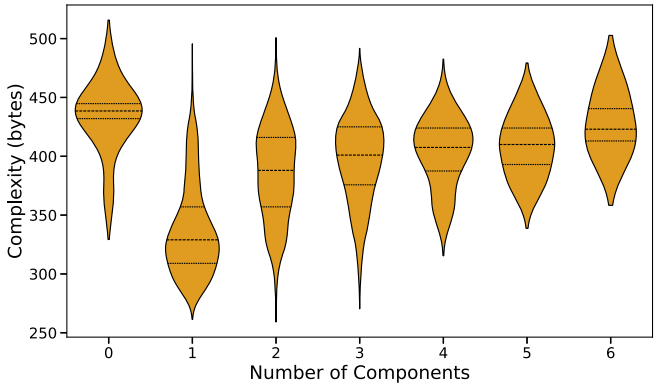


Figure 6. Source complexity as a function of the number of components fit to each *Selavy* island, with quartiles marked inside. Note the general trend of increasing complexity with the number of components.

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

Following Segal et al. (2019), we use Gaussian mixture modelling (GMM) to create a subset of sources with significantly high complexity. We fit a two-component GMM to the complexity values for each of our islands from the *Selavy* catalogue. On the cluster with the highest mean complexity, we then fit another two-component GMM using both the complexity values and signal-to-noise (S/N). Here we adopt a different definition for S/N than Segal et al. (2019), defining it as the peak flux over the background noise (both calculated by *Selavy*) of each island. The distributions of complexities for each subset are highlighted in Figure 7. After fitting both GMMs, we produce a final dataset of 3 055 significantly complex sources with a minimum complexity of 278 bytes. This value is comparable to the ∼300 byte threshold determined by Segal et al. (2019) for significantly complex sources.

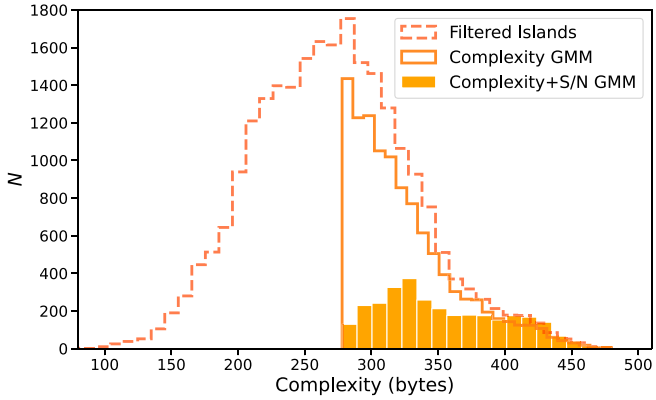


Figure 7. Comparison of distributions of the coarse-grained complexity values for all islands (dark orange dashed line), the complexity-only GMM (orange line), and the complexity+S/N GMM (light orange solid bars).

Figure 8 presents examples of the ten most and least complex sources in the significantly complex sample. The highest-complexity sources encompass a wide variety of morphologies, including radio doubles, triples, slightly bent or asymmetric sources, and diffuse, irregular emission. Extended structure or multiple bright regions clearly contribute to high apparent complexity. In contrast, the lowest-complexity examples are visually compact unresolved point-like sources. Their presence in the significantly complex sample is somewhat unexpected. Even after smoothing, small-scale noise, sharp intensity gradients in bright unresolved sources, or minor residual artefacts may inflate the compressed byte-length of a cutout. Their inclusion likely reflects algorithmic sensitivity to high S/N gradients or subtle pixel-level variations, rather than genuine morphological complexity. We explore these 'complex' point sources further in Appendix C.

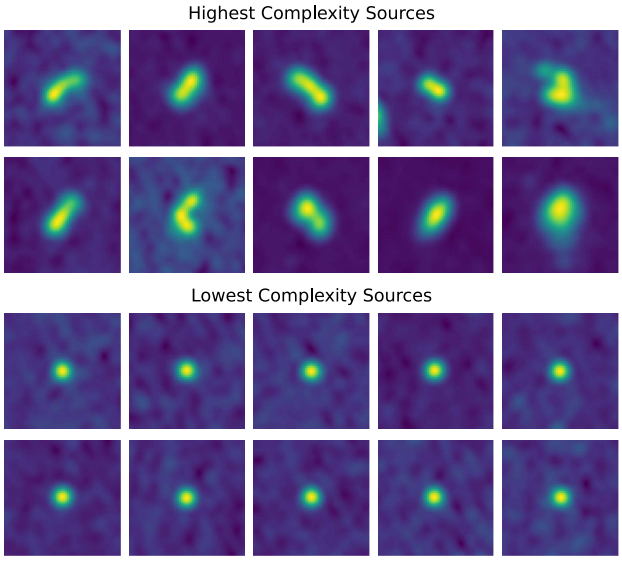


Figure 8. Examples of the ten most (top) and least (bottom) complex sources in the significantly complex dataset. The most complex sources display extended, multi-component, or diffuse morphologies, while the least complex examples are dominated by compact, unresolved sources where apparent complexity likely arises from background noise or faint nearby emission. All cutouts are 64 × 64 pixels, corresponding to an angular size of 128″ × 128″.

3.4. RG-CAT

RG-CAT (Gupta et al. 2024b) is a machine-learning (ML) detection pipeline designed to construct catalogues of radio sources (both extended and compact) from wide-field radio surveys. It leverages computer-vision methods (e.g. Gupta et al. 2023) to identify radio emission, associate related components, and locate potential infrared host galaxies within a unified framework. The model was trained on an extension of the Radio Galaxy NET dataset (Gupta et al. 2024a), containing over 5 000 annotated radio galaxies and their infrared counterparts from the EMU Pilot Survey (Norris et al. 2025). Each 8′ × 8′ cutout in the training set includes detailed annotations: radio morphological classifications (e.g. FR-I, FR-II, compact, or rare types), bounding box parameters, segmentation masks for the radio emission, and host-galaxy positions obtained via infrared cross-matching.

Gal-DINO (as described in Gupta et al. 2024a) was trained on the RG-CAT dataset to simultaneously predict bounding boxes for radio galaxies and potential keypoint positions of their infrared hosts, where a keypoint in ML refers to a specific point or landmark in images. In this work, we apply the Gal-DINO model to the EMU G09 field, without retraining, to produce a catalogue of candidate extended radio galaxies. 35 533 sources were identified in EMU-G09, with 1 823 being extended. The lower panel of Figure 9 illustrates the distribution of galaxies across the six prediction class types. RG-CAT identifies 33 710 compact sources (C), 108 FR-I, 969 FR-II, 265 FR-x (uncertain whether FR-I or FR-II), 446 resolved (R), and 35 peculiar morphologies (Pec). We will focus specifically on the 1 823 extended sources, the 33 710 compact sources identified by RG-CAT are not used in this work.

We show a comparison with the sources (and source types) of our data and the catalogue produced by Gupta et al. (2024b) for the EMU-PS in Table 1. The relative fractions of each morphological class in the G09 field are broadly consistent with those reported by Gupta et al. (2024b), suggesting that the RG-CAT model performs similarly across different sky regions. Both datasets are dominated by FR-II sources, which comprise roughly half of all extended detections, followed by smaller fractions of resolved (R) and FR-x morphologies. The proportions of FR-I and peculiar sources are also comparable between samples. Minor deviations, such as the slightly higher fraction of resolved sources and the lower fraction of peculiar morphologies in G09, likely reflect field-to-field variations in source density and morphology between the EMU-PS and G09 regions.

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

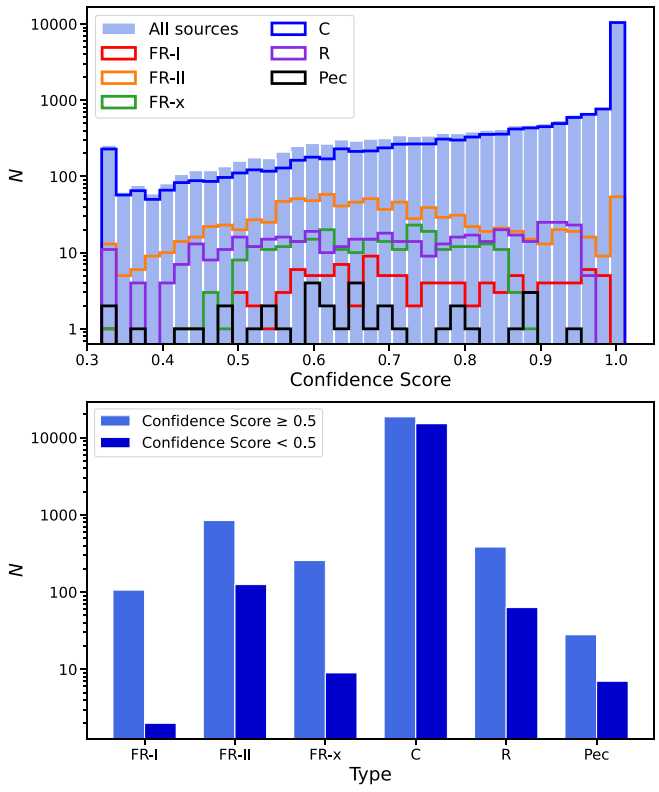


Figure 9. Top panel: Distributions of all sources (solid bars) and subsets of different classification types (coloured steps) based on the confidence score. Bottom panel: Number of different morphology types with confidence scores predicted by Gal-DINO above and below 0.5 (light blue and dark blue bars, respectively).

Table 1. Comparison of RG-CAT morphological class distributions between EMU-G09 sources and EMU-PS sources (Gupta et al. 2024b), with fractions given relative to the total number of extended sources identified in each dataset.

	EMU-G09		EMU-PS	
Type	Count	Fraction (%)	Count	Fraction (%)
FR-I	108	5.9	582	5.6
FR-II	969	53.2	5 602	53.8
FR-x	265	14.5	1 494	14.3
R	446	24.5	2 375	22.8
Pec	35	1.9	361	3.5
Total	1 823	100	10 414	100

4. Source-finder comparison

4.1. Relative performance

In general, the performance of any detection method can be assessed through comparing different outcomes; true positives (*TP*), false positives (*FP*), false negatives (*FN*), or true negatives (*TN*). A *TP* occurs when a real source is correctly identified, while an *FP* arises from a spurious detection. Conversely, an *FN* corresponds to a missed source, and a *TN* indicates correctly identifying empty background. These basic outcomes can be combined to calculate performance metrics (e.g. Cormack & Lynam 2006; Manning, Raghavan, & Schuetze 2008; Beitzel, Jensen, & Frieder 2009; Sortino et al. 2023). Below, we summarise the metrics used to evaluate each source-finder:

$$\text{Recall} = \frac{TP}{TP + FN} \text{ (or Completeness)},$$

$$\text{Precision} = \frac{TP}{TP + FP} \text{ (or Reliability)},$$

$$\text{Informedness} = \frac{TP}{TP + FN} - \frac{FP}{TN + FP}.$$

Here recall is the fraction of the truly positive instances, precision is the proportion of predicted positive instances that are actually correct. Informedness provides a single measure of the detection methods performance relative to random chance. The precision-recall curve, $P(R)$, describes the relationship between precision and recall as some threshold is varied (Beitzel et al. 2009; Gupta et al. 2023). A common evaluation metric, average precision (AP), is determined through integrating the precision-recall curve at a fixed intersection-over-union (IoU) threshold (Manning et al. 2008). The mean average precision (mAP) is then the average of AP values computed across multiple IoU thresholds (Cormack & Lynam 2006; Beitzel et al. 2009):

$$AP = \int_0^1 P(R)dR, \qquad mAP = \frac{1}{n}\sum_n AP_n.$$

Each of the three source-finders considered in this work were evaluated using slightly different sets of performance statistics, as reported in their respective papers. Gordon et al. (2023) report the performance of *DH* in terms of completeness and reliability. They estimate the completeness to be $\geq$45% for sources with $S_{3\,\mathrm{GHz}} >$ 20 mJy, rising to $\geq$85% for $S_{3\,\mathrm{GHz}} >$ 100 mJy. The reliability of their catalogue is reported as $89^{+1.2}_{-1.6}$%. Segal et al. (2019, 2023) determine the performance in a slightly different way, using recall and informedness (reported as 90% and 84%, respectively). Gupta et al. (2024b) instead make use of the average precision measurement to quantify the performance of RG-CAT, with an AP_{50} (average precision at IoU = 0.5) of 73.2% for the radio galaxy bounding box predictions and 71.7% for the infrared host keypoint positions.

In Section 3.4, we showed the performance of RG-CAT for EMU-G09 to be similar to that for EMU-PS in Gupta et al. (2024b). The performance of CG-Complexity is likely to be very similar to the quoted values from Segal et al. (2019), as it was evaluated on radio data with comparable observational parameters. For *DH* however, due to the better sensitivity and poorer resolution of EMU relative to VLASS, we cannot assume that the performance metrics will be the same. It is likely that the *DH* algorithm will produce somewhat more spurious doubles compared to the VLASS proportions, mainly due to the greater sensitivity (i.e. more nearest neighbour pairs). Formally determining the completeness and reliability of *DH* on EMU data would require using either simulated sources or injections into real images. Such analysis is beyond the scope of this work, however we performed a visual inspection of 1 000 randomly selected *DH* detections (~37% of *DH* sources, assessed in ten independent subsets of 100 by L.J.B) to obtain a lower limit on the reliability. Approximately 52.5% of sources appear to correspond to genuine double-lobed radio galaxies. This reliability is likely underestimated, as visual inspection alone, particularly at EMU's sensitivity and resolution, cannot always confirm whether a faint pair of components originates from the same physical system. Excluding sources with component flux

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

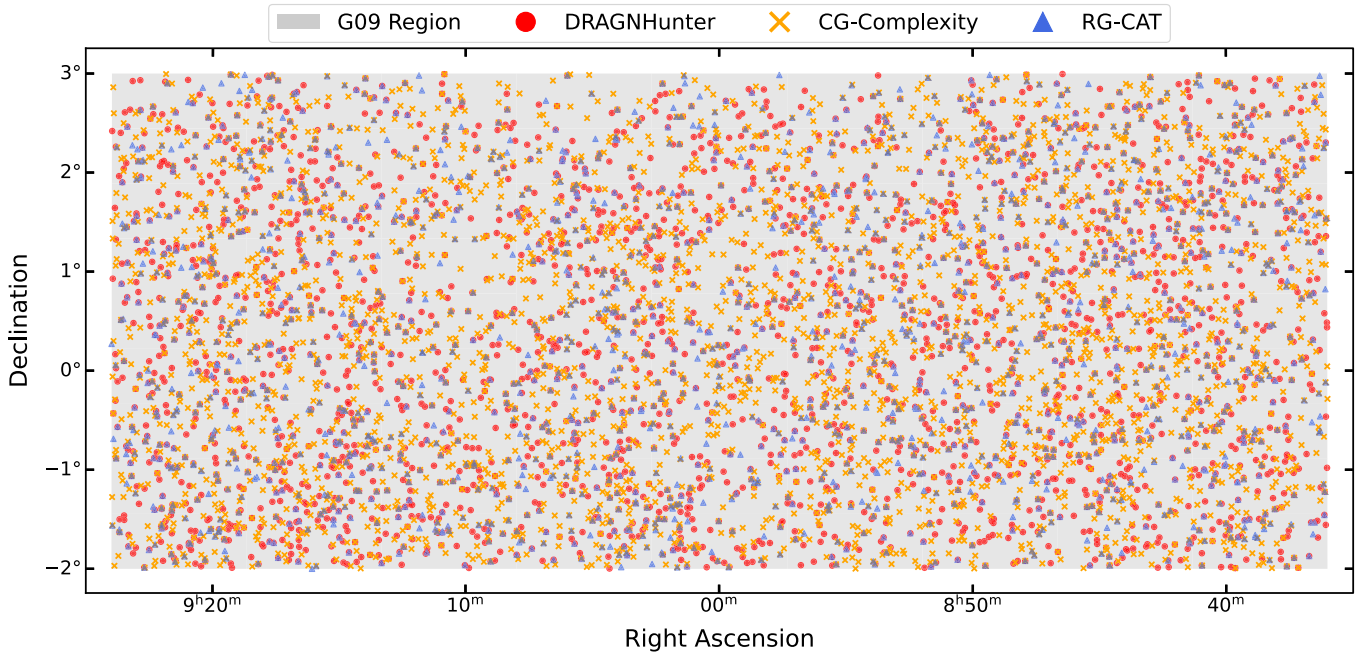


Figure 10. EMU field (grey), overlaid with the sky positions of sources detected by each finder: likely DRAGNs identified by DRAGNHUNTER (red dots), significantly complex regions from CG-Complexity (orange crosses), and extended sources detected by RG-CAT (blue triangles). At this scale, it is evident there is clustering of sources in some regions, and others showing a lack of detections from all finders. There is also very little overlap between each source-finder.

ratios < 10 or low S/N would likely further increase this reliability estimate.

While these performance statistics indicate that each source-finder performs reliably within its own design framework, the differences in depth, resolution, and evaluation criteria, mean that we cannot compare these metrics directly. Through applying each source-finder to the same field, we can instead assess the sensitivities and biases of each approach, and explore the types of radio sources for which each source finder is most effective at identifying.

4.2. Sources detected in G09

We show the positions of the likely DRAGNs (*DH*), highly complex regions (CG-complexity) and extended sources (RG-CAT) detected in the G09 region in Figure 10. The distributions of sources are not uniform across the field. Clustering on square degree scales is evident in certain regions, while other areas show a lack of detections from all finders. This spatial variation reflects underlying large-scale structure (Driver et al. 2011), and the filaments and voids in the radio source population should be investigated in future work. Table 2 summarises the total number of sources detected by each finder, with the corresponding source density calculated for EMU-G09. The higher number of detections from CG-complexity likely reflects its sensitivity to a wide range of morphologies, including irregular or multi-component emission that may not meet the criteria of *DH* and RG-CAT (as well as the detected point sources). For adjacent pairs of well-resolved sources, CG-Complexity may also identify each lobe of a radio galaxy as a significantly complex source.

To determine a reasonable matching radius for comparing sources detected by each finder, we first compute the number of matches for each unique pair of our EMU-G09 source-finder catalogues out to 1°: DRAGNHUNTER to CG-Complexity (DH-CG), DRAGNHUNTER to RG-CAT (DH–RG), and CG-Complexity to RG-CAT (CG-RG), with 1 605, 1 349, and 1 362 matches, respectively. Figure 11 shows the distribution of separations for these matched pairs. In each case, the distribution is bimodal: the first

Table 2. Number of sources detected in EMU-G09 field by each extended-source finder, and the corresponding surface densities.

Approach	Sources	Sources/deg^2
DRAGNHUNTER	2 687	44.8
CG-Complexity	3 055	50.9
RG-CAT	1 823	30.4

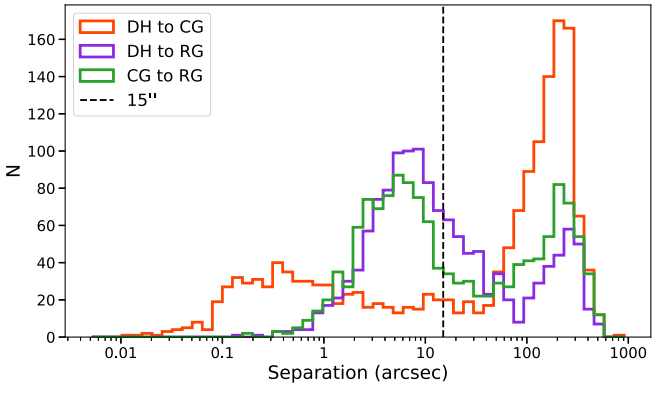


Figure 11. Separation distributions for matches between source-finder catalogues: DH-CG (dark orange), DH-RG (purple), and CG-RG (green). Each distribution is bimodal, where the first peak is likely due to genuine matches and the second is likely dominated by random associations. We adopt a 15″ radius as the threshold for genuine matches.

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

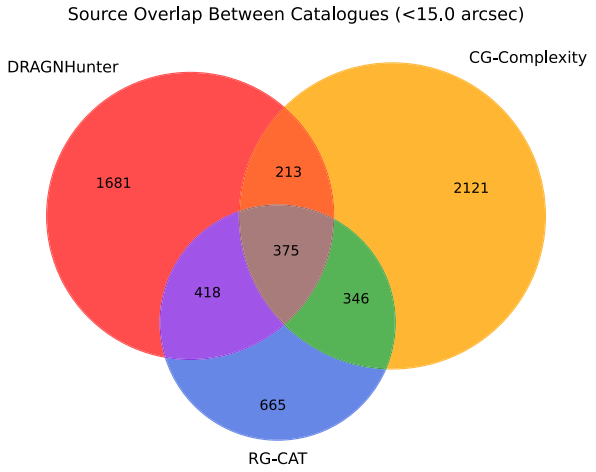


Figure 12. Overlap ($< 15''$) of sources detected by DRAGNHUNTER (red), CG-Complexity (orange), and RG-CAT (blue). Each region of the diagram represents the number of sources uniquely or jointly identified by the corresponding source-finders. Only 375 sources are common to all three, with this small overlap likely reflecting the differing selection biases of each method.

peak at small separations likely corresponds to genuine matches, while the second peak arises from random associations. For all catalogue pairs, the first peak diminishes around $15''$, above which random associations dominate. We therefore adopt a matching radius of $15''$ to define overlap between sources identified by each of our finders.

Detections from two or more source-finders within $15''$ are likely to belong to the same source. Figure 12 establishes that there is limited overlap between each approach with 375 sources being independently detected by all source-finders. This is likely due to the different sources that each source-finder is tailored to find. As the primary selection criteria for *DH* is a nearest neighbour, it may be biased towards smaller sources (i.e. smaller angular separation between components, Figure 4). RG-CAT, due to being trained on manually identified sources, may be biased towards sources that are larger and brighter (i.e. sources more easily identified by-eye). The CG-complexity measure is agnostic to predetermined morphology types, and is instead searching for significantly complex regions. These regions would often coincide with an extended source such as a DRAGN, but also includes irregular diffuse structure. As CG-complexity is applied to the cutouts of the *Selavy* islands, this can possibly explain why there are more detections than *DH* and RG-CAT (i.e. each lobe of a large, well-resolved radio source may be deemed significantly complex).

To verify this, we show cutouts of the CG-Complexity sources, as well as sources from *DH* and RG-CAT in Figure 13. The *DH* sample is dominated by classical double-lobed radio galaxies, with *DH* sources generally appearing to be relatively compact. This is reflective of the nearest-neighbour pairing strategy used by *DH*, which favours close, symmetric component pairs. Many *DH* cutouts also include faint, low S/N sources where Gaussian components have been fitted to marginal detections or possibly to bright noise fluctuations. Conversely, the RG-CAT sources often appear to be slightly larger and brighter than the *DH* sources. CG-Complexity identifies typical radio doubles as well as regions of more intricate or irregular emission. Many cutouts in the random selection of CG-Complexity sources show a simple compact source where, as discussed in Section 3.3, the surrounding structure can increase the overall complexity score.

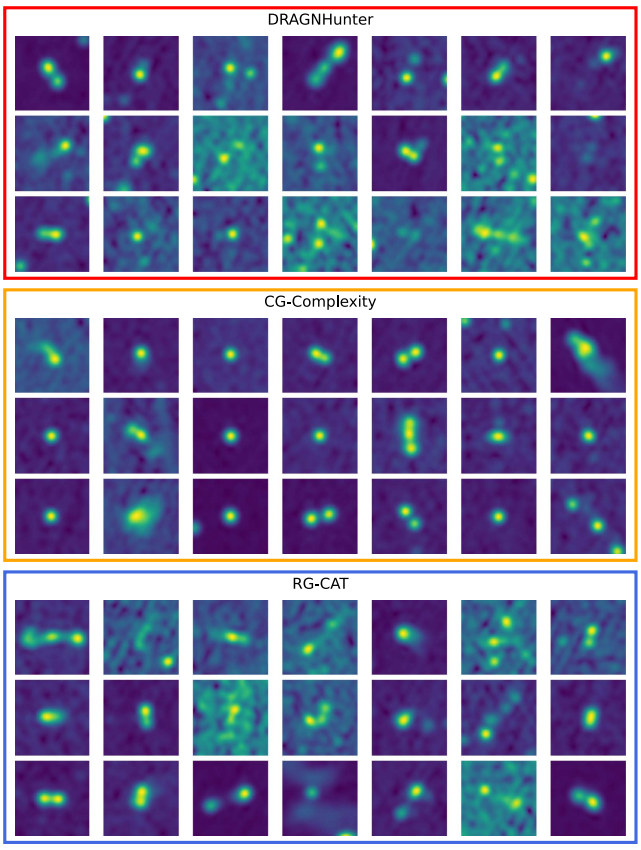


Figure 13. Example sources detected by each source-finding approach. Each panel contains a random selection of 21 sources identified by each of the respective finders. All cutouts are 64×64 pixels, corresponding to an angular size of $128'' \times 128''$. *DH* appears to predominantly identify typical double-lobed structure; CG-Complexity identifies compact sources, typical doubles and more diffuse irregular structure; RG-CAT tends to identify typical doubles as well as more complex extended structure.

To further understand the populations of sources identified by each source-finder, we compare the largest angular size (LAS), and flux density (S) in Figure 14 for all sources. Sources identified by all finders generally exhibit similar flux densities indicating that flux is not the primary driver of detection differences. While there is significant overlap in this parameter space, there are some obvious differences, particularly in LAS, between sources in our sample. The CG-Complexity sources tend to have the smallest angular size when compared to *DH* and RG-CAT sources, with RG-CAT being the largest. This is likely due to how LAS is defined for each source finder. For *DH* sources, LAS is the separation between the component positions, increased slightly by the semi-major size of each component. For the RG-CAT sources, LAS corresponds to the major axis of a bounding box predicted by Gal-DINO, designed to encapsulate the entire source emission. It is therefore likely that LAS of RG-CAT may be slightly overestimated for a given source, and slightly underestimated for *DH* sources. This is also expected given that RG-CAT is trained on visually identified radio galaxies, where bounding boxes are drawn to include all detectable emission, naturally favouring larger angular extents. For the CG-Complexity sources, LAS is the size of the major axis of each *Selavy* island, determined by the number of pixels brighter than minimum flux density threshold. As an island may correspond to a radio galaxy core, or a single lobe it is expected to have a smaller LAS than an entire extended radio source.

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

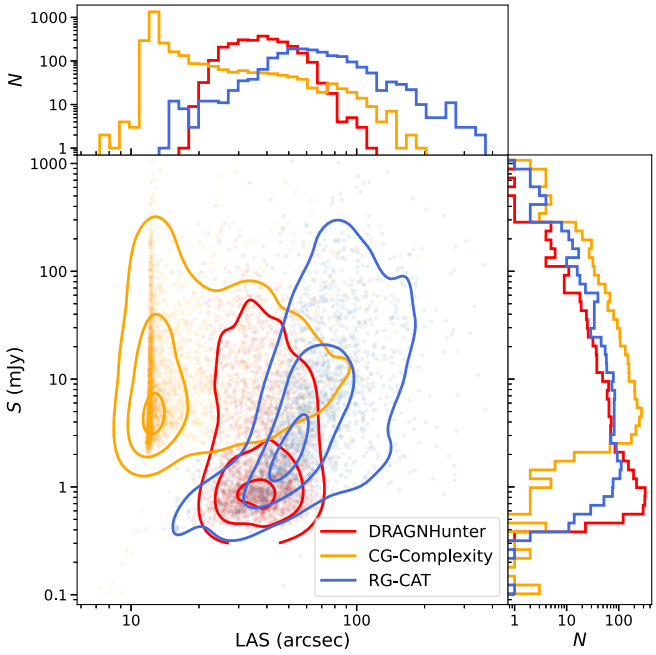


Figure 14. Distributions of LAS and integrated flux density (S) for sources detected by each source-finding approach. Density contours for the distributions of LAS and *S* for DRAGNHUNTER sources are shown in red, high complexity islands in orange, and RG-CAT sources in blue. Contours contain 90%, 50%, and 10% of the data points. While there is substantial overlap in this parameter space, CG-Complexity includes a population of smaller-angular-size sources, whereas RG-CAT tends to detect larger and brighter systems, again likely reflecting selection biases inherent to each method.

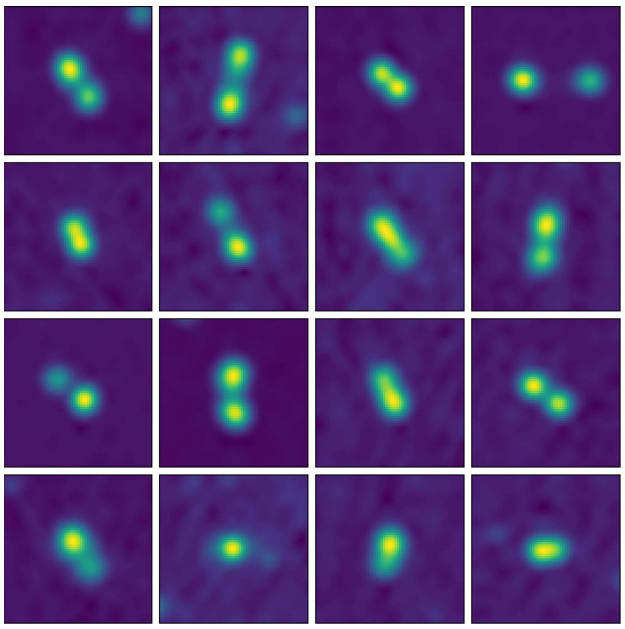


Figure 15. A random selection of 16 sources detected by all source-finders. These objects typically exhibit well-defined double-lobed morphologies of moderate angular size. All cutouts are 64 × 64 pixels, corresponding to an angular size of 128″ × 128″.

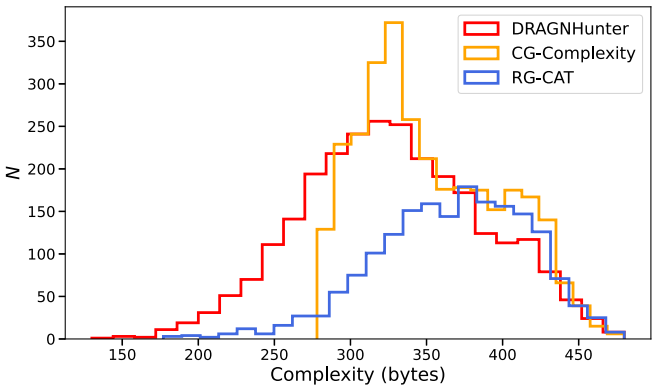


Figure 16. Comparison of calculated complexity of sources detected by DRAGNHUNTER (red), CG-Complexity (orange), and RG-CAT (blue). All sources identified by each approach have comparable apparent complexity, with RG-CAT tending to pick up the higher complexity sources. This likely reflects its training on visually classified, often larger and more irregular sources, while *DH* tends to favour more symmetric, double-lobed morphologies.

In Figure 15 we highlight a representative sample of sources that were detected by all finders. These appear to be classical double-lobed radio galaxies, with relatively bright (average of ∼31 mJy), moderately extended emission (average of ∼51″). Their morphology places them in the narrow regime where the strengths of the three methods overlap: distinct nearby components for *DH*, sufficient structural complexity to be identified by CG-Complexity, and structures similar to observed FR-I/FR-II type sources included in RG-CAT's training set. As a result, they form the most stable and consistently recovered population across the different detection strategies.

However, due to the limited positional overlap between sources identified by each finder (Figure 12), the systematically smaller LAS values measured for CG-Complexity sources (compared to *DH* and RG-CAT) likely reflect the presence of two distinct populations within the G09 field: a relatively compact population dominated by isolated components or cores, and a more extended population of complex radio sources. This interpretation is consistent with results from the GLEAM 4-Jy sample (G4Jy; White et al. 2020a,b, 2025), where the size distribution similarly indicates a more compact, marginally resolved population, and an extended radio source population. The marginally resolved population may also correspond to sources at higher redshift.

We also determined the CG-Complexity of the *DH* and RG-CAT sources by generating 64 × 64 pixel cutouts centred on their positions, applying the same processing steps as for the *Selavy* islands. The resulting distributions are shown in Figure 16. Sources identified by all approaches exhibit broadly comparable apparent complexity, with RG-CAT tending to detect slightly more complex systems than *DH*. This difference likely reflects the design of the two methods: *DH* primarily targets classical double-lobed radio galaxies, which can often appear as two relatively symmetric and circular lobes, and thus show comparatively lower apparent complexity. In contrast, RG-CAT, trained on visually identified DRAGNs, may be biased toward larger and brighter sources that naturally exhibit more extended or irregular structure. Further, the inclusion of morphological tags within RG-CAT, specifically the 'peculiar' class, is reserved for atypical irregular morphologies, therefore allowing it to capture a broader range of complex radio emission than *DH*.

Our results show that each method captures a different subset of the extended radio source population in the G09 field. *DH* appears to be effective at identifying classical double-lobed systems, while RG-CAT recovers slightly larger and morphologically complex objects. CG-Complexity in turn, identifies diffuse or irregular emission that may not follow traditional radio-galaxy

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

morphologies. Although the three approaches together detect nearly all sources in the field, none provides a complete census of the extended-source population on its own. Certain classes, such as giant radio galaxies (GRGs; Dabhade et al. 2017; Dabhade et al. 2020; Andernach 2025), remain difficult to capture without additional or visual inspection. Tailoring algorithms to such rare morphologies could improve completeness for specific classes but would risk biasing detection and increasing false positives in other regimes.

4.3. Multiwavelength cross-matching

4.3.1. WISE

We make use of the Wide-field Infrared Survey Explorer mission (WISE, Wright et al. 2010), which provides coverage for the entire sky in four mid-infrared bands, 3.4, 4.6, 12, and 22 μm ($W1$ to $W4$, respectively). Infrared cross-matching is implemented in a slightly different way for each source-finder presented in this work. The *DH* script incorporates cross-matching to the AllWISE catalogue (Cutri et al. 2012, 2013) for candidate hosts of identified DRAGNs (if one exists), or to the flux weighted centroid position between lobes (for sources without a candidate host). This cross-matching is done by querying the AllWISE database for sources in a 30″ search radius and selecting the most likely AllWISE counterpart via likelihood ratio testing of sources within the search radius. In our sample, the *DH* script identified 1 748 (65.1%) sources with an AllWISE counterpart.

For CG-Complexity sources, we perform a positional cross-match with the AllWISE catalogue using a search radius of 5″ to identify potential nearby hosts. Since the positions of some islands may correspond to single radio lobes, offset from the actual host galaxy, we expect some of these matches to be spurious. However nearby, large, and well resolved sources are much more rare than smaller more compact sources, so we do not expect many single-lobe islands in our sample. We find 1 900 (62.2%) AllWISE counterparts for the highly complex islands. The RG-CAT pipeline incorporates cross-matching with the CatWISE catalogue (Marocco et al. 2021) for identified radio sources, however the CatWISE catalogue only contains $W1$ and $W2$ magnitude measurements. Following Gupta et al. (2024b), we cross-match the CatWISE positions to the AllWISE catalogue within 1″. Not all sources are detected in CatWISE however, and some CatWISE sources may not have an AllWISE match (either because they are too faint for AllWISE, or because AllWISE is a blend of two or more CatWISE sources). If the CatWISE to AllWISE search returns no matches, we then use the radio coordinates with a search radius of 5″. Using this methodology, we find 922 (50.6%) AllWISE counterparts for the extended RG-CAT sources, with 644 from the CatWISE positions and 278 from the radio positions.

The *DH* script includes a likelihood ratio (LR) calculation that provides both an LR and a corresponding reliability estimate for each cross-matched source (see Appendix D for further details). Using the derived reliability values, we find the false-positive rate (fpr) for *DH* sources cross-matched with AllWISE to be approximately $\sim$ 10.2%. For CG-Complexity and RG-CAT sources, we estimate the fpr by artificially shifting both the RA and Dec values of the AllWISE catalogue by 1° and repeat the cross-match as described above. This resulted in 282 and 182 matches for the CG-Complexity and RG-CAT samples, corresponding to an approximate fpr of $\sim$ 9.2% and $\sim$ 10.0%, respectively.

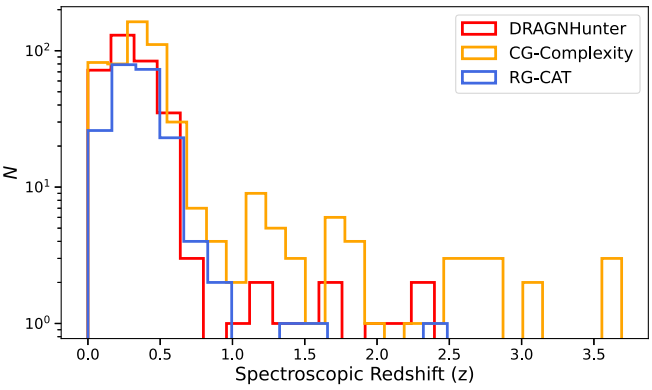


Figure 17. Spectroscopic redshift distributions of *DH*, CG-Complexity and RG-CAT sources cross-matched with GAMA-DR4 data (red, orange, and blue solid lines, respectively). All redshifts here have a quality factor > 2. The *DH* and RG-CAT sources exhibit similar redshift distributions, while CG-Complexity extends to systematically higher redshifts, consistent with its sensitivity to smaller or more compact emission regions that may correspond to unresolved high-redshift doubles.

4.3.2. GAMA

We then match the AllWISE coordinates associated with the *DH*, CG-Complexity, and RG-CAT sources with our GAMA catalogue (Section 2.1) using a 2″ search radius. This resulted in 675, 690, and 355 matches, for *DH*, CG-Complexity and RG-CAT, respectively. The distributions of spectroscopic redshifts (spec-z) from GAMA are shown in Figure 17. Only spec-z values with a redshift quality > 2 are shown (335, 524, and 210 sources for *DH*, CG-Complexity, and RG-CAT, respectively). For subsequent analyses requiring redshift, we only use sources with a quality factor> 2. The redshift distributions of the *DH* and RG-CAT sources are closely aligned, whereas CG-Complexity sources are found to extend to systematically higher redshift. This difference likely reflects a combination of selection and methodological effects. While some contribution may arise from spurious or chance associations in the CG sample, this alone cannot account for the observed shift. As CG-Complexity operates directly on *Selavy* islands, high-redshift doubles that are unresolved at EMU resolution may be represented as a single, moderately extended component. Such structures can still exhibit non-uniform or irregular brightness distributions that register as 'complex' by the CG-Complexity metric, even when not recognised as multi-component sources by algorithms like *DH* or RG-CAT.

5. Astrophysical properties of sources in G09

We determine both the luminosity ($L_{944\,\mathrm{MHz}}$) and the largest linear size (LLS) for our sample of sources (Figure 18). There is significant overlap between the RG-CAT and *DH* sources, however as before, the RG-CAT sources tend to be slightly larger than *DH* sources. The high complexity sources, conversely, always tend to be smaller than both *DH* and RG-CAT sources. Particularly for the *DH* and CG-Complexity sources, we see there is a steep increase in the luminosity of the sources at $L_{944\,\mathrm{MHz}} \approx 10^{24}$ W/Hz, which may be indicative of two populations (i.e. FR-I/FR-II, for example).

We adopt the stellar masses estimated by Taylor et al. (2011) determined using a Chabrier initial mass function (IMF), of the host galaxies identified for each source finder. We show the mass distributions in Figure 19. All three distributions are broadly similar, with each peaking around $10^{11.4}$ $M_\odot$, indicating that the typical host of an extended radio source in the G09 field is a massive

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

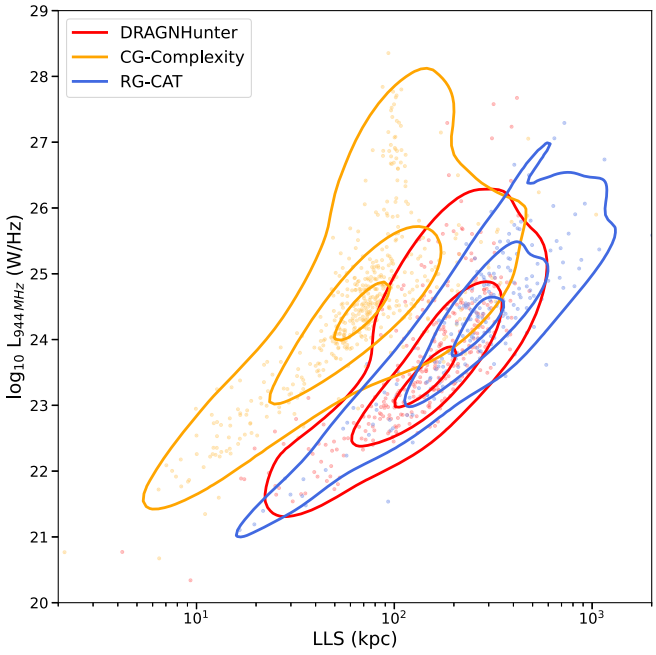


Figure 18. Distributions of LLS and radio luminosity for sources identified by *DH* (red), CG-Complexity (orange), and RG-CAT (blue). Contours enclose 90%, 50%, and 10% of the data points. LLS values are calculated using LAS, which is defined differently for each finder (see Section 4.2)

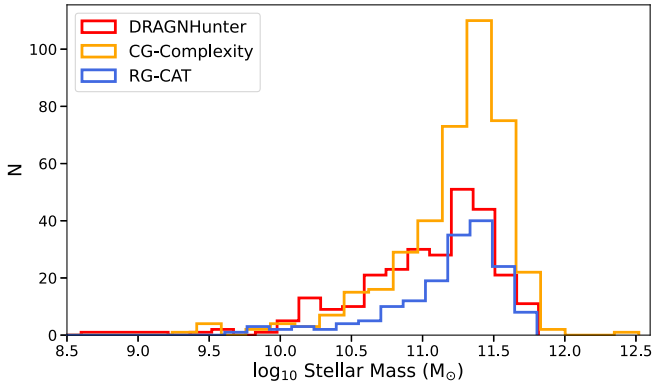


Figure 19. Stellar mass distributions of *DH*, CG-Complexity and RG-CAT sources cross-matched with GAMA-DR4 data (red, orange, and blue solid lines, respectively). All three distributions are broadly similar, peaking around $10^{11.4}$ $M_{\odot}$.

galaxy. *DH* tends to identify more low mass sources than RG-CAT, with a number of low-mass hosts comparable to those from CG-Complexity. Despite the largely distinct source samples identified by each finder, this similarity in stellar mass suggests that the underlying galaxy populations are comparable in terms of their fundamental properties, and that differences in source detection are primarily driven by the morphological or structural criteria of the source-finders rather than by host galaxy characteristics.

We implement the WISE colour diagnostic ($W1 - W2$ versus $W2 - W3$ magnitudes, Lake et al. 2012) to classify the infrared (IR) properties of the host galaxies for sources identified by each respective source-finder. Here we adopt the limits defined by Mingo et al. (2016), where the $W1 - W2$ colour separates AGN dominated systems (> 0.5) from those dominated by host-galaxy light (< 0.5). For host-dominated sources, the $W2 - W3$ colour

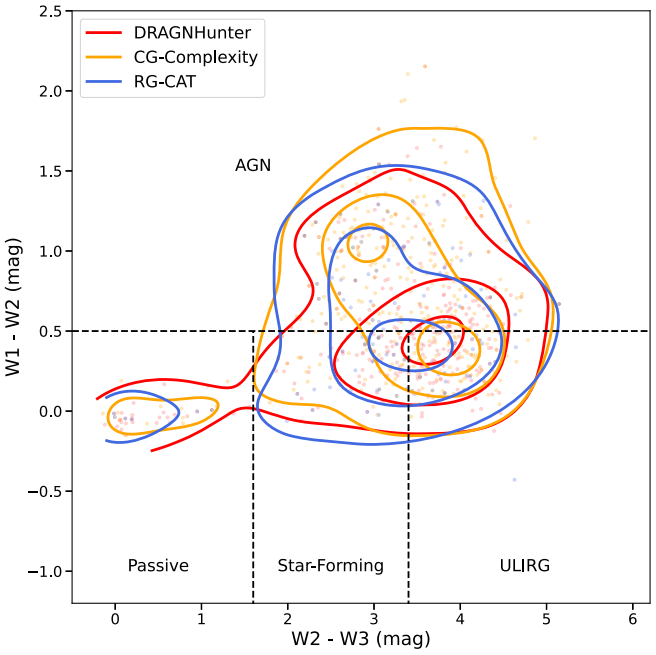


Figure 20. The WISE colour distributions for the hosts of sources identified by *DH* (red), CG-Complexity (orange), and RG-CAT (blue). The contours contain 90%, 50%, and 10% of the data points. Despite differences in detection strategy, sources identified by all three approaches have similar infrared host populations, with many sources classified in the star-forming/ULIRG region of WISE colour space. Note, only sources with a S/N > 3 in $W1$, $W2$, and $W3$ are shown.

further distinguishes between passive/elliptical galaxies (< 1.6), star-forming galaxies (1.6–3.4), and starburst systems (> 3.4), which often correspond with ultra-luminous infrared galaxies (ULIRG). We only apply these classifications to sources with S/N in all three bands >3. While this requirement biases the distribution against passive galaxies, whose $W3$ emission is often several magnitudes fainter than $W1$ or $W2$ (e.g. Cluver et al. 2020), the higher S/N threshold means that the classifications we do obtain are relatively robust. Therefore, the resulting colour distribution should be interpreted as fairly reliable, but not representative of the entire sample.

The distributions of WISE colour classifications for each source-finder are shown in Figure 20, with corresponding percentages in Table 3. Our sample contains a slightly higher fraction of star-forming hosts than reported by Gordon et al. (2023) and Gupta et al. (2024b). Our proportion of star-forming hosts are however, consistent with the findings of Mingo et al. (2016) and Mingo et al. (2019), determined for classical FR-type and bent radio sources. The infrared hosts identified for CG-Complexity sources show the largest AGN fraction, with ∼60% lying in the AGN region, and the smallest proportion of star-forming hosts, with only 11.8%. This may reflect CG-Complexity's capacity to identify sources at higher redshift. *DH* and RG-CAT show similar colour distributions, each with 41–49% in the AGN and 19–23% star-forming regions, along with comparable ULIRG fractions. Passive galaxies comprise only a small minority (4.7–8.7%) across all source-finders.

To further investigate the proportions from the WISE-colour classifications in our sample, we compare the radio luminosity and $W3$ luminosity (Figure 21). Most sources lie above the 1:1 reference line, indicating a radio excess relative to their mid-infrared emission, except for the majority of ULIRG-classified

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

Table 3. The top section highlights the percentage of sources detected in AllWISE with S/N in *W*1, *W*2 and *W*3 > 3, and those that are unclassified (either insufficient S/N or no detection in one or more bands), relative to the total number of AllWISE matches (shown under each source-finder, respectively). The bottom section shows the percentages of WISE colour classifications (AGN, Passive, Star-Forming, U-LIRG) for the S/N > 3 subset, relative to the number of sources in the S/N > 3 subset (shown under each source-finder, respectively). For all percentages, we include 1σ binomial uncertainties.

	Sources detected in AllWISE	
Source-finder	S/N > 3 (%)	Unclassified (%)
DRAGNHUNTER (AllWISE = 1 748)	18.48 ± 0.93	81.52 ± 0.93
CG-Complexity (AllWISE = 1 900)	16.95 ± 0.86	83.05 ± 0.86
RG-CAT (AllWISE = 922)	12.26 ± 1.08	87.74 ± 1.08

	WISE colour classifications			
Source-finder	AGN (%)	Passive (%)	Star-forming (%)	ULIRG (%)
DRAGNHUNTER (*N* = 323)	41.49 ± 2.74	8.98 ± 1.59	19.50 ± 2.20	30.03 ± 2.55
CG-Complexity (*N* = 322)	61.80 ± 2.71	4.66 ± 1.17	11.80 ± 1.80	21.43 ± 2.29
RG-CAT (*N* = 113)	49.56 ± 4.70	5.31 ± 2.11	23.01 ± 3.96	22.12 ± 3.90

sources. Many of the WISE-AGN sources are identified by CG-Complexity, consistent with the idea that it detects compact or marginally resolved (potentially high redshift) sources. A small number of sources lie close to the distribution of star-forming galaxies, suggesting that some are genuinely powered by star formation. However, the overall population is dominated by systems showing significant radio excess, consistent with AGN-driven radio emission.

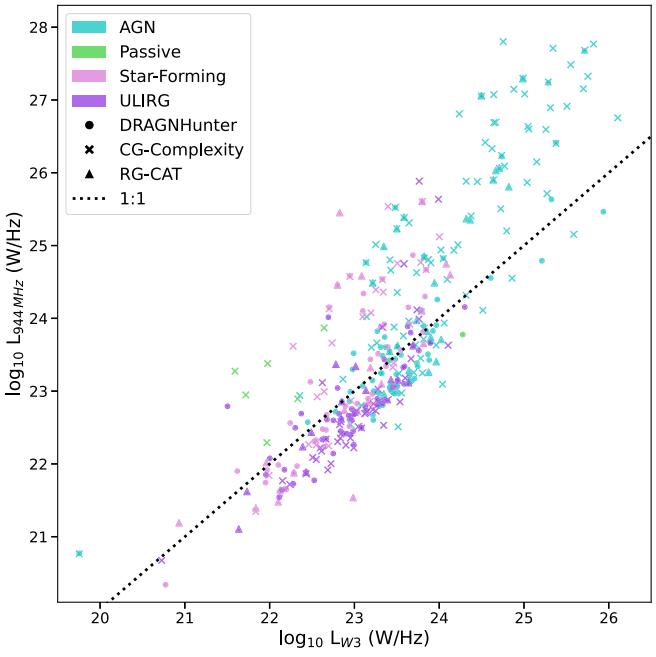


Figure 21. Radio luminosity versus *W*3 luminosity for sources identified by each source-finder, colour–coded by WISE-colour classification (passive, star-forming, ULIRG, AGN). The dashed line shows the 1:1 reference, included as a visual guide only and not representing the true radio–IR correlation (better traced by the distribution of star-forming sources). Most sources lie above the line, indicating a radio excess consistent with AGN-dominated emission, even for hosts classified as star-forming. Only sources with S/N > 3 in *W*1, *W*2, and *W*3 are shown.

We visually inspected sources classified as star-forming or ULIRG using WISE and SDSS images (Figure 22). For many of these sources, the SDSS composite image reveals a resolved/semi-resolved spiral galaxy, with the radio emission likely coming from HII star forming regions in the spiral arms. Interestingly, CG-Complexity has the smallest number of star-forming classifications, despite its capacity to detect irregular emission. It appears for *DH* and RG-CAT, that two or more nearby spiral galaxies, with radio bright cores at their centre, can confuse the algorithms into a false detection of a radio double candidate.

6. Conclusions

The identification of extended radio sources remains a central challenge in large-area surveys, where it can be difficult to identify and characterise the complex morphology of radio sources by conventional source-finding algorithms. In this work, we applied the extended radio source-finders DRAGNHUNTER, coarse-grained complexity, and RG-CAT to EMU observations of the G09 field.

We find that each source-finder identifies a largely independent set of sources, with only 375 detections common to all three catalogues. This demonstrates that the choice of source-finding algorithm can substantially shape the resulting sample and highlights the absence of a single, universally optimal method for complex radio emission.

Despite this limited overlap, the identified sources share comparable physical properties. In particular, DRAGNHUNTER and RG-CAT recover sources of similar angular and linear sizes, flux densities, and luminosities, suggesting that these algorithms are sampling comparable astrophysical populations. Their infrared colour distributions are also nearly identical, indicating that all three methods trace host galaxies with similar underlying characteristics. Notably, a large proportion of the infrared-matched hosts fall outside canonical AGN regions, implying that many of the radio bright sources in EMU-G09 may not be powered by an AGN.

These results collectively show that no single source-finder provides a complete census of extended radio sources in the EMU-G09 field. Each approach contributes complementary detections that, when combined, yield a more comprehensive view of the sky. The similarities in the astrophysical properties across finders suggest that the physical interpretation of the detected sources is robust, even though any individual catalogue will be incomplete. Future work should therefore focus on integrating the strengths of these approaches, potentially through ensemble or hybrid methods, to ensure both completeness and reliability in upcoming large-scale surveys.

Acknowledgement. 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

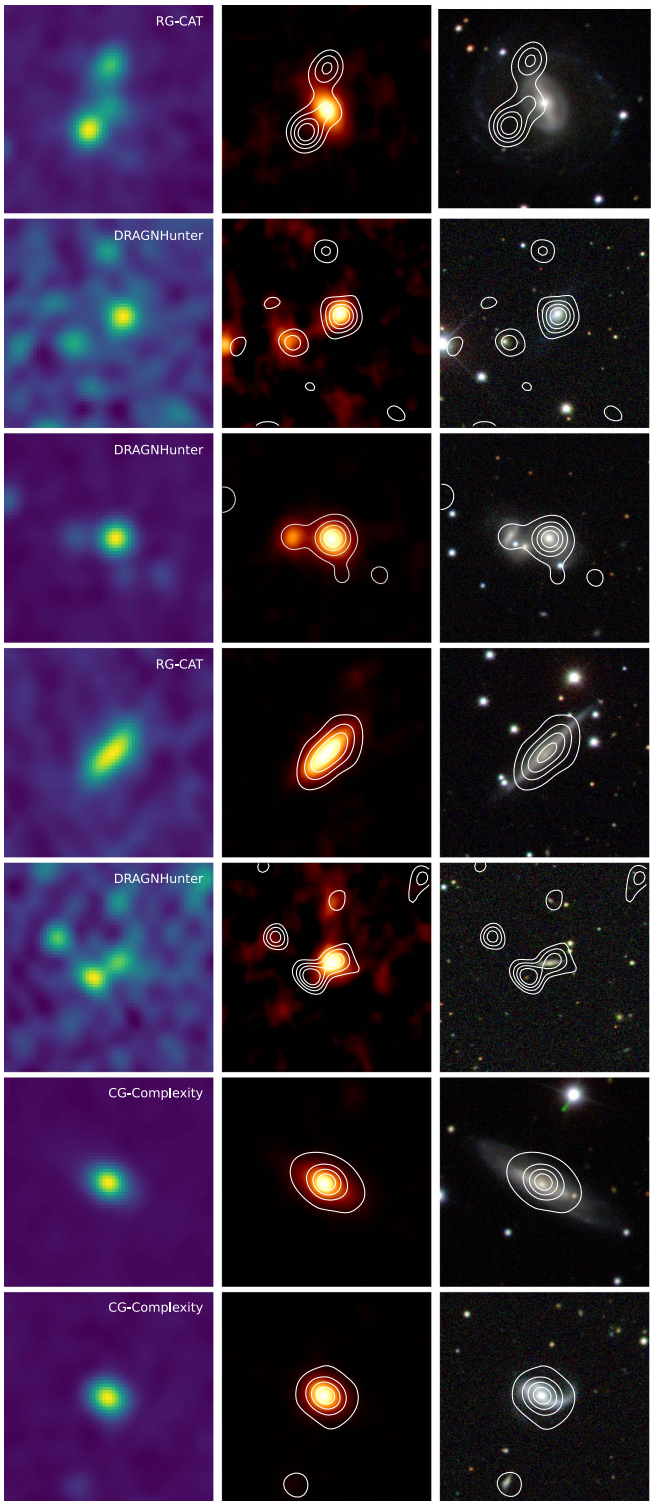


Figure 22. Example sources classified as star-forming or ULIRG based on their WISE colours. For each system, the left panel shows the radio emission with the detecting source-finder indicated in the top right, the middle panel shows the WISE $W3$ (12μm) image, and the right panel shows the corresponding SDSS composite-colour image using the g, r, i bands. Contours in both the WISE and SDSS images correspond to the radio flux. All cutouts have an angular size of $128'' \times 128''$. Many of the WISE star-forming classifications appear to be from a resolved spiral galaxy with the radio emission coming from star-forming regions in the spiral arms.

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).

This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration.

GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programmes including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is https://www.gama-survey.org/.

L.J.B acknowledges the financial support provided by the Macquarie Research Excellence Scholarship (MQRES) program, received throughout the duration of this research.

We thank Denis Leahy and Michael Cowley for providing constructive comments on this paper.

References

Aaronson, S., Carroll, S. M., & Ouellette, L. 2014, arXiv e-prints, arXiv:1405.6903
Alam, A., Pimbblet, K., & Gordon, Y. 2025, PASA, 42, e013
Andernach, H. 2025, arXiv e-prints, arXiv:2509.06132
Banfield, J. K., et al. 2015, MNRAS, 453, 2326
Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559
Beitzel, S. M., Jensen, E. C., & Frieder, O. 2009, in Encyclopedia of Database Systems, ed. L. Liu, & M. T. Özsu (Boston, MA: Springer US), 1691
Bellstedt, S., et al. 2020, MNRAS, 496, 3235
Boyce, M. M., et al. 2023a, PASA, 40, e028
Boyce, M. M., et al. 2023b, PASA, 40, e027
Brienza, M., et al. 2017, A&A, 606, A98
Cluver, M. E., et al. 2020, ApJ, 898, 20
Condon, J. J. 1992, ARA&A, 30, 575
Condon, J. J., et al. 1998, AJ, 115, 1693
Conover, W. J. 1999, Practical Nonparametric Statistics, 3rd Edition, Vol. 350 (New York: John Wiley & Sons)
Cormack, G. V., & Lynam, T. R. 2006, in Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 533
Cutri, R. M., et al. 2013, Explanatory Supplement to the AllWISE Data Release Products, 1 pp. https://ui.adsabs.harvard.edu/abs/2013wise.rept....1C
Cutri, R. M., et al. 2012, VizieR Online Data Catalog: WISE All-Sky Data Release (Cutri+ 2012), VizieR On-line Data Catalog: II/311
Dabhade, P., et al. 2017, MNRAS, 469, 2886
Dabhade, P., et al. 2020, A&A, 635, A5
Driver, S. P., et al. 2011, MNRAS, 413, 971
Driver, S. P., et al. 2016, MNRAS, 455, 3911
Driver, S. P., et al. 2022, MNRAS, 513, 439
Fan, D., Budavári, T., Norris, R. P., & Basu, A. 2020, MNRAS, 498, 565
Fan, D., Budavári, T., Norris, R. P., & Hopkins, A. M. 2015, MNRAS, 451, 1299
Fanaroff, B. L., & Riley, J. M. 1974, MNRAS, 167, 31P
Galvin, T. J., et al. 2020, MNRAS, 497, 2730
Gopal-Krishna, Biermann, P. L., & Wiita, P. J. 2003, ApJ, 594, L103
Gordon, Y. A., et al. 2023, ApJS, 267, 37
Gupta, N., Hayder, Z., Norris, R. P., Huynh, M., & Petersson, L. 2024a, PASA, 41, e001
Gupta, N., et al. 2022, PASA, 39, e051

Gupta, N., et al. 2023, PASA, 40, e044
Gupta, N., et al. 2024b, PASA, 41, e027
Hardcastle, M. J., & Croston, J. H. 2020, NewAR, 88, 101539
Hopkins, A. M., et al. 2013, MNRAS, 430, 2047
Hopkins, A. M., et al. 2015, PASA, 32, e037
Hopkins, A. M., et al. 2025, PASA, 1–32
Hotan, A. W., et al. 2021, PASA, 38, e009
Johnston, S., et al. 2007, PASA, 24, 174
Lacy, M., et al. 2020, PASP, 132, 035001
Laing, R. A., & Bridle, A. H. 2014, MNRAS, 437, 3405
Lake, S. E., et al. 2012, AJ, 143, 7
Leahy, J. P. 1993, Jets in Extragalactic Radio Sources, 421, 1
Levine, J. R. 2012, The 'application/zlib' and 'application/gzip' Media Types, RFC 6713, doi:10.17487/RFC6713
Liske, J., et al. 2015, MNRAS, 452, 2087
Ma, Z., et al. 2019, ApJS, 240, 34
Manning, C., Raghavan, P., & Schuetze, H. 2008, Introduction to Information Retrieval (Cambridge: Cambridge University Press)
Marocco, F., et al. 2021, ApJS, 253, 8
Mingo, B., et al. 2016, MNRAS, 462, 2631
Mingo, B., et al. 2019, MNRAS, 488, 2701
Mohan, N., & Rafferty, D. 2015, PyBDSF: Python Blob Detection and Source Finder, Astrophysics Source Code Library, record ascl:1502.007
Ndung'u, S., Grobler, T., Wijnholds, S. J., Karastoyanova, D., & Azzopardi, G. 2024, MNRAS, 530, 783
Norris, R. P. 2011, in Sixth IEEE International Conference on eScience, 21
Norris, R. P. 2017, NatAs, 1, 671
Norris, R. P., et al. 2011, PASA, 28, 215
Norris, R. P., et al. 2013, PASA, 30, e020
Norris, R. P., et al. 2021, PASA, 38, e046
Norris, R. P., et al. 2025, EMU and the DRAGNs I: A Catalogue of DRAGNs, arXiv:2507.23337
Riggi, S., et al. 2016, MNRAS, 460, 1486
Riggi, S., et al. 2019, PASA, 36, e037
Riggi, S., et al. 2021, MNRAS, 502, 60
Riggi, S., et al. 2023, A&C, 42, 100682
Segal, G., Parkinson, D., Norris, R. P., & Swan, J. 2019, PASP, 131, 108007
Segal, G., et al. 2023, MNRAS, 521, 1429
Shimwell, T. W., et al. 2017, A&A, 598, A104
Shimwell, T. W., et al. 2019, A&A, 622, A1
Shimwell, T. W., et al. 2022, A&A, 659, A1
Siegel, S., & Castellan, N. 1988, Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill International Editions Statistics Series, 2nd Edition (New York: McGraw-Hill)
Smith, G. A., et al. 2004, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 5492, Ground-based Instrumentation for Astronomy, ed. A. F. M. Moorwood, & M. Iye, 410
Sortino, R., et al. 2023, ExA, 56, 293
Taylor, E. N., et al. 2011, MNRAS, 418, 1587
White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D. 1997, ApJ, 475, 479
White, S. V., et al. 2020a, PASA, 37, e018
White, S. V., et al. 2020b, PASA, 37, e017
White, S. V., et al. 2025, PASA, 42, e085
Whiting, M., & Humphreys, B. 2012, PASA, 29, 371
Whiting, M., Voronkov, M., Mitchell, D., & ASKAP Team. 2017, in ASPCS, Vol. 512, Astronomical Data Analysis Software and Systems XXV, ed. N. P. F. Lorente, K. Shortridge, & R. Wayth, 431
Whiting, M. T. 2012, MNRAS, 421, 3242
Wong, O. I., et al. 2025, MNRAS, 536, 3488
Wright, E. L., et al. 2010, AJ, 140, 1868
Wu, C., et al. 2019, MNRAS, 482, 1211

Appendix A. Small pair separation

We find that 210 or the candidate DRAGNHUNTER component pairs have an angular separation $< 15''$ (0.6%), with a minimum component separation of $0.659''$. In Figure A1, we highlight the four sources with the smallest pair separations. These cases are the result of multiple Gaussian components fit to a single island (e.g. a narrow Gaussian on the bright core and a broad Gaussian for the diffuse wing) that are nested within the same island.

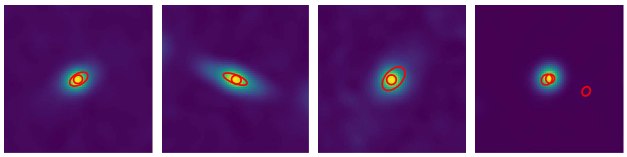


Figure A1. Sources with the lowest component pair separation. Overplotted are ellipses corresponding to the Gaussian components fit by *Selavy*. For each of these sources, the components are nested, therefore producing extremely small component pair separation. All cutouts are 64×64 pixels, corresponding to an angular size of $128'' \times 128''$.

Appendix B. Zero component islands

We find 228 islands that were not fit with a Gaussian component by *Selavy*. To further understand this, we first compare the distribution of signal-to-noise (S/N) of the islands with zero components to the entire sample (Figure B1). The zero component islands tend to have low to moderate S/N, when compared to the rest of the islands.

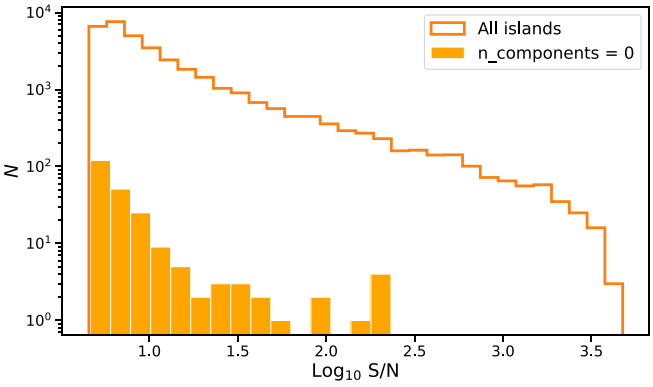


Figure B1. Signal to noise distributions of all islands (dark orange line) and islands with zero fitted components (orange bars).

We then visually inspected the zero component sample. Figure B2 shows a random selection of 16 images of islands with zero Gaussian components. This sample includes several low S/N sources which may prevent *Selavy* from fitting Gaussian components. However, there are some brighter point sources, and some morphologically complex sources present. This suggests that the absence of fitted components does not necessarily indicate poor data quality. We therefore retain the zero-component islands in our analysis. The low number of zero-component sources, relative to the high complexity sample, should not impact our analysis. Their inclusion may also partly explain why the CG-Complexity approach identifies a larger number of sources than both *DH* and RG-CAT.

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

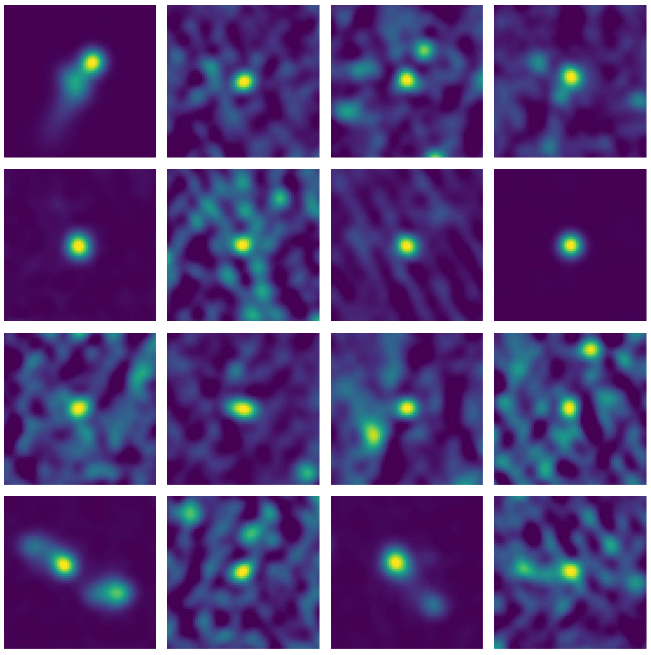


Figure B2. Cutouts of a random selection of 16 islands with zero fitted Gaussian components. All cutouts are 64 × 64 pixels, corresponding to an angular size of 128″ × 128″.

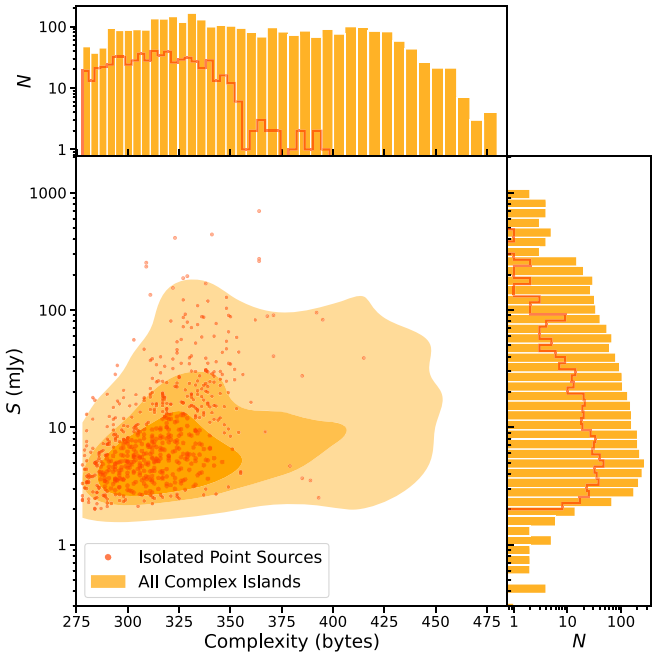


Figure C1. Distributions of complexity and flux density (*S*) for the entire high complexity sample (orange) and isolated point sources within the high complexity sample (dark orange). Density contours of the distributions of complexity and *S* for the full high complexity sample contain 90%, 50%, and 25% of the data points.

Appendix C. 'Complex' point sources

To investigate why point sources appear in the high-complexity subset, we cross-matched islands fitted with a single Gaussian component to the *Selavy* component catalogue in order to obtain their angular sizes. We then defined a threshold below which single-component islands are considered unresolved. The minimum component size among single-component islands is 14.39″. A visual inspection with different thresholds indicated that a threshold of 17″ effectively separates the majority of genuinely extended sources in the high-complexity sample from sources that appear to be unresolved. Sources where the fitted component is below this 17″ are considered point sources. We then performed a nearest-neighbour search on the point source sample using a radius that captures the full cutout size employed in the complexity calculation (see Section 3.3). Sources with no neighbours within this radius are considered isolated point sources. We further explore the isolated point sources by examining their complexity values and flux density, relative to the full highly complex sample in Figure C1.

We find that the isolated point sources generally occupy the lower end of the complexity distribution relative to the full high-complexity sample, while exhibiting a broadly similar flux density distribution. This suggests that their inclusion in the high-complexity subset is not driven purely by extreme brightness. Instead, the results are consistent with a weakly constrained relationship between apparent complexity and flux density, whereby sufficiently bright compact sources, even when isolated, could be deemed as 'complex' by the CG-Complexity measurement. This behaviour may partly reflect the signal-to-noise criterion incorporated in the second GMM used to define the highly complex sample: intrinsically bright sources naturally exhibit high S/N and may therefore be preferentially included, even when their underlying morphology is simple.

If the goal is to minimise contamination from isolated point sources prior to the complexity analysis, the procedure outlined above (component matching, application of a size threshold, and a nearest-neighbour isolation check) can be applied before running the two-stage GMM selection. This would preferentially remove compact, isolated sources that may be retained after the two step GMM filtering. In this work, however, we deliberately retain these sources to preserve a minimally tuned implementation of the CG-Complexity pipeline and to provide a representative assessment of the intrinsic behaviour of each source-finding approach under broadly consistent conditions.

Appendix D. *DH* likelihood ratio calculation

The full description of how the likelihood ratio (LR) is determined in the *DH* script is detailed in Section 4.2 of Gordon et al. (2023). For brevity, we only include the most relevant equations from Gordon et al. (2023). LR is defined as:

$$\mathrm{LR} = \frac{q(W1)f(r)}{n(W1)}, \tag{D1}$$

where $q(W1)$ is probability that the radio source has an AllWISE counterpart with a given magnitude in the WISE $W1$-band, $f(r)$ is the radial separation probability distribution function for the cross match, and $n(W1)$ is the sky distribution of AllWISE sources of a given $W1$-band magnitude. Once LR has been determined for each host, the reliability of any given match, R_i, can then be determined for the match between the radio source and ith AllWISE candidate out of N possible matches:

$$R_i = \frac{LR_i}{\Sigma_{j=1}^{N} LR_j + (1 - Q_0)}, \tag{D2}$$

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press

where Q_0 is an estimate of the fraction of radio sources with an AllWISE match. The *DH* script only adopts an AllWISE host with $R > 0.5$.

Figure D1 shows both the distribution of LR values and the relationship between LR and reliability for all *DH* sources matched with AllWISE. The upper panel highlights that the majority of matches have a low to moderate LR (median of 41.23). This is a significantly lower value than the LR peak of the VLASS data (peaking closer to ∼1 000). This indicates that the cross-matching for the EMU sources is less reliable than the cross-matching for the VLASS sources, likely due to the resolution and sensitivity differences between surveys. We find a mean reliability (R) of 0.9, corresponding to an estimated false positive rate (fpr, calculated by $1 - R$) for *DH*-AllWISE cross-matches of approximately ∼ 10.2%.

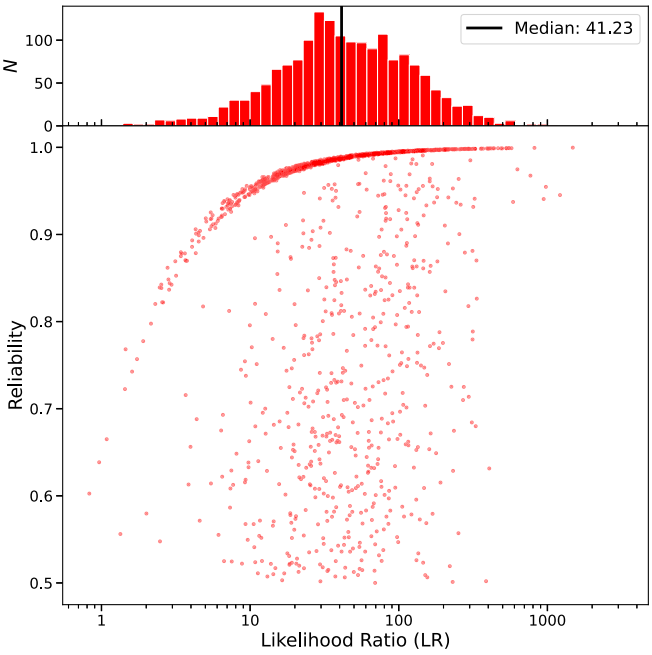


Figure D1. Top panel: Distribution of likelihood ratio values for *DH* sources cross-matched with AllWISE sources. Bottom panel: Likelihood ratio vs the calculated reliability for each cross-match.

https://doi.org/10.1017/pasa.2026.10178 Published online by Cambridge University Press