The coordinated radio and infrared survey for high-mass star formation – V. The CORNISH-South survey and catalogue

T. Irabor,1,* M. G. Hoare,2,* M. Burton,2 W. D. Cotton,3 P. Diamond,4 S. Dougherty,5 S. P. Ellingsen,6 R. Fender,7 G. A. Fuller,4,8 S. Garrington,4 P. F. Goldsmith,9 J. Green,10 A. G. Gunn,4 J. Jackson,11 S. Kurtz,12 S. L. Lumsden,1 J. Marti,13 I. McDonald,14,4 S. Molinari,15 T. J. Moore,16 M. Mutale,1 T. Muxlow,4 T. O’Brien,4 R. D. Oudmaijer,1 R. Paladini,17 J. D. Pandian,18 J. M. Paredes,19 A. M. S. Richards,3 A. Sanchez-Monge,8 R. Spencer,4 M. A. Thompson,12,20 G. Umana,21 J. S. Urquhart,13,22 M. Wieringa10 and A. Zijlstra4

Affiliations are listed at the end of the paper

Accepted 2022 December 21. Received 2022 December 17; in original form 2022 February 21

ABSTRACT
We present the first high spatial resolution radio continuum survey of the southern Galactic plane. The CORNISH project has mapped the region defined by $295° < l < 350°$; $|b| < 1°$ at 5.5 GHz, with a resolution of 2.5 arcsec (FWHM). As with the CORNISH-North survey, this is designed to primarily provide matching radio data to the Spitzer GLIMPSE survey region. The CORNISH-South survey achieved a root mean square noise level of $\sim 0.11$ mJy beam$^{-1}$, using the 6A configuration of the Australia Telescope Compact Array (ATCA). In this paper, we discuss the observations, data processing and measurements of the source properties. Above a $7\sigma$ detection limit, 4701 sources were detected, and their ensemble properties show similar distributions with their northern counterparts. The catalogue is highly reliable and is complete to 90 per cent at a flux density level of 1.1 mJy. We developed a new way of measuring the integrated flux densities and angular sizes of non-Gaussian sources. The catalogue primarily provides positions, flux density measurements, and angular sizes. All sources with IR counterparts at 8 $\mu$m have been visually classified, utilizing additional imaging data from optical, near-IR, mid-IR, far-IR, and sub-millimetre galaxy plane surveys. This has resulted in the detection of 524 H II regions of which 255 are ultra-compact H II regions, 287 planetary nebulae, 79 radio stars, and 6 massive young stellar objects. The rest of the sources are likely to be extragalactic. These data are particularly important in the characterization and population studies of compact ionized sources such as UCHII regions and PNe towards the Galactic mid-plane.

Key words: techniques: image processing – catalogues – surveys – H II regions – radio continuum: ISM.

1 INTRODUCTION
Understanding the formation and evolution of the content of our Galaxy requires studying an unbiased population of objects covering different evolutionary stages utilizing a wide range of wavebands. Cm-wave radio continuum surveys are useful for probing the ionized gas components such as H II regions and planetary nebulae.

The ultra-compact (UC) H II population provides a means to probe the early phases of massive star formation, where the young stars are still deeply embedded in their natal molecular clouds. They are characterized by physical sizes $\lesssim 0.1$ pc, high emission measures $\geq 10^7$ pc$^{-6}$, high electron densities $\geq 10^4$ cm$^{-3}$, and lifetimes typical of 10$^5$ yr (Wood & Churchwell 1989; Comeron & Torra 1996; Kurtz 2000; Kurtz & Franco 2002; Churchwell 2002). Given that these populations often form in clusters, radio and infrared observations at arcsecond resolution are required to resolve their morphologies and provide insight into their immediate environment (Churchwell 1990; Hoare et al. 2007). An unbiased survey of the population of UCHII regions will allow us to test evolutionary models of massive star formation, H II region dynamics, galactic structure, and massive star formation rate in our Galaxy (Churchwell 2002; Hoare et al. 2007; Davies et al. 2011; Urquhart et al. 2013b; Steggles 2016). Radio observations of UCHII regions are also more useful when carried out at a frequency where thermal free-free emission is optically thin ($\geq 5$ GHz; Wood & Churchwell 1989; Churchwell 1990).

IR surveys with high sensitivity and arcsecond resolution of the Galactic plane have made possible studies of an unbiased and statistically representative population of Galactic objects, thereby aiding the studies of massive star formation and stellar evolution. In the northern Galactic plane, these surveys include the the mid-infrared Galactic Legacy Infrared Mid-Plane Survey Extraordinaire (GLIMPSE) by the Spitzer satellite (Churchwell et al. 2009), the mid-infrared inner Galactic plane survey using the Multiband Infrared Photometer for Spitzer (MIPSGAL1; Carey et al. 2009), the far-infrared Herschel

* E-mail: janezoki@yahoo.com (TI); m.g.hoare@leeds.ac.uk (MGH)
1http://mipsig.ipac.caltech.edu/
Infrared Galactic Plane (Hi-GAL) survey (Molinari et al. 2010), the near-infrared Galactic plane survey (GPS) of the United Kingdom Infrared Deep Sky Survey (UKIDSS) project (Lucas et al. 2008), sub-millimetre APEX Telescope Large Area Survey of the Galaxy (ATLASGAL; Schuller et al. 2009) and the $H\alpha$ Isaac Newton Telescope Photometric Survey (IPHAS; Drew et al. 2005).

To complement these surveys, the CORNISH project delivered a uniform and high resolution (1.5 arcsec) radio continuum data set of the northern Galactic plane at 5 GHz. It achieved a sensitivity ~0.43 mJy/beam$^{-1}$ using the VLA in B and BnA configurations (Hoare et al. 2012; Purcell et al. 2013, hereafter Paper I and Paper II, respectively). These data made possible an unbiased census of UCHII regions that is the largest selection to date in the northern-Galactic plane (Kalcheva et al. 2018, hereafter Paper III). Kalcheva (2018) found that 70 per cent of UCHII regions had a cometary morphology using the CORNISH data and follow-up higher resolution radio data. Through a multiwavelength analysis, lbrar et al. (2018, hereafter Paper IV) uncovered an unbiased population of compact planetary nebulae (PNe), of which 7 per cent were newly discovered PNe. A subset of the PNe population has properties that are typical of young sources (Paper IV and Fragkou et al. 2018). The study of such objects is critical, given that the transition window from the post-AGB to the PN phase is small (usually <1000 yr). Other studies that used the CORNISH data in massive star formation studies include Urquhart et al. (2013a), Cesaroni et al. (2015), Tremblay et al. (2015), Yang et al. (2019) and Djordjevic et al. (2019). Note that the Global view on Star formation in the Milky Way (GLOSTAR) Galactic plane survey (Brunthaler et al. 2021), is using the JVLA to go deeper than the CORNISH survey of the northern Galactic plane.

In the southern-Galactic plane, the VST Photometric $H\alpha$ Survey (VPHAS +; Drew et al. 2014) and the Vista Variables in the Via Lactea (VVV) survey (Minnitii et al. 2011) delivered high resolution $H\alpha$ and near-infrared data, respectively, in addition to existing infrared and sub-millimetre data (see Table 1). Surveys of masers at radio wavelengths such as the Methanol Multibeam Survey (MMB; Green et al. 2012, 2017) and the $H_{2}O$ Southern Galactic Plane Survey (HOPS; Walsh et al. 2011) are useful tracers of massive star formation. Existing radio continuum surveys include the Molongolo Galactic Plane Survey (MGPS) (Green 1999; Murphy et al. 2007) that surveyed the 245$^\circ < l < 365^\circ$ and $|b| < 10^\circ$ region at 843 MHz with a resolution of 45 arcsec, the 1.4-GHz Southern Galactic Plane Survey (SGPS) that mapped the 253$^\circ < l < 358^\circ$; $|b| < 1.5^\circ$ region, with a resolution of 100 arcsec and sensitivity of 1 mJy beam$^{-1}$ (McClure-Griffiths et al. 2005; Haverkorn et al. 2006), the GaLactic and Extragalactic All-sky Murchison Widefield Array (GLEAN) survey at 72–231 MHz, with a resolution of 4–2 arcmin that covers all the southern plane (Hurley-Walker et al. 2019) and the TIFR GMRT Sky Survey (TGSS) survey at 150 MHz and 25 arcsec resolution covering the sky above declination ~53$^\circ$ (Intema et al. 2017). These surveys are too low a resolution to resolve UCHII regions and are at frequencies where these objects are optically thick.

### Table 1. Multiwavelength high resolution surveys in the southern Galactic plane.

<table>
<thead>
<tr>
<th>Survey</th>
<th>Bands</th>
<th>Resolution and sensitivity</th>
<th>Coverage</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLIMPSE</td>
<td>3.6, 4.5, 5.8, and 8.0 $\mu$m</td>
<td>&lt;2 arcsec, 0.4 mJy at 8 $\mu$m (5$\sigma$)</td>
<td>295$^\circ &lt; l &lt; 360^\circ$; $</td>
<td>b</td>
</tr>
<tr>
<td>MIPSGAL</td>
<td>24 $\mu$m</td>
<td>~10 arcsec, 1.3 mJy (5$\sigma$)</td>
<td>295$^\circ &lt; l &lt; 350^\circ$; $</td>
<td>b</td>
</tr>
<tr>
<td>VVV</td>
<td>J, H, Ks</td>
<td>~1 arcsec, 16.5 mag in K (5$\sigma$)</td>
<td>295$^\circ &lt; l &lt; 350^\circ$; $</td>
<td>b</td>
</tr>
<tr>
<td>VPHAS +</td>
<td>$H_\alpha$, $r$, $i$</td>
<td>~1 arcsec, 20.56 mag in $H_\alpha$ (5$\sigma$)</td>
<td>+210$^\circ &lt; l &lt; 440^\circ$; $</td>
<td>b</td>
</tr>
<tr>
<td>Hi-Gal</td>
<td>70, 160, 250, 350, and 500 $\mu$m</td>
<td>~6–40 arcsec, ~13–27 mJy (1$\sigma$)</td>
<td>~70$^\circ &lt; l &lt; 68^\circ$; $</td>
<td>b</td>
</tr>
<tr>
<td>ATLASGAL</td>
<td>870 $\mu$m</td>
<td>~18 arcsec, ~50–70 mJy beam$^{-1}$ (1$\sigma$)</td>
<td>280$^\circ &lt; l &lt; 60^\circ$; $</td>
<td>b</td>
</tr>
</tbody>
</table>

In this paper, we present a new radio continuum survey from the CORNISH project of the southern-Galactic Plane. The observations are described in Section 2. The calibration and imaging are discussed in Sections 3 and 4. Data quality and measurement of source properties are presented in Sections 5 and 6 along with a discussion of the catalogue and its reliability. Furthermore, example sources and a comparison of the CORNISH catalogue with other surveys are presented in Section 6.

### 2 OBSERVATIONS

The CORNISH programme observed the southern-Galactic plane with the Australia Telescope Compact Array (ATCA$^3$) using the 2-GHz bandwidth of the Compact Array Broadband Backend (CABB) correlator (Wilson et al. 2011). Observations were carried out with the 6A$^4$ array configuration for about 400 h at two different frequency bands (4.5–6.5 GHz and 8–10 GHz) in full polarization, centred on 5.5 and 9 GHz, simultaneously. This paper focuses on the 4.5–6.5 GHz band. Data reduction of the 8–10 GHz data is still ongoing and will be presented in a future paper (Ibrar et al., In preparation). Observations were between 2010 and 2012 and the observation parameters are summarized in Table 2.

The survey utilized on-the-fly mapping, such that the antennas were scanning continuously, while the phase centres were sequentially moved in a traditional mosaic pattern. This resulted in the doubling of the uv coverage in a single run and a primary beam that is elongated in the scanning direction.$^5$ The phase centres were spaced at 7.4 arcmin in a mosaic pattern that is a scaled version of the hexagonal mosaic implemented in the NVSS survey (Condon

---

$^2$https://www.narrabri.atnf.csiro.au/

$^3$https://www.narrabri.atnf.csiro.au/cgi-bin/obstools/baselines.cgi?array=6a

et al. 1998). This spacing delivers a sensitivity that is uniform to 10 per cent at 5.5 GHz.

The survey region was divided into 33 blocks (33 d of observations) covering 110 deg² of the Galactic plane, defined by 295° < l < 350° and |b| < 1°. The mosaic was scanned in Galactic latitude and required ~18 pointings to cover the 2° of a row (i.e. −1° to +1°). At a scan rate of 7 arcmin/10 s, each row was completed in 3.2 min, including the turnaround time. A secondary calibrator was observed for 2 min after observations of eight rows. A further eight rows were then observed with a secondary calibrator to complete one uv cut of a block of observation, summing up to 54 min. To achieve an optimum uv coverage, this was repeated eleven times at different LSTs, resulting in 1.8 min of on-source time. 12 h were spent on each block, including flux calibration and set-up time. 16 rows correspond to 1.7° in longitude, hence 33 of these 1.7° × 2° blocks were required to cover the survey area. PKS B1934–638 was observed at the end of each block of observations as the primary flux calibrator. PKS B0823–500 was also observed at the beginning of each block of observations as a backup flux calibrator. The secondary calibrators and corresponding days of observation are presented in Table 3.

The observations fall naturally into two epochs (see Fig. 1), based on the observation period (between 2010 and 2012). Fields that were missed or days with observation difficulties due to bad weather or correlator problems were repeated to achieve as full a coverage as possible within the allocated time.

3 DATA REDUCTION PIPELINE

For us to achieve a similar uniform processing as in the CORNISH-North survey, we implemented a similar semi-automated calibration and imaging pipeline (see fig. 1 in Paper II). The pipeline was implemented in PYTHON language using MIRPY6 to directly interface with the Multichannel Image Reconstruction, Image Analysis, and Display (MIRIAD) software packages (Sault, Teuben & Wright 1995).

![Figure 1](https://academic.oup.com/mnras/article/520/1/1073/6979835)

Table 3. Secondary calibrators for each block of observations and the corresponding longitude range.

<table>
<thead>
<tr>
<th>Date</th>
<th>Calibrator (s)</th>
<th>Longitude range</th>
<th>Date</th>
<th>Calibrator (s)</th>
<th>Longitude range</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-12-21</td>
<td>1714–397</td>
<td>348.5–350.0</td>
<td>2011-12-21</td>
<td>1511–55</td>
<td>317.8–319.4</td>
</tr>
<tr>
<td>2010-12-22</td>
<td>1729–37</td>
<td>346.8–348.5</td>
<td>2011-12-22</td>
<td>1352–63</td>
<td>304.1–305.7</td>
</tr>
<tr>
<td>2010-12-23</td>
<td>1729–37</td>
<td>345.1–346.8</td>
<td>2011-12-23</td>
<td>1511–55</td>
<td>316.0–317.6</td>
</tr>
<tr>
<td>2010-12-25</td>
<td>1646–50</td>
<td>341.7–343.4</td>
<td>2011-12-25</td>
<td>1352–63</td>
<td>310.9–312.5</td>
</tr>
<tr>
<td>2010-12-26</td>
<td>1646–50</td>
<td>340.0–341.7</td>
<td>2011-12-26</td>
<td>1352–63</td>
<td>309.2–310.8</td>
</tr>
<tr>
<td>2010-12-27</td>
<td>1646–50</td>
<td>338.3–340.0</td>
<td>2011-12-27</td>
<td>1148–671</td>
<td>297.2–298.8</td>
</tr>
<tr>
<td>2010-12-29</td>
<td>1646–50</td>
<td>334.8–336.5</td>
<td>2011-12-29</td>
<td>1148–671</td>
<td>299.0–304.0</td>
</tr>
<tr>
<td>2010-12-30</td>
<td>1646–50</td>
<td>331.4–333.0</td>
<td>2011-12-30</td>
<td>1352–63</td>
<td>305.8–307.5</td>
</tr>
<tr>
<td>2010-12-31</td>
<td>1646–50</td>
<td>329.7–331.4</td>
<td>2011-12-31</td>
<td>1148–671</td>
<td>302.4–304.0</td>
</tr>
<tr>
<td>2011-01-04</td>
<td>1511–55</td>
<td>324.6–326.3</td>
<td>2012-01-03</td>
<td>1148–671, 1352–63</td>
<td>294.3–295.4</td>
</tr>
<tr>
<td>2011-01-05</td>
<td>1511–55</td>
<td>322.9–324.6</td>
<td>2012-01-04</td>
<td>1352–63</td>
<td>312.6–314.2</td>
</tr>
<tr>
<td>2011-01-06</td>
<td>1511–55</td>
<td>321.2–322.9</td>
<td>2012-01-05</td>
<td>1148–671, 1352–63</td>
<td>301.5–326.2</td>
</tr>
<tr>
<td>2011-12-20</td>
<td>1646–50, 1511–55</td>
<td>333.1–334.7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

in the bandpass at the time of observation, which would lead to false structures in the final images. We performed both automated (PGFLAG) and manual flagging (UVFLAG) on the flux calibrator, secondary calibrators, and science data.

Flagging parameters were determined manually for each block of observations and written to a master configuration file that was applied automatically when the pipeline was run. After an initial flagging of the primary and secondary calibrators, bandpass and flux calibrations were performed. For blocks with two secondary calibrators, the solutions in phase and amplitude were estimated for each calibrator and then combined to produce a global calibration table. A second pass of flagging was then performed on the calibrated data before performing a second and final calibration. This was to ensure that the final calibration was performed on properly flagged data. The final global flagging and calibration tables produced were then applied to the science data and split into individual pointings.

4 IMAGING

The calibrated visibilities of individual fields were imaged by iteratively cleaning down to the maximum residual peak flux (MRF; see Section 4.1) using multifrequency synthesis (mfs; Sault & Conway 1999). This is implemented in MIRIAD using INVERT with the ‘MFS’, ‘SDB’ options and the multifrequency CLEAN MFCLEAN tasks. Multifrequency imaging accounts for the spectral variation across the observation bandwidth of 2 GHz. Thus, the resulting image consists of the normal flux component and the flux times the spectral component (fν), where Sν ∝ να.

The dirty images were created using a robust weighting scheme (Briggs 1995) of 0.5 robustness and gridded to an image pixel size of 0.6 arcsec, ~, which is about one third of the minor axis of the synthesized beam. The robust weighting scheme provides a trade-off between uniform and natural weighting, which is a trade-off between resolution and sensitivity. The choice of 0.5 allows an improved sensitivity without sacrificing the resolution. To ensure uniform resolution, we have forced a restoring circular Gaussian beam with an FWHM of 2.5 arcsec using RESTOR (see Section 5.2).

The residuals were not corrected for the change in resolution, so care should be taken when summing values below the maximum residual flux (see Section 4.1).

4.1 Maximum residual flux

The maximum residual flux (MRF) is an important parameter in the deconvolution process and should be carefully chosen. CLEANing deeper than necessary would result in many artefacts in the form of faint spurious sources and underestimated peak fluxes and flux densities of real sources. This is due to the CLEAN algorithm cleaning noise spikes and redistributing flux from point sources to noise peaks and is known as the clean bias effect (White et al. 1997; Condon et al. 1998). This effect is more pronounced for sources with lower signal-to-noise ratio (SNR), given that the amount of distributed flux is independent of the flux density of the source. According to Prandoni et al. (2000), the clean bias effect can be reduced by setting the MRF above the noise level. However, a shallow CLEAN, where the MRF is too far above the noise level, will create residual images dominated by sidelobes. Therefore, it is important that MRF be carefully chosen, especially for a blind survey such as CORNISH. Unlike for the CORNISH-North survey the MFCLEAN task does not employ windowing. Equation (1) is an expression for the MRF, where f_mrf is a constant multiplicative factor and rms is the root mean square noise level.

\[
\text{MRF} \simeq f_{\text{mrf}} \times \text{rms} \tag{1}
\]

Given that the clean bias is also dependent on the rms noise level, we have determined the rms in equation (1) by imaging the central portion of each field’s dirty map in Stokes V. Stokes V maps usually have very few sources or no sources, since there are few circularly polarized sources at 5.5 GHz (Roberts et al. 1975; Homan & Lister 2006). Hence, the Stokes V maps are dominated by thermal noise.

In order to determine an average and appropriate f_mrf value for the CORNISH-South data, we followed the same procedure as in Paper II. Artificial point sources and sources with simple morphologies were introduced into the uv data of an empty field and the field was imaged with f_mrf values from 0.5 to 5, using our imaging pipeline. The flux densities of these artificial point sources were then measured and compared with the original flux densities. Fig. 2 shows the resulting plot of the fraction of recovered flux density against the f_mrf values. The range of f_mrf values resulted in a >98 per cent recovered flux density. Below f_mrf = 1.5 the field is over CLEANed with a >40 per cent decrease in the rms noise level, compared to the Stokes V map.

To further constrain the choice of an average f_mrf value, for each iteration, the residual images of the sources with simple morphologies were inspected for sidelobe structures. At f_mrf = 1.5, the Stokes I map is CLEANed deeply enough to remove sidelobes and recover >99 per cent of the flux density. Based on these tests, an average value of f_mrf = 1.5 was used in the imaging process. For fields with imaging artefacts as a result of very bright sources (>1 Jy) and/or extended sources, we have reimaged with higher values of f_mrf and manually CLEANed, where necessary. We discuss the procedure for estimating the effect of the clean bias on the CORNISH-South data in Section 5.4.

4.2 Imaging extended sources

The shortest baseline of the ATCA 6A configuration is ~337 m, which corresponds to a spatial scale of ~40 arcsec. Nevertheless,

---

the fraction of recovered flux density as a function of the injected Gaussian FWHM (top panel) and a plot of the measured sizes against the injected FWHM (bottom panel).

For sizes above 17 arcsec, the source was fitted with two Gaussians by AEGEAN; the FWHM plotted in Fig. 3 (bottom panel) is the Gaussian with the larger $\theta_{maj}$ (red line). Our pipeline combines the multiple Gaussians into a polygon and measures the geometric mean as the angular size (see Section 6.1), which is traced by the black line. The maximum FWHM of the fitted Gaussian is $\sim 17$ arcsec (red line) as illustrated in Fig. 3 (bottom panel). However, by switching to a polygon that encloses the emission, we can recover up to $\sim 24$ arcsec (black line). Above 17 arcsec, the corresponding flux density is reduced by $\geq 40$ percent as the flux density steeply drops off. For the real data, we expect that the reduction in flux density and limited size of an extended source should be influenced by morphology as well as rms noise level. Compared to the northern counterpart with an estimated maximum size of $\sim 14$ arcsec (Paper II), we expect to see more extended sources in the CORNISH-South catalogue. Most structures that are less than 15 arcsec should give reasonable estimates of the angular size and the integrated flux density is expected to be within a factor of 2 of the true value.

4.3 Mosaicking

Individually imaged fields were linearly mosaicked onto 20 arcmin × 20 arcmin grid tiles, using LINMOS, a MIRIAD task that overlaps by 1 arcmin). The tiles are arranged in equatorial coordinates (J2000) and 1825 tiles were needed to cover the survey region, extending to the edges of the survey. Tiles covering the edges for $|b| > 1^\circ$ may not be full; i.e. less than 20 arcmin, and some sources at the edges of the survey region may have poor image fidelity due to poor $uv$ coverage. Wide-band primary beam correction, which is the inverse of the primary beam response as a function of the radius and frequency, was performed by LINMOS for each field before linearly mosaicking them to form a tile. LINMOS uses the $\sigma$ $t$ plane and takes into account the OTF scanning during the primary beam correction. In order to improve the accuracy of the primary beam correction, the ‘BW’ option in LINMOS was used to specify the bandwidth of the images (2 GHz).

Linear mosaicking is performed in LINMOS using the standard mosaic equation by minimizing the rms noise (Sault, Staveley-Smith & Brouw 1996). In order to properly account for the geometry and avoid interpolation problems during mosaicking, the overlapping fields were put on the same pixel grid using the ‘OFFSET’ key in INVERT. If this is not taken into account, the position of sources in overlapping tiles will be altered because LINMOS does not automatically account for the geometric correction during linear mosaicking.

5 DATA QUALITY

5.1 Calibrators

The six secondary calibrators along the Galactic plane are shown in Fig. 1. For observation days with two secondary calibrators, calibration was done separately and then the solutions were combined. Before imaging the science data, the calibrators were imaged. This was to inspect the images for artefacts, jets, or anything else that could affect the data. All the secondary calibrators as shown in Fig. 4

\cite{https://www.atnf.csiro.au/computing/software/miriad/doc/lmmos.html}
are point sources with no jets or any other structure \( \geq 5\sigma \) within the field.

B1934–638 was the preferred primary flux calibrator with a flux density of 4.95 Jy at 5.5 GHz. An additional backup flux calibrator (B0823–500) with a flux density of 2.93 Jy at 5.5 GHz was also observed at the beginning of each day’s observation. B0823–500 was used for flux calibration when the B1934–638 data were bad or when it could not be observed due to time constraints e.g. for blocks 28 and 35 (2011-12-30 and 2012-01-07). The flux densities of the secondary calibrators from the ATCA calibrator manual\(^9\) at 5.5 GHz are 1.064 ± 0.007 Jy (1352–63), 1.557 ± 0.007 Jy (1148–671), 2.338 ± 0.009 Jy (1511–55), 2.063 ± 0.018 Jy (1646–50), 1.172 ± 0.006 Jy (1714–397), and 1.73 ± 0.01 Jy (1729–37). Because the calibrators are standard calibrators, MFCAL uses the appropriate flux density variation with frequency during calibration.

With the available data, the flux densities of the secondary calibrators were measured after flagging and calibration. Fig. 5 shows the deviation in percentage from the median flux densities of the six secondary calibrators over the 35 d of observations. Each point in Fig. 5 represents a single day’s observations. Different calibrators were used for the different epochs with the exception of 1511–55 that overlaps both epochs. The secondary calibrators show a percentage flux density deviation from the mean flux density that is less than 10 percent and a standard deviation of 4.0 percent. Based on the scatter in the flux density deviation of the secondary calibrators in Fig. 5, we have adopted a 10 percent calibration error for the CORNISH-South data. The mean positional accuracy of all six calibrators\(^10\) is 0.1 arcsec.


\(^10\)https://www.narrabri.atnf.csiro.au/calibrators/calupdate.html

---

**Figure 4.** Images of the six secondary calibrators 1352–63 (1.064 ± 0.007 Jy), 1148–671 (1.557 ± 0.007 Jy), 1511–55 (2.338 ± 0.009 Jy), 1646–50 (2.063 ± 0.018 Jy), 1714–397 (1.172 ± 0.006 Jy), 1729–37 (1.73 ± 0.01 Jy). The quoted flux densities are from the ATCA calibrator manual. The images are all 1 arcmin by 1 arcmin with the size of the beam shown at the bottom left.

**Figure 5.** Percentage deviation from the median flux density versus block/day number for the six secondary calibrators. Each point represents measurements from a block’s observations.

---

### 5.2 Synthesized beam

Fig. 6 (bottom right) shows the distribution of the unconstrained major and minor axes of the dirty beam. The major axis has a median of \( \sim 2.5\) arcsec and extends up to 5.5 arcsec. The minor axis distribution shows a narrower distribution with a peak about 1.8 arcsec. For uniformity across the survey region, the median value of the major axis distribution (2.5 arcsec) was chosen as the size of a circular restoring beam. This results in a super-resolution (major axis/restoring beam of 2.5 arcsec) distribution presented in Fig. 6 (bottom middle). Ninety per cent (90 per cent) of the fields have super-resolution that is less than 1.3. We also plan to release the calibrated uv dataset so that users can re-image fields with whichever beam they require, as well as convolving the residuals with their chosen beam.

The resulting elongation of the synthesized beam before super-resolution, i.e. the ratio of the major to the minor axis, is shown in Fig. 6 (bottom right). Ninety-six per cent (96 per cent) of the fields have elongation less than 2 with a peak at \( \sim 1.4 \). This means that there are a few fields (<4 per cent) where the major axis is 2 or more times...
greater than the minor axis. The variation of the beam’s major axis across the survey region is presented in Fig. 6. The elongation of the synthesized beam of an East-West array like ATCA is a function of the declination of the field, hence the greater elongation for \( \lambda > 344^\circ \) for more equatorial declinations. This will also explain the fields with major axis \( >3 \) arcsec (4 arcsec > \( \theta_{\text{maj}} > 3 \) arcsec) seen within the longitude region \( >344^\circ \) (Fig. 6: Top panel). Fields with higher major axes (\( >3.5 \) arcsec) in Fig. 6 fall within the longitude 333° to 335° region. Fig. 6 (top panel) also highlights the epochs. Epoch II shows a less elongated beam, with major axes lower than 3 arcsec for ~93 per cent of the fields. The fields with higher major axes within the longitude 333° to 335° region are seen to come from epoch II. This is due to the poor \( uv \)-coverage resulting from less than the typical 12 scans for the block observed on 2011-12-20. However, only a few fields were affected, making up ~6 per cent of the epoch II data.

5.3 Sensitivity/Root Mean Square (rms) noise level

The rms noise level achieved across the survey region in CLEANed Stokes I maps is shown in Fig. 7. The noise level is fairly uniform, having a mean of 0.11 mJy beam\(^{−1}\). The rms noise level around a few very bright source clusters is particularly high as expected. The noisy region between longitude 333° and 335° is seen to reflect the poor \( uv \)-coverage seen in Fig. 6. Fig. 7 (bottom panel) further shows a distribution of the rms noise with an elongated tail that corresponds to regions with poor \( uv \) coverages. The two peaks correspond to the different epochs, where the second peak is dominated by epoch I data. The noise level of epoch II data (hatched grey shaded region) is better than that of epoch I (hatched red shaded region). Given that the two epochs are separated by 11 months, the intervening series of system maintenance and tests\(^{11}\) could have improved the efficiency and overall performance of the array. The region of epoch II data with rms noise >0.11 mJy beam\(^{−1}\) corresponds to fields where there were fewer than the average 12 scans (see Section 2).

of the flux densities compared to the statistical and absolute flux density calibration uncertainty of 10 per cent.

6 CATALOGUE

6.1 Source finding and characterization

The automated source finding and source characterization part of our pipeline utilizes the AEGEAN software package\(^\text{12}\) (Hancock et al. 2012; Hancock, Trott & Hurley-Walker 2018). AEGEAN uses the flood-fill algorithm, where two thresholds (\(\sigma_{\text{f}}\); seeding threshold, \(\sigma_{\text{f}}\); flooding threshold) are defined, such that \(\sigma_{\text{f}} \geq \sigma_{\text{f}}\). The seeding threshold is used to seed an island, while the flooding threshold is used to grow the island (see Hancock et al. 2012). Detected pixels are then grouped into contiguous islands and characterized by fitting with one or more overlapping 2D Gaussians.

We have used the background and noise estimation function, BANE (Hancock et al. 2012) to compute the background and rms noise for each tile. BANE uses a grid algorithm that estimates the rms noise (\(\sigma_{\text{BANE}}\)) and background level within a sliding box of a defined size that is centred on a grid point. The pixel values within the box are then subjected to sigma clipping (3\(\sigma\)), which reduces any effect source pixels may introduce (Hancock et al. 2012; see also Bertin & Arnouts 1996). Given that radio images do not have very complicated backgrounds, we do not expect the noise properties across a 20 arcmin \(\times\) 20 arcmin tile to change much. Therefore, we used the default boxcar size of 30\(\theta_{\text{bin}}\) (\(\theta_{\text{bin}}\): synthesized beam size) for the CORNISH-South data, which is about 75 arcsec. This size has been demonstrated to be optimized for the completeness and reliability of compact sources (Huynh et al. 2012). For source finding, we defined a 4.5\(\sigma\) seeding threshold and 4.0\(\sigma\) flooding threshold to create an initial catalogue of the CORNISH-South data.

6.2 Quality control

6.2.1 Elimination of duplicate sources

With a 60 arcsec overlap of the tiles, sources closer to the edges of the tiles were detected more than once. However, because overlapping regions are formed from the same fields, the difference in position would be a fraction of the synthesized beam size. This will also affect the peak flux, depending on the local rms noise. An extended source fitted with multiple Gaussians could have different parameters from one tile to another because it is closer to the edge on one tile and fully imaged on another tile.

To eliminate such duplicated sources, we searched for sources with similar positions, \(<2.0\) arcsec and similar peak flux, \(\text{Peak}_{\text{max}}/\text{Peak}_{\text{max}} > 0.7\). In addition to both conditions, the distance of each duplicated source, relative to the centre of the tile, was calculated and the source closer to the centre of a tile was retained, over the ones closer to the edges.

6.2.2 Elimination of spurious sources

The choice of a cut-off threshold for a catalogue, in terms of the SNR, is a trade-off between completeness and reliability. A low threshold catalogue, e.g. 3\(\sigma\), will result in more real sources but with many unreal sources as well, while a high threshold will result in a highly reliable catalogue but miss real sources with low SNR. In order to

\(^\text{12}\)https://github.com/PaulHancock/Aegaean/wiki/Quick-Start-Guide
A plot of \( f(\sigma) \) is presented in Fig. 8, assuming the total number of possible detections equals the number of beams within the CORNISH survey region (2.02 × 10^8 beams). With this assumption, the total number of spurious sources is underestimated (blue dashed line). However, the number of sources can be allowed to be a free parameter, resulting in a fit represented by the black line. The black line appears to predict the number of spurious sources at 4.5\( \sigma \), but falls off rather too steeply, compared to the number of spurious sources. Fitting \( f(\sigma) \) to bins higher than 5\( \sigma \), and adjusting the width of the Gaussian (\( \sigma = 0.8\sigma_{\text{gauss}} \)) results in a better fit (green dashed line) and predicts the number of spurious sources to be less than 10 at 7\( \sigma \). Fig. 8 shows that the fraction of spurious sources decreases from 25 to 5 per cent to 1 per cent at 5\( \sigma \), 5.5\( \sigma \), and 6\( \sigma \), respectively. Based on this analysis, the cut-off threshold for a reliable CORNISH-South catalogue to be accepted is set at 7\( \sigma \).

### 6.2.3 Gaussian sources

For compact sources fitted with a single Gaussian by AEGEAN, the integrated flux densities we adopt those given by AEGEAN. However, for the rms noise we re-measure this in an annulus around an aperture (\( \sigma_{\text{ann}} \)). The source aperture was defined by an elliptical aperture which extends to 3\( \sigma \) of the Gaussian major (\( \theta_{\text{maj}} \)) and minor (\( \theta_{\text{min}} \)) axes. An annulus with the same shape as the source aperture of width 15 arcsec, offset at 5 arcsec from the source aperture, was then defined to measure the rms noise and background level. The choice of an annulus offset of 5 arcsec allows an estimation of the background around the immediate locale of the source. An annulus of width 15 arcsec provides a statistically large area in pixels over which to compute the background and rms noise level, compared to the synthesized beam area of 19.7 pixels. This is more local than \( \sigma_{\text{beam}} \), and is consistent with what was used for the CORNISH-North survey. Paper II provides equations and further details on the aperture photometry (see also Paper IV).

### 6.2.4 Extended non-Gaussian sources

Extended non-Gaussian sources were automatically detected by searching for contiguous islands with more than one overlapping Gaussian. A single optimal 2D polygon was then defined using the CONVEX HULL algorithm to trace the outer outline of the Gaussians, enclosing the emission. The assumption is that overlapping Gaussians trace a single extended source. The extent of the generated 2D polygon is strongly affected by the extent of the individual Gaussians. Thus, before generating the polygons, there was the need for manual intervention to trace the outline of the real emission. Such cases account for \( \sim \)10 per cent of non-Gaussian sources.

Given the defined polygons, new intensity weighted centres (\( \alpha_0, \delta_0 \)) and diameters were then determined. The diameter of the 2D polygon, defined by n-sides and n-vertices, was calculated by determining the radius of each vertex to the intensity weighted centre and then estimating the geometric mean of the radii and multiplying by a factor of 2. Fig. 9 shows an example of two sources that were first fitted by multiple Gaussians and the subsequently generated polygons. For these extended sources, aperture photometry was used to measure the source properties within the defined polygon that...
encloses the emission and the rms noise level within the annulus (see Paper II and Paper IV). The geometric mean of the 2D polygon is \( \theta_e = 2 \left( \frac{1}{N} \sum_{i=1}^{N} r_i \right) \), where \( r_1, r_2, \ldots, r_N \) are the radii of the vertices.

Having removed duplicates, spurious sources, and sources \(<7\sigma_o\) (see Section 6.2.2), we then used visual inspection to further eliminate artefacts due to sidelobes that are close to very bright sources. This was aided by comparison with the GLIMPSE (Churchwell et al. 2009) data. As we are primarily interested in a complete sample of UCHII regions, radio sources in areas with artefacts that have clear IR counterparts were retained. Often these artefacts are caused by bright H II regions themselves and we expect other H II regions to be present in such clustered star forming regions. After these eliminations, the final CORNISH-South catalogue has a total of 4701 high-quality sources above \( 7\sigma_o \).

### 6.3 Measurements and uncertainties

In order to create a uniform catalogue that is similar to the CORNISH-North catalogue, we have used the same sets of equations given in Paper II to estimate the properties and associated errors of the sources (see also Condon 1997). For the well-defined and unresolved sources, defined by a single Gaussian fit, the AEGEAN Gaussian fit measurements are the catalogued properties. However, for the catalogued rms noise level, we have remeasured within an annulus around the source for both extended and non-extended sources. This was to create a CORNISH (North and South) catalogue with uniform noise measurements, given that we have implemented sigma clipping to remove sources within the annulus (see Paper II).

The integrated flux density and associated error are given by

\[
S = \frac{A \pi}{4 \ln(2)} \frac{\theta_{maj} \theta_{min}}{\theta_{e}^2} \tag{4}
\]

and

\[
\frac{\sigma^2}{S^2} \approx \frac{\sigma^2}{A^2} + \frac{\theta_{maj}^2}{\theta_{maj}^2 + \theta_{min}^2} \left[ \frac{\sigma^2(\theta_{maj})}{\theta_{maj}^2} + \frac{\sigma^2(\theta_{min})}{\theta_{min}^2} \right] \tag{5}
\]

where \( A \) is the peak amplitude, \( \theta_{maj} \) is the minor axis, \( \theta_{maj} \) is the major axis, and \( \theta_{e} \) is the synthesized beam size. \( \sigma(\theta_{maj}) \) and \( \sigma(\theta_{min}) \) are the errors on the Gaussian fits. The catalogued angular size and associated error is the geometric mean of the major and minor axes, which can be calculated from

\[
\theta_{mean} = \sqrt{\theta_{maj} \theta_{min}} \tag{6}
\]

and

\[
\sigma(\theta_{mean}) = \frac{\theta_{mean}}{2} \sqrt{\frac{\sigma^2(\theta_{maj})}{\theta_{maj}^2} + \frac{\sigma^2(\theta_{min})}{\theta_{min}^2}} \tag{7}
\]

The integrated flux density of the polygonal sources, measured using aperture photometry, is given by

\[
S = \frac{\left( \sum_{i=1}^{N_{pol}} A_i - N_{arc} B \right)}{\theta_{min}} \tag{8}
\]
Figure 10. A plot of the intensity-weighted diameter against the geometric mean diameter for the polygonal sources. The one-to-one line (black) shows that the geometric mean diameter is consistently larger than the intensity weighted diameter. The best-fitting line (green) is given by $y = 0.44x - 0.41$.

$$\sigma_g^2 = \frac{\sigma_0^2 + \pi N_{sky} \sigma_N^2}{2_{sky}}$$

where $\sigma_0^2$ is the variance and $N_{sky}$ is the number of pixels in the annulus. For CORNISH-North, the angular sizes for the polygonal sources were intensity-weighted diameters and were given by

$$d_w = \frac{\sum_i N_{src} A_i}{\sum_i A_i},$$

where $d_w$ is the intensity-weighted diameter and $\sum_i N_{src} A_i$ is the sum of the flux within the defined source aperture. This works well for simple extended sources. However, the sizes of very extended and double-lobed sources could be underestimated by $\geq 50$ per cent, in some cases. Fig. 10 shows a comparison of the intensity-weighted diameters and the geometric mean diameters for the polygonal sources (see Section 6.2.4). The intensity-weighted diameters are consistently smaller compared to the geometric mean diameters. The best-fitting line to the scatter is given by: $y = 0.44x - 0.41$. Based on this, the catalogue size for a polygonal source is the geometric mean. This can be useful in interpreting and rescaling the CORNISH-North sizes for extended sources.

6.4 Completeness

It is important to demonstrate the completeness of our 5.5-GHz catalogue as a function of flux density. In order to quantify this, artificial point sources were injected into the calibrated $uv$ data of nine fairly empty tiles having no imaging artefacts, fairly homogenous noise distribution and from both epoch I and epoch II. The flux densities of the artificial point sources were chosen to be in the range of 0.2 mJy ($\sim 2\sigma$) to 4 mJy ($\sim 40\sigma$). The positions and flux densities were randomly assigned, while avoiding positions of real sources and making sure that the artificial sources do not overlap. In total, 5000 sources were injected into the tiles. The tiles were imaged, and the properties of the injected sources were measured with the same pipeline that produced the catalogue. This procedure was repeated 10 times and then the average values of the measured properties over the 10 iterations were compared to the injected properties.

Fig. 11 shows the completeness level measured by the percentage of detected sources as a function of their injected flux densities. The mean completeness level (per centage) is also shown as a black line with the ‘+’ symbol. The percentage completeness from the graph shows $>90$ per cent for 1.5 mJy and essentially 100 per cent for $>3$ mJy. The completeness for $<0.3$ mJy is 0 per cent because it is below the seeding threshold of 4.5$\sigma$, Table 4 shows the noise level and completeness for the individual tiles at 50 and 90 per cent. Tiles from epoch I with higher noise levels show lower completeness level compared to tiles from epoch II. The stated rms noise level is an average across each tile and so the completeness will be affected by local rms noise surrounding a given source. At 1.4 mJy ($\sim 7\sigma$), the percentage completeness is about 90 per cent for the worst case (Tile 1614). Based on the mean completeness level, for point sources, the CORNISH-South data is 90 per cent complete at 1.1 mJy. The completeness will be worse around very bright sources, but as can be seen from Fig. 7, this represents less than 0.3 per cent of the total area of the survey.

Table 4. Completeness level of the 5.5-GHz CORNISH-South data across nine representative tiles at 50% and 90%.

<table>
<thead>
<tr>
<th>Tile</th>
<th>Epoch</th>
<th>rms (mJy)</th>
<th>50% (mJy)</th>
<th>90% (mJy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>II</td>
<td>0.09</td>
<td>0.63</td>
<td>0.85</td>
</tr>
<tr>
<td>62</td>
<td>II</td>
<td>0.09</td>
<td>0.43</td>
<td>0.64</td>
</tr>
<tr>
<td>81</td>
<td>II</td>
<td>0.09</td>
<td>0.45</td>
<td>0.68</td>
</tr>
<tr>
<td>109</td>
<td>II</td>
<td>0.10</td>
<td>0.63</td>
<td>0.88</td>
</tr>
<tr>
<td>1614</td>
<td>I</td>
<td>0.19</td>
<td>0.81</td>
<td>1.38</td>
</tr>
<tr>
<td>1655</td>
<td>I</td>
<td>0.14</td>
<td>0.80</td>
<td>1.30</td>
</tr>
<tr>
<td>1731</td>
<td>I</td>
<td>0.12</td>
<td>0.72</td>
<td>1.27</td>
</tr>
<tr>
<td>1733</td>
<td>I</td>
<td>0.12</td>
<td>0.51</td>
<td>1.02</td>
</tr>
<tr>
<td>1752</td>
<td>I</td>
<td>0.16</td>
<td>0.42</td>
<td>1.14</td>
</tr>
<tr>
<td>Mean</td>
<td>I &amp; II</td>
<td>0.12</td>
<td>0.60</td>
<td>1.09</td>
</tr>
</tbody>
</table>
6.5 Catalogue ensemble properties

Figs 12 to 14 present the distribution of the ensemble physical properties of the CORNISH-South sources. We identified 4701 sources above the 7σ limit, of which the properties of 608 are measured with a polygon.

6.5.1 Angular size

The catalogued angular size for both the Gaussian and non-Gaussian sources is the geometric mean (see section 6.3). Based on the mean error of the angular sizes, which is ~0.3 arcsec and the size of the restoring beam (2.5 arcsec), resolved sources are defined as sources with angular sizes >2.8 arcsec for the CORNISH-South catalogue. The angular size distribution in Fig. 12 is dominated by unresolved sources (66.3 per cent) and accounts for the obvious peak at ~2.5 arcsec. Resolved sources (1584) account for 33.6 per cent of the catalogue, of which 38 per cent are polygonal sources (608). The distribution of resolved sources is fairly flat out to 30 arcsec after a steep drop from 2.5 to 5 arcsec. This is consistent with the maximum recoverable size (see Fig. 3).

As a characteristic of all interferometric observations, very extended emission will not be properly imaged due to missing information on large scale structures, limited by the shortest baseline in the array. Thus, caution should be applied in interpreting the angular sizes and flux densities of very extended sources (>17 arcsec). This is also demonstrated in Section 4.2.

6.5.2 Galactic latitude and longitude distributions

Figs 13(a) and (b) show the distributions of the CORNISH-South sources in Galactic latitude and longitude, respectively. The coverage of the CORNISH-South survey is complete within the |b| ≤ 1.0 region as shown in Fig. 13(a). The distributions are similar compared to the Galactic distributions of the CORNISH-North catalogue (Paper II).

The latitude and longitude distributions of the resolved sources correspond to the Galactic region traced by high-mass star formation (Urquhart et al. 2009, 2011). Known star formation complexes G333 and G338.398+00.164, can be seen at l = 333° and l = 338°, respectively (Urquhart et al. 2013a). Based on the Galactic distribution of the resolved sources, they are expected to be dominated by H II regions that are concentrated towards the Galactic mid-plane (Urquhart et al. 2013a, Paper III).

6.5.3 Flux density distribution

The integrated flux density and peak flux distribution in Fig. 14 shows similar distributions compared to the CORNISH-North sources (Paper II). The flux density distribution peaks at ~1 mJy, below which the number of sources drops off due to increasing incompleteness (see Section 6.4). For the resolved/extended sources, the flux density distribution peaks at ~3 mJy and gently falls off, extending up to 10^4 mJy. Compared to the CORNISH-North, we have picked up more faint sources as expected due to the sensitivity being two times better.

6.6 Source classification and example sources

Initial classification of the CORNISH-South sources has utilized the availability of comparable high-resolution and high-sensitivity surveys (GLIMPSE, VVV, VPHAS, HI-GAL and ATLASGAL) of the Galactic plane. As with the CORNISH-North survey, one of us (MGH) has

---

Figure 12. Angular size distribution of the 7σ CORNISH-South sources.

Figure 13. Galactic latitude (13a) and longitude (13b) distributions of the 7σ CORNISH-South sources.
Figure 14. Distributions of the integrated flux density of the $7\sigma$ CORNISH-South catalogue. Resolved/extended sources are represented by the hatched regions. Fig. 19 in Paper II shows the distribution of the CORNISH-North.

visually inspected the multiwavelength images of each source on the CORNISH-South website. Examples of multiwavelength images of the main different types of sources are shown in Fig. A1. Using experience gained from classification of sources in the RMS (Red MSX Source) survey (Lumsden et al. 2013), the following visual classification criteria were used.

II regions have strong, and usually extended, mid-IR, far-IR-, and sub-millimetre counterpart emission. The morphology of the IR emission is usually irregular and complex, as well as often being part of a clustered environment. Their radio emission is usually fairly strong and when resolved can have a cometary, shell, or irregular morphology. The mid-IR emission, dominated by polycyclic aromatic hydrocarbons (PAH) emission, arises from just outside the radio emitting region and often reflects the same morphology (Hoare et al. 2007). An initial sub-classification on angular size has been used in the catalogue, pending distance information. All H II regions less than 5 arcsec in size have been labelled as UCHIIIs. This corresponds to the typical 0.1 pc size of UCHIIIs if they were at a typical distance of ∼4 kpc for the more nearby part of the population of UCHIIIs (Paper III). The larger ones were mostly labelled as H II regions if the radio emission was clearly identifiable as part of a single source across the radio and IR bands. If the radio emission was due to over-resolution of a much larger, multiple, and complex source in the IR then it was labelled as a diffuse H II region. A few H II regions were hidden behind large amounts of dust extinction at 8 μm and were labelled as IR-Dark H II regions.

PNe also have strong mid-IR counterparts due to dust and PAH emission (Smith & McLean 2008; Guzman-Ramirez et al. 2014; Cox et al. 2016), but are much fainter at far-IR wavelengths than H II regions, and are usually undetected in sub-millimetre plane surveys. The SEDs of PNe generally peak at ∼24 μm but some young and dense PNe have their peaks extending up to 70 μm and beyond (see Paper IV; Anderson et al. 2012; Urquhart et al. 2013a). A few of the bipolar, Type I, PNe can have more far-IR and sub-millimetre emission, but are still significantly weaker than H II regions. Morphologically they are much simpler than H II regions and are isolated rather than being in clustered, complex environments.

Radio stars are point sources in every waveband they are detected in. They are also isolated sources in the field. Depending on the type of radio star, they can either have blue or red colours in the optical and IR. Very red radio stars like dusty symbiotic stars are difficult to distinguish from unresolved PN without further information (Irabor et al. 2018).

Due to the sensitivity of the CORNISH-South survey, the radio emission from a few known massive young stellar objects (MYSOs) was detected. They share all the IR characteristics of H II regions, but have very weak radio emission compared to H II regions in general and are unresolved or jet-like (Persur et al. 2016). It can be difficult to distinguish weak, unresolved UCHIIIs powered by B3 stars from MYSOs as they have similar radio luminosities (Persur et al. 2016).

Most extragalactic radio sources are undetected in corresponding optical, IR and sub-millimetre Galactic plane surveys used here and therefore they are straightforward to identify. In the radio, they are usually unresolved single sources and we classify them as IR quiet sources in the catalogue. A small number of these may be radio stars that are so distant or obscured as to remain undetected in the optical and IR surveys. A significant number of extragalactic sources show the classic double radio source morphology, and these are classified as ‘Radio Galaxy (Lobe)’ in the catalogue. In some cases, the core of the radio galaxy is also seen and is classified as such. If both lobes, or a lobe and a core, or both lobes and a core are visible, but part of the same extended source in the catalogue, they are referred to as ‘Radio Galaxy (Both)’. It is possible that some normal star-forming galaxies, as opposed to AGN, are detected in CORNISH-South. Some radio sources had very faint, resolved counterparts in the mid-IR, where it was not clear if they are very distant PNe or galaxies. Further studies will be required to rule out a PN classification.

A few sources did not fit in to any of the above categories, either because they are known sources of unusual type, or their origin is currently unknown. These were classified as ‘Other’.

At the time of publication, all sources with detectable flux at 8 μm seen within the same aperture used for radio fluxes have been classified. This should account for the vast majority of Galactic sources. Fig. 15 shows the distribution of classifications of these sources. Resolved sources with infrared counterparts are dominated by PNe and H II regions. We find more UCHII regions compared to the northern counterpart and have also identified six known MYSOs. With a sensitivity of 0.11 mJy beam$^{-1}$, we have also detected PNe with lower radio flux density compared to CORNISH-North. We expect that the vast majority of the sources that remain to be visually classified will be extragalactic and under the IR Quiet or Radio Galaxy categories. The latest classifications will be those on the CORNISH-South website.
Table 5. An excerpt of the version catalogue available online. The flux densities have not been corrected for the clean bias effect.

<table>
<thead>
<tr>
<th>Source Name</th>
<th>α(J2000)</th>
<th>δ(J2000)</th>
<th>S6.7 cm</th>
<th>S6.1 cm</th>
<th>S6.6 cm</th>
<th>S6.8 cm</th>
<th>S6.9 cm</th>
<th>RMS (mJy beam$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G295.175 - 0.5744</td>
<td>11:43:38.84 (0.85)</td>
<td>62.2517.5 (0.44)</td>
<td>11:43:25.27 (0.85)</td>
<td>62.2626.46 (0.02)</td>
<td>62.2626.48 (1.28)</td>
<td>62.2214.9 (0.35)</td>
<td>62.2525.24 (0.03)</td>
<td>11:29.13 (0.30)</td>
</tr>
<tr>
<td>G289.593 - 0.2695</td>
<td>12:02:23.29 (0.96)</td>
<td>62.2626.44 (0.02)</td>
<td>62.2626.48 (1.28)</td>
<td>62.2214.9 (0.35)</td>
<td>62.2525.24 (0.03)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G295.175 - 0.5744</td>
<td>11:43:38.84 (0.85)</td>
<td>62.2517.5 (0.44)</td>
<td>11:43:25.27 (0.85)</td>
<td>62.2626.46 (0.02)</td>
<td>62.2626.48 (1.28)</td>
<td>62.2214.9 (0.35)</td>
<td>62.2525.24 (0.03)</td>
<td>11:29.13 (0.30)</td>
</tr>
</tbody>
</table>

Table 6. Comparison of the RMS 6 cm (4.8-GHz) and the CORNISH-South observation parameters.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>CORNISH-South</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rest frequency (GHz)</td>
<td>5.5</td>
<td>4.8</td>
</tr>
<tr>
<td>Array</td>
<td>6A</td>
<td>6C/6D</td>
</tr>
<tr>
<td>Bandwidth (GHz)</td>
<td>2</td>
<td>0.128</td>
</tr>
<tr>
<td>Synthesised beam</td>
<td>2.5 arcsec</td>
<td>2.5 arcsec</td>
</tr>
<tr>
<td>Typical image rms (mJy beam$^{-1}$)</td>
<td>0.11</td>
<td>0.27</td>
</tr>
<tr>
<td>Image pixel size</td>
<td>0.6 arcsec</td>
<td>0.6 arcsec</td>
</tr>
</tbody>
</table>

6.6.1 Catalogue format

An excerpt of the catalogue is presented in Table 5 and the columns are arranged in the following format: Column (1) – CORNISH Source name (1 + b); Columns (2) and (3) – right ascension (α) and declination (δ) in (J2000) with their associated errors in brackets; Column (4) – Peak flux and associated error in mJy beam$^{-1}$; Column (5) – integrated flux density and associated error in mJy. Because the clean bias is close to the rms noise level and within the errors, the flux densities were not corrected for the clean bias effect. Column (6) – Angular size and associated error in arcsec; Column (7) – Gaussian FWHM major axis and error in arcsec; Column (8) – Gaussian FWHM minor axis and error in arcsec; Column (9) – position angle (E of N) of elliptical Gaussian; Column (10) – local rms noise level, ($\sigma_{\text{rms}}$), in mJy beam$^{-1}$ as measured in the annulus described in Section 6.2; Column (11) – SNR of the source given by Peak flux divided by $\sigma_{\text{rms}}$; Column (12) – Type of source tells if the source is Gaussian fitted (G) or non-Gaussian, in which case a polygon (P) was drawn around the source; Column (12) – This column indicates the classification of the source. Gaussian sources have both the Gaussian fitted sizes and geometric mean sizes in the final catalogue. The final version of the full table is made available on the CORNISH website (http://cornish.leeds.ac.uk/public/index.php).

6.7 Astrometry and flux density quality check

6.7.1 GLIMPSE

In order to check the astrometry of the CORNISH-South catalogue, we cross-matched a subset of the CORNISH-South sources with the GLIMPSE point source catalogue. To avoid mismatches and multiple matches, the CORNISH-South catalogue was limited to classified sources, excluding extended H$\text{I}$ regions, radio-galaxies, and infrared quiet sources. Additionally, sources with angular size $>3$ arcsec were excluded. The cross-match returned 218 sources and Fig. 16 shows the distributions of the offsets in right ascension (α) and declination (δ).

The distribution of the angular offsets in Fig. 16 (top left) peaks at $\sim$0.4 arcsec and steeply falls to 1.5 arcsec before continuing gently out to 3 arcsec. The offset distribution in α is tightly peaked at about 0 arcsec compared to the distribution in δ that has double peaks at 0.5 and $-0.2$ arcsec. The spread in δ offset can be attributed to the distribution of the intrinsic beam along the major axis seen in the epoch I data (see Fig. 6).

6.7.2 The Red MSX Source (RMS) 6 cm ATCA survey

Targeted radio continuum observations were conducted to identify UCHII regions and PNe as part of the Red MSX Source (RMS) survey (Lumsden et al. 2013). The observations were carried out with the ATCA within the $235^\circ < 1 < 350^\circ$ region at 3.6 and 6 cm (Urquhart...
et al. 2007b). Table 6 compares the observational parameters of the RMS 4.8 GHz (6 cm) and the CORNISH-South 5.5 GHz. Both surveys have similar observational properties but the CORNISH-South observing bandwidth of 2 GHz provides better image fidelity compared to the 128 MHz of the RMS survey.

For further checks on the CORNISH-South astrometry and flux densities, the CORNISH-South 5.5-GHz catalogue was cross-matched with the RMS 4.8-GHz catalogue (Urquhart et al. 2007a) within a 5 arcsec radius. For a one-to-one match, the CORNISH-South catalogue was limited to only sources that have been visually classified, excluding diffuse H ii regions and radio galaxies. Fig. 17 (top left) shows the distribution of the angular separation between the CORNISH-South and RMS for 186 radio sources. The distribution shows a tight correlation with a sharp peak at 0.3 arcsec and then a steep fall to ~1.5 arcsec before gently falling off to ~4.3 arcsec.

75 per cent of the cross-matched sources fall within 1.5 arcsec and 94 per cent fall within 3 arcsec. Compared to the distribution of the offset in $\delta$, the offset in $\alpha$ shows a narrower distribution that is strongly peaked at 0 arcsec. Similar offset distribution in $\delta$ is seen in the CORNISH-South-GLIMPSE cross-match that is attributed to the spread in the intrinsic major axis distribution. The mean offset in $\alpha$ and $\delta$ is 0.1 and 0.2 arcsec, respectively. Based on this, and the CORNISH-South-GLIMPSE cross-match, we adopt a positional accuracy of 0.22 ± 0.11 arcsec for the CORNISH-South catalogue. This is also in line with the mean positional accuracy of the secondary calibrators (see Section 5.1).

In Fig. 18, the flux densities of the cross-matched sources are compared. A one-to-one line (black line) shows that the CORNISH-South 5.5-GHz flux densities are higher than the RMS 4.8-GHz flux densities on average. The few sources where the RMS flux densities are higher are found to be H ii regions with larger angular sizes. The two surveys used different ATCA configurations (see Table 6), which are sensitive to different angular sizes. Compared to the shortest baselines of the 6C (153 m) and 6D (77 m) configurations used for the RMS survey (Urquhart et al. 2007b), the shortest baseline of the CORNISH-South 6A configuration (337 m) makes it less sensitive to extended structures. Hence, for optically thin and angularly large H ii regions, the RMS flux densities are expected to be higher, as seen in Fig. 18. However, the better $uv$ coverage of the CORNISH-South recovers more extended emission on some sources.

A best-fitting line (dotted green line) predicts ~24 percent flux density increase for the CORNISH-South counterparts that is more than the typical calibration error of 10 percent. According to Urquhart et al. (2007b), the distribution of spectral indices, ($\alpha$, where $S_\nu \propto \nu^\alpha$) between the 3.6 and 6 cm data is slightly skewed towards positive indices, which suggests optically thick sources in
their sample. This provides a possible explanation for the compact sources where the 5.5 GHz CORNISH-South flux densities are higher than the 4.8 GHz RMS ones.

7 CONCLUSION AND FUTURE WORK

The CORNISH program has successfully mapped the southern Galactic plane at radio wavelengths with unprecedented resolution and sensitivity. We have presented radio continuum ATCA data at 5.5 GHz, covering the 295° < l < 350°; |b| ≤ 1° region of the southern Galactic plane. The resolution of 2.5 arcsec and noise level of 0.11 mJy beam\(^{-1}\) deliver radio data matched to the existing high resolution, multi-wavelength surveys GLIMPSE, VVV and VPHAS + of the southern Galactic plane.

Utilizing the MIRIAD program for data reduction and AEGEAN source finding algorithm, we have identified 4701 sources above 7σ. Data arising from fields with poor uv coverage make up to only 2 per cent of the data set. In addition to several measures undertaken to ensure data quality, visual inspection has also been used to exclude artefacts and hence the data are highly reliable. The survey has a 90 percent completeness level at a flux density of 1.1 mJy. The measured properties show distributions that are similar to that of the CORNISH-North catalogue presented in Paper II.

All sources with potential counterparts at 8 μm (GLIMPSE) have been visually classified. This corresponds to 43 percent of the sources. We have identified 530 H\textsc{ii} regions, of which 257 are UCH\textsc{ii}s. Additionally, we identified 287 PNe and 79 radio stars. The vast majority of the remaining unclassified sources without infrared counterparts are expected to be extragalactic. With the sensitivity of the CORNISH-South survey being two times better than the CORNISH-North counterpart, sources with lower flux densities have been detected, including a few MYSOs. A detailed analysis of the statistical properties of individual catalogues will be presented in future papers. The CORNISH-South survey also carried out simultaneous observations of the entire field at 9 GHz, which will be presented in a separate paper. The 9 GHz data provide higher spatial resolution at reduced sensitivity and will enable an examination of spectral indices of the sources.

The CORNISH-South data are particularly important in the characterization of compact ionized sources towards the Galactic mid-plane, as no radio survey has previously covered the southern Galactic plane in such resolution and sensitivity. Previous surveys of the southern Galactic plane were limited in resolution and coverage, especially within the [b] < 1° region, hence not suitable for studies of compact ionized regions. New radio facilities such as the, ASKAP\(^{13}\), MeerKAT\(^{14}\) (Jonas & MeerKAT Team 2016), and ultimately the SKA\(^{15}\) (Braun et al. 2019) are exploring the southern Galactic plane at greater depth. The CORNISH-South catalogue is well positioned as a very useful resource to characterize the population of radio sources seen in the MEERKAT L-band survey (8 arcsec resolution and 10 mJy beam\(^{-1}\) noise level; Goedhart, In preparation.) and also for follow-up observations with ALMA.\(^{16}\) It will also be useful in combination with existing multiwavelength surveys of the southern Galactic plane to address key questions in stellar formation and evolution. This will allow multiwavelength exploration and statistical studies of compact ionized regions, which can then be compared with population synthesis models. To date, the CORNISH programme has delivered the most sensitive and unbiased compact radio source catalogue towards the southern Galactic mid-plane. Data products in the form of FITS images of individual sources and catalogues are available on the CORNISH-South Website\(^{17}\).

With deeper and larger radio surveys on the way, it is also important that we look at building and improving existing machine-learning models for image classification (e.g. Adhiambo, in preparation). Having used multiwavelength visual classification, the CORNISH sources provide a good training and validation set for machine-learning models that will allow automated classification of radio sources in these upcoming larger and deeper surveys.

ACKNOWLEDGEMENTS

TI acknowledges the support of the Science and Technology Facilities Council of the United Kingdom (STFC) through grant ST/P00041X/1 and the OSAPND scholarship (Nigeria). The work done by PFG was carried out in part at the Jet Propulsion Laboratory, which is operated by the California Institute of Technology under a contract with the National Aeronautics and Space Administration (80NM0018D0004). JM acknowledges financial support from the State research agency, Spanish Ministry of Science and Innovation under grant PID2019-105510GB-C32. JMP acknowledges financial support from the State Agency for Research of the Spanish Ministry of Science and Innovation under grant PID2019-105510GB-C31 and through the Unit of Excellence Marfa de Maeztu 2020–2023 award to the Institute of Cosmos Sciences (CEX2019-000918-M). GAF acknowledges support from the Collaborative Research Centre 956, funded by the Deutsche Forschungsgemeinschaft (DFG) project ID 184018867.

The Australia Telescope Compact Array (ATCA) is part of the Australia Telescope National Facility which is funded by the Australian Government for operation as a National Facility managed by CSIRO. We acknowledge the Gomeroi people as the traditional own-

---

\(^{13}\)https://www.atnf.csiro.au/projects/askap/index.html
\(^{14}\)https://www.sarao.ac.za/gallery/meerkat/
\(^{15}\)https://www.skatelescope.org/
\(^{16}\)https://www.almaobservatory.org/en/home/
\(^{17}\)https://cornish-south.leeds.ac.uk/
ers of the Observatory site. This work made use of Montage, aplypy and astropy python libraries in the batch processing of FITS files.

**DATA AVAILABILITY**

All data (images and catalogues) are described in the text and available on the cornish website at [https://cornish-south.leeds.ac.uk](https://cornish-south.leeds.ac.uk) to download in csv (catalogues), uv files and FITS (images) formats, catalogues are already available. Fits images and uv-data will be available by the end of February 2023.

**REFERENCES**


Briggs D. S., 1995, PhD thesis, New Mexico Institute of Mining and Technology


Carey S. J. et al., 2009, PASP, 121, 76


Churchwell E., 1990, A&ARv, 2, 79


Churchwell E. et al., 2009, PASP, 121, 213


Minniti D. et al., 2011, Bol. Asociacion Argentina de Astron. La Plata Atenta, 54, 265

Molinari S. et al., 2010, PASP, 122, 314


**SUPPLEMENTARY INFORMATION**

Supplementary data are available at *MNRAS* online.

**Table 6.** An excerpt of the 5.5-GHz CORNISH-South catalogue. **Figure A1.** From the left: 5.5-GHz radio (CORNISH), 3-colour GLIMPSE, 70 μm image (Hi-Gal) and ATLASTGAL 850 μm image. **Figure A2.** Continuation from Fig. A1. Images are available online.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors.
Any queries (other than missing material) should be directed to the corresponding author for the article.

**APPENDIX A: EXAMPLE SOURCES IN THE CORNISH-SOUTH CATALOGUE THAT HAVE BEEN CLASSIFIED**

![Figure A1](https://example.com/image.png)

*Figure A1.* From the left: 5.5-GHz radio (CORNISH), 3-colour GLIMPSE, 70 μm image (HI-GAL) and ATLASGAL 850 μm image. From top to bottom: Example H II region, UCHII region, PN, radio star, radio galaxy, and infrared-quiet sources. CORNISH radio images are 80 arcsec by 80 arcsec; GLIMPSE images are 100 arcsec by 100 arcsec; Hi-Gal and ATLASGAL images are 180 arcsec by 180 arcsec. The red polygons/gaussian overlays are that of the 5.5-GHz CORNISH sizes. Images are available online.
Figure A2. Continuation from Fig. A1. Images are available online.

1. Physics and Astronomy, University of Leeds, LS2 9JT, UK
2. Armagh Observatory and Planetarium, College Hill, BT61 9DB, UK
3. The National Radio Astronomy Observatory, Charlottesville, VA 22903, USA
4. Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Manchester, M13 9PL, UK
5. The ALMA headquarters, Santiago, Alonso de Córdova 3107, Chile
6. School of Natural Sciences, College of Sciences and Engineering, University of Tasmania, Hobart 7001, TAS, Australia
7. Department of Physics, University of Oxford, Keble Road, Oxford, OX1 3RH, UK
8. Physikalisches Institut, University of Cologne, Zülpicher Str. 77, 50937 Köln, Germany
9. Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
10. CSIRO Space and Astronomy, PO Box 1130, Bentley WA 6102, Australia
11. Green Bank Observatory, 155 Observatory Rd, PO Box 2, Green Bank, WV 24944, USA
12. Instituto de Radio Astronomy and Astrophysics, National Autonomous University of Mexico, 58089 Morelia, Michoacán, México
13. Departamento de Física (EPSI), Universidad de Jaén, Campus Las Lagunillas s/n, A3, E-23071 Jaén, Spain
14. Department of Physics and Astronomy, Open University, Walton Hall, Milton Keynes, MK7 6AA, UK
15. Istituto Nazionale di Astrofisica – IAPS, Via Fosso del Cavaliere 100, I-00133 Roma, Italy
16. Astrophysics Research Institute, Liverpool John Moores University, Twelve Quays House, Egerton Wharf, CH41 1LD, UK
17. Infrared Processing Center, California Institute of Technology, Pasadena, CA 91125, USA
18. Department of Earth and Space Sciences, Indian Institute of Space Science and Technology, Trivandrum 695547, India
19. Cosmos Science Institute, University of Barcelona, ICCUB, Martí i Franquès, 1, 08028 Barcelona, Spain
20. Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK
21. INAF – Osservatorio Astrofisico di Catania, Via S. Sofia 78, 1-95123, Catania, Italy
22. Centre for Astrophysics and Planetary Science, University of Kent, Canterbury, CT2 7NH, UK

This paper has been typeset from a \LaTeX file prepared by the author.